

LA LOGICA FUZZY

I PARADOSSI DELLE VAGHEZZA (Versione light)

Giangiacoimo Gerla

*Facoltà di Scienze, Università di Salerno,
Via Ponte don Melillo Fisciano (SA).*

e-mail gerla@unisa.it

INDICE

INTRODUZIONI

PRIMA INTRODUZIONE (il contenuto del libro)

SECONDA INTRODUZIONE (invito a leggere)

TERZA INTRODUZIONE (mettere la mani avanti)

CAPITOLO 1

LA VAGHEZZA ED I SUOI PARADOSSI

1. Verità graduate	1
2. Come si riconoscono le nozioni vaghe	4
3. Vaghezza, mucchi di grano ed uomini calvi	7
4. Neanche la scienza è immune dai paradossi della vaghezza (perché la teoria dell'evoluzione è assurda)	9
5. Un vero scienziato deve evitare i concetti vaghi ?	12
6. Ma sembra difficile fare a meno della vaghezza	15
7. La soluzione ai paradossi proposta dalla logica fuzzy	19
8. Altre soluzioni: il finitismo stretto	22
9. Altre soluzioni: la teoria delle supervalutazioni	24
10. Il paradosso del mentitore	27
11. Non solo paradossi: la scienza ingenua	30
12. Probabilità e fuzziness: esiste un rapporto ?	31

CAPITOLO 2

MODELLI MATEMATICI PER LA VAGHEZZA

1. Insiemi fuzzy: un modello matematico per la vaghezza	37
2. Qualche giustificazione per la definizione proposta	40
3. Inclusione, unione, intersezione e complemento	42
4. Norme, conorme e negazioni	46
5. E' l'uso dei connettivi ad imporre la normalizzazione	50
6. Altri connettivi: implicazione, equivalenza, negazione	53
7. Interpretazioni fuzzy	55
8. Semantica vero-funzionale	58
9. Esempi: preordini ed ordini fuzzy	60
10. Esempi: le similarità	62
11. Una dualità tra similarità e distanze	63
12. Esempi: strutture algebriche fuzzy	65
13. Similarità e gruppi fuzzy di trasformazioni	67

14. Quanto è elastica una struttura fuzzy ?	69
15. La nozione di taglio e le supervalutazioni	71
16. Il Cervino di Varzi, oggetti crisp, oggetti vaghi	74
17. Oggetti vaghi, proprietà, gruppi	78
18. Oggetti in matematica e supervalutazionismo	80
19. Il paradosso di Evans e quello della nave di Teseo	84
20. Qualche considerazione finale sugli oggetti vaghi	85

CAPITOLO 3

RAGIONAMENTI APPROSSIMATI

1. La logica come strumento per trattare informazione	87
2. L'operatore di conseguenza logica	89
3. Caratteristiche di una dimostrazione classica	91
4. Dimostrazioni che coinvolgono nozioni vaghe	94
5. Un esempio: risolvere il paradosso dell'indiscernibilità	97
6. Ma meglio indebolire la proprietà transitiva	100
7. Fondere l'informazione fornita da dimostrazioni diverse...	105
8. Oltre l'intervallo dei reali	107
9. Altre forme di informazione: constraints e bireticoli	108
10. Perché il principio di induzione è falso	111
11. La logica come studio degli operatori di chiusura	115
12. Il problema della compattezza	115
13. La compattezza come continuità	118
14. Paradigma della logica fuzzy e logica probabilista	119

CAPITOLO 4

ESEMPI DI APPARATI INFERENZIALI FUZZY

1. L'approccio "graded" alla logica di Łukasiewicz	123
2. Metalogica fuzzy per logiche crisp	125
3. Ogni situazione ha la sua norma: idempotenza	127
4. Ogni situazione ha la sua norma: non idempotenza	130
5. La logica fuzzy è monotona ?	133
6. Metalogica fuzzy: ragionamenti con sinonimia	137
7. La programmazione logica classica	140
8. Programmazione logica fuzzy	144
9. Usare la tecnica dei punti fissi	147
10. Sinonimia e corruzione in programmazione logica	149
11. Modelli di Herbrand	150
12. Il controllo: usare l'esperienza di un praticone ?	154
13. La procedura proposta dal controllo fuzzy	156
14. Logica e modelli di Herbrand per il controllo fuzzy	160

15. Quantità contro qualità nella scienza moderna	164
16. Ritorno alla qualità: il successo del controllo fuzzy	167

CAPITOLO 5

ESISTENZA POTENZIALE, EFFETTIVITA'

1. Esistenza potenziale ed esistenza attuale	171
2. Un esempio: la geometria senza punti	172
3. Un altro esempio: la teoria degli insiemi enormi	176
4. Ultrapotenze per passare dal potenziale all'attuale	178
5. Ma manca un approccio totalmente soddisfacente	180
6. Effettività e vaghezza	182
7. Le macchine di Turing fuzzy	185
8. Ma tali macchine non si dimostrano adeguate	188
9. Effettività come limite di approssimazioni calcolabili	189
10. Una tesi di Church per la computabilità fuzzy	191
11. Sfumare o precisare per ottenere decidibilità	193
12. Conclusioni (?)	195

Indice	197
---------------------	-----

Bibliografia	203
---------------------------	-----

PRIMA INTRODUZIONE

(cosa contiene il libro)

Questo manoscritto deriva dal libro omonimo che può essere acquistato dal sito [ilmiolibro](http://ilmiolibro.kataweb.it/): <http://ilmiolibro.kataweb.it/>. Rispetto al libro mancano solo le maggioranza delle dimostrazioni ed alcuni argomenti di carattere più tecnico.

Il libro riguarda le proprietà che potremmo chiamare “*graduate*”. Con tale termine intendo quelle proprietà che possono essere attribuite ad un oggetto, persona od animale con gradi diversi. Esempi sono le proprietà “*essere veloce*”, “*essere piccolo*”, “*essere alto*”, “*essere giovane*”, Nella letteratura sull’argomento tali proprietà vengono usualmente chiamate “*proprietà vaghe*” e, anche se tale modo di indicarle non mi sembra quello più adeguato, nel libro mi atterrò a questa tradizione.

Il problema del significato delle proprietà vaghe è di natura filosofica, tuttavia questo è un libro di matematica, nel senso che parla di come si possano (eventualmente) trovare modelli matematici capaci di rappresentare in modo quanto più possibile adeguato il fenomeno della vaghezza. Più precisamente vuole essere un libro di logica in quanto i modelli matematici che si propongono hanno lo scopo (ambizioso) di descrivere e riprodurre i processi deduttivi dell’uomo nel momento in cui tali processi coinvolgono nozioni vaghe, cosa che avviene molto più spesso di quanto non si ritenga. Insomma è un libro di logica matematica e, come tutti i libri di logica matematica, è piuttosto rozzo ed impreciso come necessariamente è ogni tentativo di definire una teoria matematica di una cosa tanto complessa come il comportamento razionale dell’uomo.

Il contenuto del libro è la “logica fuzzy” dal punto di vista che a volte viene chiamato *graded* ed in cui la nozione di *ragionamento approssimato* assume un ruolo centrale. Tale logica nasce da un articolo di L. A. Zadeh del 1965 in cui viene proposta la nozione di *sottoinsieme fuzzy* come modello matematico per i concetti vaghi.¹ Successivamente, a partire da

¹ Un precursore, in Italia, della teoria degli insiemi fuzzy è stato Angelo Fadini che ha insegnato presso la Facoltà di Architettura di Napoli e che fin dal 1962 ha proposto un calcolo delle classi legato alle logiche a tre valori [Fadini 1962]. Quando ero appena laureato ho avuto modo di

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.

Da notare comunque che la logica fuzzy non nasce dal nulla poiché si innesta nella tradizione degli studi sulle logiche a più valori iniziata, fin dagli anni venti, all'interno della scuola polacca di logica. Tuttavia, mentre per quanto riguarda l'apparato semantico non esistono sostanziali differenze tra le logiche a più valori e logica fuzzy, per quanto riguarda gli aspetti deduttivi, e quindi di elaborazione dell'informazione, gli scopi e i metodi della logica fuzzy sono completamente nuovi. Infatti, nella tradizione delle logiche a più valori l'apparato inferenziale punta alla generazione di tutte e sole le formule che per qualche motivo si ritengono logicamente valide. Probabilmente tale scelta è stata influenzata dal fatto che nella logica classica del primo ordine è possibile ridurre la nozione di deduzione da ipotesi a quella di formula logicamente valida.² Purtroppo nelle logiche a più valori le cose non sono altrettanto semplici e quindi nasce l'esigenza di focalizzare l'attenzione direttamente sulla deduzione da ipotesi. Più precisamente la logica fuzzy, almeno nel suo approccio "graduato", propone apparati inferenziali capaci di "calcolare", per ogni fissato sottoinsieme fuzzy di ipotesi, il sottoinsieme fuzzy di formule che da esso si possono dedurre. In altri termini, in logica fuzzy sia la nozione di "essere ipotesi" che quella di "essere deducibile" da date ipotesi sono concetti vaghi.

conoscerlo e lo ricordo con piacere per la sua grande apertura e gentilezza e per i suoi ampi interessi culturali.

² Infatti, in base al teorema di compattezza ed a quello di deduzione, data una teoria T ed una formula α , risulta che

$T \models \alpha \Leftrightarrow$ esistono $\alpha_1, \dots, \alpha_n$ in T tali che $\alpha_1 \wedge \dots \wedge \alpha_n \rightarrow \alpha$ è logicamente valida.

SECONDA INTRODUZIONE (superare la giusta diffidenza)

A dispetto della sua semplicità (in realtà grazie a tale semplicità) è un fatto che la logica fuzzy ha avuto un enorme successo, almeno nel mondo delle applicazioni tecnologiche. Sicuramente a favore di tale successo ha giocato il fenomeno “moda” che come tutte le mode ora sembra affievolirsi. Poi ha giocato la capacità propulsiva che sembra avere il tono messianico presso molte persone ed il fascino che hanno molti per tutto ciò che è “alternativo”. Basti pensare al successo editoriale che ha avuto il libro di Kosko “Il fuzzy pensiero” la cui prefazione inizia con

“Un bel giorno ho capito che la scienza non era vera. Non ricordo il giorno ma ricordo il momento. Il dio del ventesimo secolo non era più dio.”

e continua, parlando del libro, con

“L’obiettivo è ... di mostrare all’opera la visione fuzzy del mondo ... per farlo bisogna aver dubitato del dio della scienza e provato un po’ della sua collera”.

Una tale prefazione non poteva non portare un grande successo di pubblico (ed una reazione giustificatamente astiosa della critica). Se a questo stile irritante si aggiunge il fatto che in una larga parte della produzione scientifica dei fuzzisti, almeno nella fase iniziale, si ripresentano ovvi risultati con un nuovo linguaggio, si capisce perché non ci sia mai stata nel mondo scientifico una accoglienza favorevole della logica fuzzy.

Eppure, anche se i toni di molti fuzzisti sono irritanti, ritengo che effettivamente la logica fuzzy porti in sé qualche cosa di interessante sia dal punto di vista tecnologico sia, principalmente, da un punto di vista filosofico. In particolare voglio sottolineare:

1. La possibilità di stabilire un ponte tra modo di procedere di tipo quantitativo e modo di procedere di tipo qualitativo nella scienza e nella tecnologia.
2. La possibilità di formalizzare e di studiare il ruolo dell’apprendimento dall’esperienza (e conseguente attività di *tuning*) senza ridurre tale apprendimento al processo cieco delle reti neurali.

3. Possibili soluzioni di paradossi come quello del mucchio di grano, quello di Poincaré, quello del mentitore quello della nave di Teseo.

4. Spinta ad interpretare il significato delle parole come risultato di un complesso gioco linguistico di “contrattazione” tra gli interlocutori.

TERZA INTRODUZIONE **(mettere la mani avanti)**

Concludo queste introduzioni “mettendo le mani avanti” in varie direzioni.

Per prima cosa preciso che l’ambizione di questo libro non è di dare una completa ed equilibrata esposizione della logica fuzzy (che d’altra parte è ancora ben lontana dall’avere una sistemazione definitiva). Infatti, come dicevo nella prima introduzione, mi riferisco solo al punto di vista “graded” come proposto da Goguen, Pavelka, Novak, ed altri autori. Tale punto di vista, anche se è minoritario nell’ambito delle ricerche sulla logica fuzzy, è più vicino ai miei interessi poiché tenta di riprodurre in qualche modo il fenomeno delle argomentazioni che coinvolgono concetti vaghi. Di fatto il filone più sviluppato e dove esistono risultati matematici più avanzati è quello che fa capo a Peter Hajek e che ha in Italia punte avanzate. Esso è più strettamente legato alla tradizione degli studi delle logiche a più valori ed ai suoi aspetti algebrici ed usa una sorta di traduzione della logica a più valori in quella classica tramite l’introduzione di una infinità di costanti proposizionali. Esposizioni ampie e meditate della logica fuzzy si possono trovare, ad esempio, in [Gottwald 2001] e [Dubois et al. 2007].

Metto poi le mani avanti sul lato del rigore: è chiaro che molte di quelle che io indico nel libro come “dimostrazioni” sono piuttosto “argomentazioni a favore di” in quanto spesso non mi muovo nell’ambito di una teoria formalizzata.

Metto le mani avanti per quello che riguarda le osservazioni di carattere filosofico presenti nel libro. Ho letto (ma non necessariamente capito) sì e no il 10% della letteratura filosofica sull’argomento della vaghezza e questo mi dovrebbe indurre a stare zitto. D’altra parte anche se è discutibile che tutti abbiano il diritto di costruirsi la propria teoria delle equazioni differenziali del secondo ordine, è vero che tutti hanno diritto ad elaborare una propria filosofia sulle cose su cui lavorano.

CAPITOLO 1

LA VAGHEZZA ED I SUOI PARADOSSI³

The communication in the natural language is a cooperative game where in each person the meaning of the fuzzy words may change in a learning process (Hájek).

1. Verità graduate

Come dicevo nell'introduzione, questo libro riguarda le proprietà che potremmo chiamare "*graduate*" cioè quelle proprietà che possono essere attribuite ad un oggetto, persona od animale con gradi diversi. Esempi sono le proprietà "*essere piccolo*", "*essere alto*", "*essere giovane*", Nella letteratura sull'argomento tali proprietà vengono indicate come "proprietà vaghe" ma questo modo di indicarle non mi sembra adeguato. Infatti se vediamo qualche vocabolario italiano troviamo definizioni di "vago" del tipo:

Non ben definito, poco chiaro, poco preciso, incerto, indeterminato ...

Inoltre vengono proposte come frasi tipiche le seguenti:

un'informazione vaga, un vago indizio, un vago sospetto, rispondere in modo vago, avere una vaga speranza, notare una vaga somiglianza.

Mentre sui vocabolari inglesi troviamo definizioni del tipo:

Unclear by virtue of being poorly expressed, not coherent in meaning, not clearly or explicitly stated or expressed.

e frasi del tipo:

the conservative manifesto is a model of vagueness; these terms were used with a vagueness that suggested little or no thought about what each might convey.

Come si vede si tende a dare al termine un sapore negativo spesso inerente ad una voluta mancanza di precisione da parte di chi parla. Una "risposta vaga" significa una risposta reticente, in cui

³ Chi fosse interessato solo agli aspetti matematici della logica fuzzy può passare direttamente al secondo capitolo.

si cerca di dire il meno possibile anche se magari si evita di mentire.

Ora, se è vero che esiste una ambiguità di interpretazione in proprietà graduate come “piccolo”, “alto”, “giovane”, è anche vero che tale ambiguità può riguardare anche altre espressioni linguistiche. Ad esempio se dico una frase del tipo

“ieri ho incontrato Carlo Esposito”

e nel luogo in cui vivo esistono molte persone che hanno questo nome e cognome, allora tale frase essendo “non ben definita, poco chiara, poco precisa”, rientra nella definizione data di vaghezza. La stessa cosa può essere detta per una frase del tipo

“passerò dal salumiere tra le 10 e le 11”.

Tentiamo ora di precisare che cosa invece intendiamo in questo libro per “vago” ed a tale scopo analizziamo il ruolo delle proprietà vaghe all’interno dei meccanismi di comunicazione tra persone. Supponiamo ad esempio che A e B parlino al telefono ed A dica:

“Mario portava un cappello”

B capisce una tale frase poiché i termini “Mario”, “cappello” e “portare” denotano per B le stesse cose che per A. Possiamo anche riscrivere in modo più preciso quanto detto da A:

Esiste un oggetto x tale che x è un “cappello” ed x è in relazione “essere sulla testa di” con l’uomo che entrambi conosciamo sotto il nome di “Mario”.

Inoltre, seguendo i suggerimenti di Tarski, possiamo interpretare tale affermazione in modo rigoroso interpretando:

- la parola “cappello” tramite l’insieme dei cappelli,
- l’ espressione “essere sulla testa di” tramite una opportuna relazione binaria, cioè tramite un opportuno insieme di coppie,
- il nome “Mario” con un particolare elemento del dominio delle persone.

Supponiamo ora che A dica:

“Mario portava un grande cappello”.

La comunicazione riesce lo stesso, viene trasmessa informazione che ancora può essere espressa nella forma:

Esiste un oggetto x tale che x è un “cappello grande” ed x è in relazione “essere sulla testa di” con l’uomo che entrambi conosciamo sotto il nome di “Mario”.

La comunicazione riesce nonostante il fatto che in questo caso il termine “cappello grande” abbia probabilmente significati

diversi per A e B , oppure abbia un significato che può evolversi e precisarsi durante la conversazione. In questo caso è evidente che Tarski con tutto il suo bagaglio insiemistico non risulta sufficiente.⁴ Tuttavia questo esempio suggerisce una prima indicazione di che cosa si debba intendere per nozione vaga.

Definizione (euristica). Una nozione vaga è esattamente il tipo di informazione che viene trasmessa da A a B tramite l'espressione "grande cappello".

Inutile dire che questa non è una definizione matematica ma è di tipo "ostensivo" nel senso che indica in qualche modo "la cosa" che si vuole studiare e formalizzare. D'altra parte che il fenomeno della vaghezza esista ed abbia un ruolo importante è indiscutibile come è indiscutibile che in una frase del tipo "*Mario portava un grande cappello*" venga trasmessa informazione. Se ad esempio si chiede di tradurre l'espressione "grande cappello" dall'italiano all'inglese non si esita a tradurla in "big hat". Tale traduzione viene considerata esatta nel senso che si ritiene che le due espressioni abbiano lo stesso significato. Che cosa è questo significato? Una cosa è certa: se la logica matematica non riesce a "catturarlo" questo significato, questo è un limite della logica matematica e non certo della vaghezza.

Come vedremo nel seguito i fuzzisti identificano i concetti vaghi con oggetti matematici chiamati "insiemi fuzzy" e che in un certo senso sono l'estensione di tali concetti. Questa soluzione fornisce risultati interessanti ed è quella su cui si basa questo libro. Tuttavia essa appare fortemente discutibile da un punto di vista filosofico. Sono infatti molto più vicino all'idea che la vaghezza si spieghi all'interno di giochi linguistici in cui essa si rende necessaria alla massimizzazione di un qualche "guadagno". Una interessante analisi del significato "pragmatico" dei concetti vaghi nell'ambito delle conversazioni può essere trovato in [Jucker et al. 2003].

⁴ D'altra parte solo ad un matematico può apparire naturale interpretare "cappello" tramite l'insieme dei cappelli ed "essere sulla testa di" con un insieme di coppie: quando si fabbrica un nuovo cappello e quindi cambia l'insieme dei cappelli cambia anche l'interpretazione? Tuttavia i matematici, che come è noto non sono molto pignoli, si accontentano di un tale modo di interpretare le cose.

2. Come si riconoscono le nozioni vaghe

Abbiamo già proposto una definizione “ostensiva” di nozione vaga. Nel seguito evidenziamo semplici metodi per distinguere una proprietà vaga da una che non lo è. E’ possibile assumere l’insieme di tali metodi come una sorta di definizione operativa della nozione di vaghezza.

Il metodo dei modificatori linguistici:

Una proprietà è vaga se si può ragionevolmente modificare con espressioni del tipo “un po’”, “quasi”, “molto”, “abbastanza” (che a volte vengono chiamate “*modificatori linguistici*” e che rientrano nella tradizione degli operatori modali).

Infatti appaiono sensate frasi del tipo:

Napoli e Salerno sono abbastanza vicine, Maria è un po' chiacchierona, questa casa è molto cara, è preferibile nettamente la cioccolata al latte rispetto alla fondente.

Questo perché i predicati “vicino”, “chiacchierone”, “caro”, “preferibile”, sono vaghi. Se si considerano invece le frasi

Questo numero è un po' pari, Questo numero è molto pari, si nota immediatamente una ridicola stonatura. Questo proprio perché la nozione di essere pari non è vaga. I seguenti esempi pongono in risalto l'effetto comico dell'applicazione di un modificatore linguistico ad un predicato che non è vago:

Il mio fidanzato ha un difetto . . .

- è un po' sposato.

La mia fidanzata ha un problema . . .

- è un po' incinta.

Come è andato l'esame ?

mah! . . . in effetti sono stato leggermente bocciato.

Dottore come è andata l'operazione ?

- discretamente ma mi è sembrato che alla fine il paziente fosse un po' morto . . .

Il metodo delle leggi logiche. Un proprietà è vaga se non verifica le usuali leggi della logica.

Ad esempio una delle leggi della logica classica (*legge di non contraddizione*) afferma che non può valere una asserzione e allo stesso tempo la sua negazione, in simboli, l'asserzione $A \wedge \neg A$ è falsa qualunque sia il valore di verità di A . Se si considera la frase

Il numero n è pari e non è pari

come ci suggerisce la logica classica essa risulta falsa qualsiasi sia il numero n . Invece quando vengono coinvolte nozioni vaghe, asserzioni di tale tipo appaiono ragionevoli. Se chiedo:

- *la casa che hai comprato è grande ?*

una risposta potrebbe essere:

- *beh!, ... è grande e non è grande*

e tale risposta non è affatto contraddittoria. Piuttosto è un modo per esprimere il fatto che il proprio giudizio circa la grandezza della casa si situa in una posizione intermedia. Similmente nella logica classica risulta che $A \vee \neg A$ è vera qualunque sia il valore di verità di A (*legge del terzo escluso*). E' evidente che nel caso in cui A sia una proprietà vaga una tale legge non vale. Ad esempio, mentre risulta vero che "*un numero o è primo oppure non lo è*" non può essere detto lo stesso per l'affermazione "*uno studente o è bravo o non lo è*" affermazione che farebbe solo un insegnante incapace di vedere la complessità e la gradualità del processo di maturazione di uno studente.

Il metodo della comparazione. Una proprietà è vaga se si può affermare che un oggetto (o persona) verifica tale proprietà più di un altro. In altre parole se può essere la base per un giudizio comparativo.

Ad esempio, è sensato affermare che "*Maria è più bella di Luisa*" oppure "*ogni giorno Maria è un po' più bella*". Non ha invece senso dire che "*5 è più dispari di 7*". Ancora una volta coinvolgere in un giudizio comparativo predicati non vaghi porta ad ottenere frasi che hanno un sapore ironico:

"Mario è un po' più sposato di Luigi", "Mario è più napoletano di Luigi", "tutti gli uomini sono uguali davanti alla legge ma alcuni sono più uguali degli altri".

Addirittura un tale coinvolgimento di un predicato non vago in un giudizio comparativo cambia la sua natura trasformandolo in uno vago. Così "*sposato*" non è più l'espressione dello stato SI-NO

che risulta all'anagrafe, denota invece una modalità di vita che ha a che fare con proprietà vaghe come "non essere troppo libero di uscire da solo", "avere impedimenti a cercare avventure", "avere preoccupazioni" ... e così via. La parola "*napoletano*" non rappresenta il fatto storico di essere nato a Napoli, invece rappresenta un modo di essere che ha a che fare con proprietà vaghe come "essere allegro", "non essere ligio alle regole", "essere furbo", "essere imbroglione", ... e così via.

Il metodo dei casi indeterminati. Una proprietà è vaga se esistono casi limite (borderline) per quello che riguarda il partecipare o meno di tale proprietà.

Supponiamo che in una classe un professore dica

"tutti i ragazzi alti alzino la mano".

E' probabile allora di assistere al fenomeno di alcune mani che oscillano tra alto e basso incerte su cosa fare. Un po' questo avviene in quanto non è chiaro che cosa intenda il professore per alto, un po' in quanto potrebbero esistere ragazzi per i quali in ogni caso sarebbe discutibile sia una collocazione tra gli alti, sia una collocazione tra i non alti. Supponiamo che invece il professore chieda

"tutti i ragazzi che hanno portato il vocabolario alzino la mano". La classe si ripartirà senza nessuna esitazione tra studenti con la mano alzata e studenti con la mano abbassata. Evidentemente ciò accade perché "essere alto" è una proprietà vaga mentre "aver portato il vocabolario" non lo è.

Stranamente l'identificazione della vaghezza con l'esistenza di casi borderline è la più diffusa presso i filosofi. Tuttavia a me sembra essere il modo meno significativo per caratterizzare la vaghezza. Infatti, supponiamo che nella classe due studenti siano nettamente bassi, che uno sia altissimo e che i rimanenti siano di altezza variabile ma indiscutibilmente alti. In questo caso i ragazzi non avrebbero esitazioni nella decisione se alzare la mano o meno e quindi verrebbero a mancare casi borderline. Se allora si applicasse il criterio dei casi borderline la nozione di "alto" nel contesto della classe non dovrebbe essere considerata vaga. Tuttavia è un fatto che anche in tale esempio continua ad essere possibile l'applicazione di modificatori linguistici e quindi la possibilità di considerare predicati come "molto alto",

“decisamente non alto”. Sono anche possibili giudizi comparativi e dire di un ragazzo che è più alto di un altro. Pertanto mi sembra ragionevole considerare vago tale predicato anche nel contesto della classe presa in considerazione.⁵

3. Vaghezza, mucchi di grano ed uomini calvi

I predicati vaghi creano molti problemi e forse questo è uno dei motivi per cui si è tentato sempre di eliminarli dal discorso scientifico. Per fare un esempio, esponiamo un famoso paradosso detto “del mucchio di grano” che coinvolge la nozione vaga di *piccolo*. Tale paradosso fu proposto da Eubulide di Mileto, un filosofo greco della seconda metà del IV secolo, e consiste nel provare il seguente teorema.

Teorema 3.1. 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. $\square$

Per analizzare meglio tale paradosso proviamo a formalizzare il ragionamento ora esposto. Dato un numero naturale n indichiamo con $\underline{n}$ un numerale che lo rappresenti, inoltre indichiamo con $Pic(\underline{n})$ l'asserzione “un mucchio con n chicchi è piccolo”. Allora possiamo riscrivere le due assunzioni a) e b) come segue:

- a) $Pic(1)$
- b) $Pic(\underline{n}) \Rightarrow Pic(\underline{n+1})$,

⁵Anticipando la definizione di insieme fuzzy che introdurremo nel prossimo capitolo, un tale esempio corrisponde ad un insieme fuzzy s che assume valori di verità o prossimi ad 1 oppure prossimi a 0 e comunque lontani dal valore 0.5. In tale caso, dovendo decidere se un elemento x appartiene o meno ad s , è ragionevole rispondere “sì” se $s(x) > 0.5$, “no” se $s(x) < 0.5$.

dove, per evitare di coinvolgere inutilmente la logica del primo ordine, intendiamo b) come *schema di assioma*, cioè come l'insieme infinito di formule che si ottengono fissando in tutti i modi possibili il numerale $\underline{n}$. Utilizziamo poi come regola di inferenza il *Modus Ponens*, che indicheremo in breve con *MP*. Come è noto, tale regola permette di inferire da due premesse del tipo α e $\alpha \Rightarrow \beta$ la conclusione β . Fissiamo allora un qualunque numero m e proviamo che vale $Pic(\underline{m})$. Infatti:

- a partire dalla coppia di formule $Pic(1)$ e $Pic(1) \Rightarrow Pic(2)$ ed applicando *MP*, si ottiene $Pic(2)$.
- a partire da $Pic(2)$ e $Pic(2) \Rightarrow Pic(3)$, applicando una seconda volta *MP*, si ottiene $Pic(3)$,
- ...
- a partire da $Pic(\underline{m-1})$ e $Pic(\underline{m-1}) \Rightarrow Pic(\underline{m})$, applicando una $(m-1)$ -esima volta *MP*, si ottiene $Pic(\underline{m})$.

Possiamo rappresentare il ragionamento fatto con il seguente schema:

$$\begin{array}{c}
 \frac{Pic(1) ; Pic(1) \Rightarrow Pic(2)}{Pic(2)} \\
 \frac{Pic(2) ; Pic(2) \Rightarrow Pic(3)}{Pic(3)} \\
 \dots \\
 \frac{Pic(\underline{m-1}) ; Pic(\underline{m-1}) \Rightarrow Pic(\underline{m})}{Pic(\underline{m})}
 \end{array}$$

Con tale formalizzazione possiamo riscrivere il Teorema 3.1 al modo seguente.

Teorema 3.2. Assumiamo a) e b), allora per ogni modo di fissare m esiste una dimostrazione di $Pic(\underline{m})$. Tale dimostrazione utilizza $m-1$ volte la regola *MP*.

Un paradosso simile è quello dell'*uomo calvo*, dovuto ad Euclide di Megara. Per "*capelluto*" intendiamo un uomo con molti capelli.

Teorema 3.3. Un uomo con nessun capello è capelluto.

Dim. Indichiamo con $C(\underline{n})$ l'affermazione per cui un uomo con n capelli è capelluto e consideriamo le asserzioni

a) $C(10.000)$

b) $C(n) \Rightarrow C(n-1)$.

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

Si noti che mentre abbiamo presentato il paradosso del mucchio di grano tramite un ragionamento "*dal basso all'alto*" (da $Pic(1)$ si giunge a $Pic(\underline{m})$), per quello dell'uomo calvo invece abbiamo esposto un ragionamento "*dall'alto verso il basso*" (da $C(\underline{m})$ si giunge a $C(0)$). I due tipi di ragionamento, in un certo senso, sono uno riconducibile all'altro. Ad esempio, indicando con $NPic$ la negazione di "essere un mucchio piccolo", si sarebbe potuto partire da $NPic(\underline{m})$ e provare $NPic(1)$. Infatti il fatto stesso che $Pic(\underline{m})$ appaia assurdo significa che posso ragionevolmente accettare l'assunzione $NPic(\underline{m})$, inoltre le implicazioni *b*) possono essere riscritte nella forma $NPic(n) \Rightarrow NPic(n-1)$. È evidente allora che ciò consente di provare $NPic(1)$, e quindi, nel nostro caso, giungere alla paradossale conclusione che un mucchio con un solo chicco non è piccolo. Tuttavia, come vedremo nel seguito, i due modi di ragionare per certi aspetti sono diversi tra loro.

4. Neanche la scienza è immune dai paradossi della vaghezza

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

Teoria dell'evoluzione. Per prima cosa proviamo che la teoria dell'evoluzione della specie umana è sbagliata.⁶ 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 indiscutibile in quanto immediata conseguenza del significato che si è dato al predicato

⁶ Tratto da [Gerla 2010] mentre una forma non molto diversa di tale paradosso è esposta in [Sanford 1975].

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 alla madre di x . Infatti è una caratteristica della specie umana quella di non potere procreare prima dei dieci anni.

Proposizione 4.1. 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 4.2. Ho avuto antenati extraterrestri, pertanto la teoria dell'evoluzione della specie umana di Darwin è sbagliata.

Dim. Dalla proposizione 4.1 segue che io ho avuto un antenato vissuto almeno $10 \times 10^{100.000}$ anni fa. Conseguentemente, poiché la terra $10 \times 10^{100.000}$ anni fa non esisteva ancora, $10 \times 10^{100.000}$ anni fa tale antenato era extraterrestre. $\square$

Naturalmente non è una tale conclusione ad essere paradossale: anzi, i creazionisti sarebbero contenti di confutare Darwin. Quello che è paradossale è il fatto che si possa giungere a tale conclusione per via puramente logica (logica classica, ovviamente) senza alcun riferimento ad osservazioni di fatto.

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:

Teorema 4.3. La meccanica classica è una teoria in cui (quasi) tutti i teoremi sono errati. Tuttavia è una teoria verificata dall'esperienza quotidiana.

Che la teoria classica, che indicheremo in breve con 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 piccole⁷ le proposizioni provate dalla meccanica relativistica coincidono (approssimativamente) con quelle provate nella meccanica classica. Se accettiamo come valida, come attualmente viene fatto, la meccanica relativistica, allora possiamo anche dire che per velocità molto piccole la meccanica classica dice cose vere. 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 $\alpha_R = "g = 0.567662220"$.

Allora è presumibile che

- in TC sia dimostrabile un'asserzione del tipo

$$\alpha_C = "g = 0.567662221".$$

D'altra parte essendo $g=g$ un assioma logico, dalle due asserzioni α_R e α_C si ricava la contraddizione $(g=g) \wedge \neg(g=g)$. Pertanto 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 ritrovare i paradossi che nascono dalle nozioni vaghe. Ad esempio si può immaginare una successione finita di riferimenti $R_1, R_2, \dots, R_n$ 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, poiché il

⁷ Ancora una nozione vaga.

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. In altre parole il paradosso del mucchio di grano può essere riprodotto nell'ambito del rapporto tra meccanica classica e quella relativistica.

In conclusione sembra che anche per la meccanica e la teoria dell'evoluzione, due paradigmi della razionalità scientifica, nascano problemi logici dal momento che non è possibile per tali teorie evitare il coinvolgimento di predicati vaghi.

5. Un vero scienziato deve evitare i concetti vaghi ?

Ritorniamo ad esaminare il paradosso del mucchio di grano. Se lo si vuole bloccare esistono tre possibilità:

- a) si può mettere in discussione la correttezza del ragionamento
- b) si possono mettere in discussione le premesse del ragionamento
- c) si può dichiarare che la presenza di una nozione vaga rende insensato ogni tipo di ragionamento e che quindi i concetti vaghi vanno espulsi dal discorso razionale.

Ora la possibilità a) non sembra molto credibile. La formalizzazione del ragionamento che conduce al paradosso ne mostra la struttura estremamente semplice. Si tratta di una lunga, ripetitiva catena di applicazioni del MP. Per quanto riguarda la possibilità b), non si può certo negare la validità di $Pic(1)$, cioè, che un mucchio con un solo chicco di grano sia piccolo. D'altra parte, negare che le implicazioni $Pic(n) \Rightarrow Pic(n+1)$ siano tutte vere significa (nell'ambito della logica classica) ammettere l'esistenza di un intero k tale che $Pic(k) \Rightarrow Pic(k+1)$ sia falsa e quindi, per la tavola di verità dell'implicazione, tale che $Pic(k)$ risulti vera e $Pic(k+1)$ falsa. Ciò condurrebbe ad ammettere la seguente asserzione:

esiste un numero magico k tale che un mucchio con k chicchi di grano è (ancora) piccolo, ma cessa di essere piccolo non appena gli si aggiunge un solo chicco.

Questo punto di vista da "goccia che fa traboccare il vaso" sembra poco accettabile, se non altro per il fatto che contrasta con la flessibilità della nozione di essere piccolo. D'altra parte il significato del predicato Pic comporterebbe anche che $Pic(n)$ debba essere vero per ogni $n \leq k$ e falso per ogni $n \geq k+1$. Ciò è in contrasto con la vaghezza di tale predicato.

L'unica via di uscita sembra allora quella di dichiarare l'insensatezza di qualunque forma di ragionamento che coinvolga nozioni vaghe come quella di *piccolo*. 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 limite 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 concetti vaghi. 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).

Ne segue che per Frege bisogna tentare di precisare ogni concetto vago se lo si vuole inserire in un discorso razionale. Della stessa opinione è Russell che vede nel linguaggio logico lo strumento per evitare le insidie della vaghezza di cui è impregnato il linguaggio naturale:

"Voi sapete che io ho inventato un linguaggio speciale allo scopo di evitare la vaghezza. Purtroppo tale linguaggio non può essere usato in una occasione pubblica come questa.⁸ Pertanto, per quanto riluttante, io vi parlerò in inglese e quindi, ogni volta che si troverà vaghezza nelle mie parole, tale vaghezza deve essere addebitata ai nostri antenati per non essersi sufficientemente interessati alla logica."

Potremmo chiamare un tale punto di vista *normativo* in quanto più che esaminare e studiare le varie forme della razionalità umana così come esistono esso cerca di imporre ad essa delle *norme* di corretto funzionamento. D'altra parte l'eliminazione della vaghezza è a volte inevitabile⁹. Ad esempio, a volte ci si

⁸ Russell sta parlando alla radio.

⁹ Chiunque, dopo avere deciso di invitare al proprio matrimonio solo i parenti stretti e gli "amici intimi", si troverà nella difficile compito di

trova nella necessità di formulare una regola o 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. Comportano ad esempio 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.

Purtroppo che la vaghezza sia eliminabile dal diritto non è affatto credibile. Ad esempio vediamo cosa dice il Prof. Luigi Vallauri Lombardi un prestigioso studioso di filosofia del diritto.

“... qual è il valore e qual è la portata della vaghezza immancabile, essenziale, delle norme giuridiche? Dico essenziale perché la vaghezza non è qualcosa che distingue norme vaghe da altre non vaghe, ma è una caratteristica della norma per il fatto di essere norma. Io direi che il doversi misurare con il problema vaghezza è la nobiltà, è lo specifico, è la responsabilità del giurista.

.... Quindi i problemi di vaghezza definiscono proprio l'identità del giurista. Se non ci fosse vaghezza non occorrerebbero giuristi.”

stilare la lista degli inviti e quindi di trasformare la nozione vaga di *“amici intimi”* in un insieme classico.

6. Ma sembra difficile fare a meno della vaghezza

(due atteggiamenti) ... *quello espresso da Frege per cui la presenza di espressioni vaghe in un linguaggio lo rende intrinsecamente incoerente, ed il punto di vista opposto di Wittgenstein secondo cui la vaghezza è un aspetto essenziale del linguaggio* (M. Dummett, 1975).

Nonostante il fatto che i paradossi esposti evidenzino le difficoltà che si incontrano nel coinvolgere i concetti vaghi, la vaghezza gioca un ruolo fondamentale in buona parte dell'attività razionale dell'uomo. Ad esempio, evidenziamo i seguenti punti.

Elasticità. Nel primo paragrafo abbiamo citato una frase come “*Mario portava un grande cappello*” evidenziando come il coinvolgimento nel discorso di concetti vaghi permette comunque di rappresentare un qualche tipo di informazione. Si potrebbe obiettare che esistono metodi più efficienti e precisi, e che, ad esempio, trasmettendo le esatte misure del cappello di cui si parla sia possibile ottenere tale precisione. Ma facciamo un esempio un po' più complesso ed immaginiamo la seguente conversazione telefonica tra due persone: Maria ed un commesso di una pasticceria:

Maria – *Devo fare una festa, mi può mandare 3 bottiglie di Coca-Cola da due litri ?*

commesso – *subito signora !*

Maria – *poi avrei bisogno di una torta grande*

commesso – *va bene ... molto grande ?*

Maria – *senza esagerare*

commesso – *... per molte persone ?*

Maria – *non sono sicura ma potrebbero essere una quindicina di bambini*

...

Naturalmente la prima richiesta “*3 bottiglie di Coca-Cola da due litri*” non crea nessun problema: la trasmissione di informazione è perfetta ed il commesso sa esattamente quello che deve fare. Invece la nozione di “*torta grande*” crea problemi. Infatti non è detto che l'idea che ha il commesso di torta grande coincida con

quella di Maria. Se Maria è abituata a mangiare molto e se vive in una famiglia numerosa, allora potrebbe considerare piccola una torta che tutti considererebbero grande. D'altra parte su torte che non sono nettamente grandi o nettamente piccole la stessa Maria in momenti diversi potrebbe dare giudizi diversi. Questo in quanto Maria non è una bilancia di precisione.

Non è affatto chiaro in che cosa consista una tale fenomeno comunicativo. Quello che avviene nella mente del commesso quando sente parlare di una torta grande è la "messa in gioco" di tutta una collezione di *possibili torte* che potrebbero verificare la proprietà di essere "torta grande". Durante la chiacchierata telefonica potrebbe essere che tale messa in gioco si modifichi, che alcuni modelli divengano più plausibili ed alcuni meno. In tale senso sembra che sia un elemento essenziale della comunicazione che coinvolge concetti vaghi il fatto che le idee si precisino man mano che l'informazione aumenta. Tuttavia si ha la sensazione che, in ogni istante di tale gioco, nella mente dei conversanti ci sia un solo modello che rappresenti una sorta di media tra i possibili modelli e che tale modello perda genericità man mano che la conversazione va avanti.¹⁰

¹⁰Tale fenomeno non è caratteristico solo della vaghezza. Nel leggere un racconto tendiamo a costruirci fin dall'inizio una immagine di quanto letto anche se l'informazione disponibile è totalmente incompleta. Se all'inizio di un racconto leggo

"Maria avanzò rapidamente lungo la strada del paesino."

sarò indotto a farmi una immagine di Maria, della strada e del paesino. Forse è proprio l'espressione vaga "rapidamente" quella che crea meno problemi. Ad esempio l'immagine che io mi sono formato è di una donna di età tra i 20 e 30 anni, bruna e non brutta. Questo perché Maria è un nome meridionale ed anche il repertorio dei paesini della mia mente è meridionale. Poi il "rapidamente" evoca giovinezza. Non è chiaro perché escludo che sia una bambina bionda o una vecchietta che ha fatto tardi. Potrebbe darsi che io mi riferisca alla media delle mie esperienze. Se poi leggo

Ad un tratto incontrò la sua maestra ...

ecco che cambia completamente l'immagine che mi ero formata senza che questo mi crei alcun disagio. Il termine "*maestra*" crea un abbassamento dell'età ed anche i capelli si schiariscono un po' in quanto in genere le bimbe hanno i capelli più chiari degli adulti. Si potrebbe anche dire che il significato di quanto man mano si legge si realizzi nella mente del lettore solo dopo la lettura di alcuni paragrafi o alla fine

Qualunque cosa passi per la mente del cliente della pasticceria e del commesso è certo che siamo in presenza di comunicazione tra due persone razionali e che tale comunicazione non appartiene al mondo inferiore dei non-scienziati (ho dimenticato di dire che Maria è Professore Ordinario di Logica matematica ed il commesso un precario laureato in fisica e con due dottorati ... ma sono sicuro che il lettore questi particolari li aveva già intuiti). D'altra parte ogni tentativo di eliminare l'imprecisione ed i concetti vaghi sarebbe negativo per la comunicazione tra esseri umani. Se Maria e Mario abitassero in un mondo, che chiamiamo *Pignolia*, in cui è di obbligo la precisione allora nascerebbero seri problemi. Infatti in *Pignolia* tutti dovrebbero avere sempre un metro ed una bilancia in tasca per potere parlare. Ad esempio Maria dovrebbe dire per telefono qualcosa del genere:

Maria – poi avrei bisogno di una torta tonda con un diametro di 73 centimetri e 38 millimetri e con un altezza di 11 centimetri e 22 millimetri.

Allora il compito del commesso sarebbe semplificato rispetto alla interpretazione delle parole ma reso più complicato rispetto al problema di rispondere positivamente alla richiesta. Infatti supponiamo che il commesso si accorga che tutte le torte della pasticceria hanno un diametro più piccolo di quello richiesto ma che ne esista una il cui diametro è di 73 centimetri e 37 millimetri e la cui altezza è di 12 centimetri ed inoltre a forma di cuore. Allora in *Pignolia* il commesso dovrebbe rispondere negativamente e dire che non può esaudire la richiesta. Oppure dovrebbe calcolare il volume della torta che ha in negozio e confrontarlo con il volume della torta richiesta. Naturalmente in questo caso tutti concorderebbero nel dire che un tale comportamento sarebbe irrazionale. In effetti la vaghezza della richiesta di Maria è anche voluta ed è espressione di una certa tolleranza su come si debbano interpretare le parole nel gioco linguistico quotidiano.

di un capitolo. Ma ciò non mi sembra corrispondere alla realtà poiché spesso si comincia a "gustare" la lettura dopo avere letto poche righe e questo significa che un qualche significato viene trasmesso sin dall'inizio.

In Pignolia sarebbe anche difficile fare delle buone leggi. Per esempio una legge che dica

“chi esercita molestie sessuali è passibile da uno a due anni di carcere”

non potrebbe essere accettata per la difficoltà di collocare le azioni dell'accusato all'interno di uno spettro che va da azioni classificabili come “corteggiamento insistente” ad azioni classificabili come “violenze sessuali”. Di fatto un giudice ha una larga possibilità di interpretazione della parola “molestie sessuali” e questa discrezionalità è sicuramente un fatto negativo perché un giudice maschilista potrebbe dare una interpretazione non soddisfacente. Ma:

è possibile fare una legge più precisa ?

Da notare che deve essere fatta una legge che valga per persone di tutte le età, di tutte le condizioni sociali, di tutti i luoghi di un paese. Inoltre la legge dovrebbe resistere almeno qualche anno all'evoluzione dei costumi.

Ragionamenti vaghi. Senza i concetti vaghi non sarebbe possibile usare la ragione nella vita di tutti i giorni poiché la maggior parte dei ragionamenti che facciamo coinvolgono tali concetti. Esaminiamo ad esempio la seguente forma di argomentazione

SE

l'albergo è vicino alla stazione

E

l'Università è vicinissima alla stazione

ALLORA

l'albergo non è lontano dall'Università

Un altro esempio è il seguente:

SE

Maria è bella

E

Luisa assomiglia a Maria

ALLORA

Luisa non può essere troppo male.

Non ci sono dubbi sul fatto che queste siano forme di argomentazioni valide in cui sono coinvolti predicati vaghi come “*essere vicino*” e “*non essere lontano*” ...

Algoritmi vaghi : Moltissimi algoritmi usano nozioni vaghe. Basta pensare che ogni ricetta di cucina è un *programma* che usa istruzioni vaghe e se ad esempio vogliamo cucinare una crema dobbiamo applicare una ricetta del tipo:

1. metti una pentola quasi piena di latte sopra una fiamma debole
2. aggiungi un uovo
3. aggiungi un pizzico di sale
4. Se il contenuto è liquido allora effettua un giro di cucchiaino altrimenti vai all'istruzione 5
5. spegni i fuochi.

Che in questo modo si sia descritto un algoritmo non ci sono dubbi in quanto ogni essere umano è in grado di eseguire tali istruzioni. Il fatto che l'esecuzione di tali istruzioni può avvenire con maggiore o minore successo vuole dire ancora una volta che in presenza di concetti vaghi non si riesce a trasmettere informazione in modo completo. Per chi voglia imparare a cucinare non bastano i libri di cucina, l'apprendimento dall'esperienza gioca infatti un ruolo essenziale.

In definitiva una larga parte del comportamento razionale dell'uomo, almeno quello della vita di tutti i giorni, coinvolge la trasmissione e l'elaborazione di informazioni coinvolgenti nozioni vaghe. Dichiarare che informazioni di tale tipo vanno tenute fuori da un rigoroso discorso scientifico vuol dire accettare che la razionalità abbia spazio solo in alcuni campi tradizionali della ricerca scientifica (la matematica, la fisica, la chimica, forse la biologia). Soluzione, questa, non solo, per così dire, rinunciataria, ma anche illusoria se, come abbiamo visto con i paradossi, nemmeno in tali campi è possibile evitare di coinvolgere nozioni vaghe ed i paradossi che ne conseguono.

7. La soluzione ai paradossi proposta dalla logica fuzzy

Vediamo ora come la logica fuzzy affronta i paradossi della vaghezza. Consideriamo ad esempio il paradosso del mucchio di grano e riprendiamo in esame le implicazioni $Pic(n) \Rightarrow Pic(n+1)$ sulle verità delle quali, a differenza della formula $Pic(1)$, si è

portati a nutrire un qualche sospetto. Non ci sentiamo di negarle ma nemmeno di accettarle come totalmente vere. Tentiamo allora di formalizzare la modalità “non completamente vero” coinvolgendo valori di verità diversi da 0 ed 1 (che come al solito, denotano rispettivamente il falso ed il vero), valori quindi che siano capaci di rappresentare livelli intermedi di verità. Ad esempio possiamo considerare come possibili valori di verità i numeri dell’intervallo $U = [0,1]$ nel campo dei numeri reali e quindi assumere:

$Pic(1)$ con grado di verità 1
 $Pic(n) \Rightarrow Pic(n+1)$ con grado di verità 0.9.

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

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

dove $\otimes$ è una opportuna operazione binaria in U che esprime la dipendenza di cui si diceva. Una applicazione di tale regola sarà letta al modo seguente:

SE α è stata provata con grado x

E $\alpha \Rightarrow \beta$ è stata provata con grado y

ALLORA β si può ritenere provata con grado $x \otimes y$.¹¹

L’operazione $\otimes$ corrisponde alla valutazione di quanto sia lecito accettare la premessa α e la premessa $\alpha \Rightarrow \beta$. Pertanto deve essere una ragionevole interpretazione della congiunzione in una logica a più valori. Ad esempio possiamo supporre che $\otimes$ coincida con l’usuale prodotto tra numeri reali.

Ritorniamo ora all’argomentazione che conduce al paradosso del mucchio di grano¹², e formalizziamola al modo seguente:

Poiché

¹¹ Come vedremo nel capitolo 3, in realtà il significato di una asserzione del tipo “ β è provata con grado λ ” è che si ritiene che il grado di verità di β sia almeno λ (e non necessariamente uguale a λ).

¹² Anticipiamo la nozione di dimostrazione in un contesto di informazione vaga, nozione che riprenderemo e preciseremo nel seguito.

$Pic(1)$ [vale con grado 1]
e

$Pic(1) \Rightarrow Pic(2)$ [è accettato con grado 0.9]

allora, applicando *MP*,

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

Poiché

$Pic(2)$ [è provato con grado 0.9]

e

$Pic(2) \Rightarrow Pic(3)$ [è accettato con grado 0.9]

allora, applicando una seconda volta *MP*,

$Pic(3)$ [è provato con grado 0.9^2]

...

Poiché

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

e

$Pic(\underline{m-1}) \Rightarrow Pic(\underline{m})$ [è accettato con grado 0.9]

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

$Pic(\underline{m})$ [è provato con grado 0.9^{m-1}]

D'altra parte, se ci riferiamo alla norma del prodotto allora, se m è sufficientemente grande, 0.9^{m-1} è quasi uguale a zero. In definitiva abbiamo provato il seguente teorema:

Teorema 7.1. Per ogni modo di fissare m esiste una dimostrazione del fatto che un mucchio di m chicchi è piccolo ma per m grande il grado di validità di tale dimostrazione è prossimo a zero.

Si osservi che tale modo di risolvere il paradosso corrisponde ad una idea di “piccolo” come predicato applicabile con grado non nullo anche a mucchi enormi (anche se tale grado è prossimo a zero). Ciò equivale ad una idea di “grande” come predicato che non può essere 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.

Per quanto riguarda gli altri paradossi credo che si possano risolvere in maniera simile anche se per quello che riguarda il paradosso della teoria dell'evoluzione non mi appare completamente chiaro come si possa fare. Il paradosso dell'uomo calvo verrà ripreso nel seguito.

8. Altre soluzioni: il finitismo stretto

Esaminiamo brevemente due risposte al paradosso del mucchio di grano che sono diverse da quelle proposte dalla logica fuzzy e che quasi sempre sono preferite dai filosofi. Si tratta di quella data dal “*finitismo stretto*”, e quella data dalla “*teoria delle supervalutazioni*”.

Il finitismo stretto. A meno che non si voglia giungere a rifiutare il Modus Ponens, un modo per negare la validità del ragionamento con cui si prova il paradosso potrebbe essere 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. Tale punto di vista non solo esclude insiemi infiniti di operazioni ma anche insiemi “troppo grandi” di operazioni. Ne segue che solo dimostrazioni piccole possono essere considerate accettabili e quindi che il ragionamento che conduce al paradosso del mucchio di grano, 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.

Tuttavia mi sembra che esistano molte obiezioni al finitismo stretto. Eccone alcune.

1. Se si accetta il finitismo stretto ogni teorema recente deve essere considerato inaffidabile.

Infatti se si prende la dimostrazione π di un qualche recente teorema di matematica la sua lunghezza è di fatto enorme. Questo perché quasi sempre la dimostrazione di un teorema utilizza cose già dimostrate nel passato utilizzando espressioni del tipo “*D'altra parte per il Teorema T la cui dimostrazione si può trovare in ... risulta che ...*”.

Allora è evidente che nel calcolare la lunghezza di π si deve tenere conto anche della lunghezza della dimostrazione π' del teorema T . A sua volta nel calcolare la lunghezza di π' è probabile che si debba considerare la lunghezza delle dimostrazioni di altri teoremi utilizzati in π' e così via. Dovremmo allora dubitare di tutti i teoremi dimostrati recentemente. Questo è ovviamente irragionevole.

2. D'altra parte, se con il finitismo stretto si riesce a risolvere il paradosso del mucchio di grano non è difficile trovare paradossi analoghi che, riferendosi ad una dimostrazione con un numero molto piccolo di passi, non risultano risolti. 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 $s_1, \dots, s_{20}$ 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.

3. Da notare inoltre che se si accetta il principio di induzione¹³ allora il finitismo stretto non permetterebbe di bloccare nemmeno il paradosso del mucchio di grano. Infatti proviamo ad accettare, insieme alle due formule

$$(i) Pic(1)$$

$$(ii) \forall n (Pic(n) \Rightarrow Pic(n+1))$$

anche il principio di induzione applicato al predicato Pic

¹³ Moltissime dimostrazioni in matematica si basano sul principio di induzione che si può esprimere dicendo che se P è una proprietà definita nell'insieme dei numeri naturali e risulta

- che 1 verifica P
- che se il numero n verifica P allora anche $n+1$ verifica P

allora posso concludere che ogni numero naturale verifica P . Se ci riferiamo ai sottoinsiemi di N il principio di induzione si presenta come una formula del secondo ordine del tipo

$$\forall X (1 \in X \rightarrow ((\forall n (n \in X \rightarrow n+1 \in X) \rightarrow X = N)).$$

Se ci riferiamo alle proprietà dei numeri interi che possono essere espresse in un dato linguaggio, si ottiene una forma più debole del principio di induzione considerando lo schema di assioma

$$P(1) \rightarrow ((\forall n (P(n) \rightarrow P(n+1)) \rightarrow \forall n P(n)).$$

che rappresenta le infinite formule del primo ordine che si ottengono ponendo al posto di P una qualunque formula del linguaggio fissato.

(iii) $Pic(1) \Rightarrow ((\forall n(Pic(n) \Rightarrow Pic(n+1)) \Rightarrow \forall n Pic(n))$.

Allora due semplici applicazioni del *MP* permettono di provare la formula $\forall n Pic(n)$ e quindi, per particolarizzazione, $Pic(\underline{m})$ per ogni prefissato m .¹⁴

4. Il finitismo non sfugge dal coinvolgimento della vaghezza. Infatti nell'affermare che solo le dimostrazioni con un numero piccolo di passi sono valide il finitismo deve confrontarsi con la nozione vaga di "piccolo". Ne segue che esso non può sfuggire a paradossi del tipo "mucchio di grano" applicati alla nozione di dimostrazione. Infatti è ragionevole supporre che una dimostrazione fatta di un solo passo sia piccola e che se ad una dimostrazione piccola si aggiunge una semplice conseguenza si ottenga ancora una dimostrazione piccola. Allora da tale due ipotesi, che sono difficilmente confutabili, segue che non esistono limiti di lunghezza per le dimostrazioni piccole.¹⁵

9. Altre soluzioni: la teoria delle supervalutazioni

Un modo diverso di risolvere il paradosso è di utilizzare la teoria delle supervalutazioni. Essa è stata delineata da Dummett nell'articolo citato e successivamente sviluppata da K. Fine. Per Fine le nozioni vaghe si caratterizzano per il fatto di avere molte possibili "determinazioni" (da intendere come determinazioni classiche), 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

¹⁴ E' interessante osservare che la lunghezza di una tale dimostrazione non dipende dalla grandezza di m come invece avveniva nella formulazione che abbiamo esposto del paradosso del mucchio di grano.

¹⁵ Se si considera la cosa da un punto di vista probabilistico e ci si rivolge alle dimostrazioni come a fatti concreti, allora il punto di vista finitista non appare poi tanto bizzarro. Infatti supponiamo che la probabilità che un passo della dimostrazione sia corretto sia un valore δ che, pur essendo vicinissimo ad 1, sia diverso da 1. Allora la probabilità che una dimostrazione di lunghezza n sia corretta è δ^n (in ipotesi di indipendenza degli eventi). Poiché $\lim_{n \rightarrow \infty} \delta^n = 0$, per n molto grande tale probabilità è quasi uguale a zero. Il fatto che attualmente si accettino dimostrazioni lunghissime in quanto i suoi pezzi sono effettuate da calcolatori, rende abbastanza plausibile tale discorso.

chicchi e così via. In un certo senso il significato di un concetto vago è dato proprio dall'insieme delle sue possibili determinazioni.¹⁶ 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 .

Definizione 9.1. Una asserzione α è *SV-vera* (rispettivamente *SV-falsa*) quando risulta vera (falsa) per tutte le possibili determinazioni¹⁷ dei predicati vaghi che contiene. Diremo che α è *indeterminata* se non risulta essere né *SV-vera* né *SV-falsa*.

Ad esempio l'asserzione $Pic(1000)$ è da considerare indeterminata essendoci dei contesti in cui un mucchio di 1000 chicchi è grande e contesti in cui è piccolo. Invece $Pic(1)$ è vera ed è ragionevole considerare $Pic(10^{1000})$ falsa.

Proposizione 9.2. Il Modus Ponens è una regola di inferenza valida rispetto alla nozione di *SV-verità*. Inoltre tutte le tautologie classiche continuano a valere. Tuttavia la logica risultante non è vero-funzionale se si considera “indeterminato” come un valore di verità.

Poiché il Modus Ponens è valido, se si accettassero come *SV-veri* tutte le implicazioni $Pic(\underline{n}) \Rightarrow Pic(\underline{n+1})$, il paradosso continuerebbe a sussistere.

Teorema 9.3. Alcune delle implicazioni $Pic(\underline{n}) \Rightarrow Pic(\underline{n+1})$ non sono *SV-veri*, pertanto la teoria delle supervalutazioni permette di

¹⁶ La nozione di *taglio* di una struttura fuzzy, che verrà trattata nel seguito, potrebbe essere il corrispondente, nell'ambito della teoria degli insiemi fuzzy, della nozione di “determinazione”.

¹⁷ L'espressione “tutte le possibili determinazioni” non deve essere intesa nel senso utilizzato, ad esempio, per definire le tautologie. Invece si deve intendere determinazioni che di fatto in varie circostanze ha fatto una data comunità di esseri umani. Ciò comporta che asserzioni come $2x-6=0 \Rightarrow x=3$ (che ovviamente non sono tautologie) sono da considerare vere poiché la sua interpretazione sugli interi è unica e tale interpretazione la rende vera.

escludere l'argomentazione alla base del paradosso del mucchio di grano. Allo stesso tempo mentre la formula $\exists n(Pic(n) \wedge \neg Pic(n+1))$ risulta SV-vera, per nessun intero n possiamo affermare che $Pic(\underline{n}) \wedge \neg Pic(\underline{n+1})$ sia SV-vera.

Una tale soluzione al paradosso è abbastanza convincente, tuttavia sono possibili alcune obiezioni.

1. 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. Di fatto, si finisce sempre con il buttare nel cestino dei rifiuti tutte le argomentazioni basate su nozioni vaghe e quindi non si riesce a dare conto del perché tali argomentazioni siano tanto frequenti ed utili nella vita quotidiana.

2. Il passare dalla categoria vero-falso della logica classica a quella a tre valori vero-falso-indeterminato della teoria delle supervalutazioni non sembra rendere conto della ricchezza e flessibilità dei concetti vaghi.

3. Sembra molto discutibile l'assunzione per cui il significato di un concetto vago possa essere identificato con l'insieme delle sue possibili determinazioni. Non si deve confondere vaghezza e dipendenza dal contesto. Esaminiamo ad esempio la nozione giuridica di "maggiorenne", tale nozione può assumere in epoche e paesi differenti significati diversi, ognuno dei quali esattamente determinato dall'aver fissato una età-soglia. Sarebbe allora meglio dire che il significato di "maggiorenne" più che essere vago dipende dal contesto. Invece per l'analoga nozione "essere maturo" qualunque sia il paese e l'epoca resta il fatto che si può essere più o meno maturi e che una persona può essere più matura dell'altra. Allo stesso modo qualunque sia il contesto in cui si usa la nozione di mucchio piccolo, rimane il fatto che tale nozione rimane vaga come mostra il fatto che un mucchio può essere considerato più piccolo di un altro.

4. Quando si rapportano tra loro più proprietà vaghe entrano in gioco i gradi con cui tali proprietà sono soddisfatte. Ad esempio consideriamo, l'implicazione

se sei vicino ad un casello autostradale allora devi procedere con velocità piccola.

In tale espressione non solo sono coinvolti due concetti vaghi, *vicino* e *piccolo*, ma viene espresso anche un qualche collegamento tra i gradi con cui tali concetti sono verificati. Il significato sottinteso è che tanto più x è vicino ad un casello quanto più la sua velocità deve essere 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 determinazioni dei concetti vaghi in gioco, ciascuna ragionevole se presa a se stante. La corrispondente determinazione della frase sopra scritta sarebbe allora

se sei a meno di 500 metri da un casello autostradale allora devi procedere 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!

5. La principale obiezione è tuttavia che esistono proprietà di strutture vaghe che non sono riconducibili a proprietà delle possibili determinazioni di tali strutture: la teoria delle supervalutazioni non riesce a dare conto di questo fatto. Ad esempio questo avviene per la proprietà transitiva. Per una giustificazione di tale affermazione, che richiede alcune nozioni che verranno date nel seguito, si rimanda al paragrafo 14 del secondo del secondo capitolo.

10. Il paradosso del mentitore

A supporto della logica fuzzy viene anche messa la sua capacità a risolvere il famosissimo *paradosso del mentitore* (Zadeh [1979]). Tale paradosso, nella forma attribuita ad Eubulide di Mileto (IV secolo a.C.), procede al modo seguente. Consideriamo la frase

$A = \text{"io sto mentendo"}$.

Allora se A è falsa vuol dire che afferma il vero. Se A è vera allora vuol dire che afferma il falso. Esistono varie riformulazioni di un tale paradosso. Ad esempio Aristotele considera l'ordine:

"disobbedisci all'ordine che ti sto dando in questo momento"

Allora se si obbedisce all'ordine non si obbedisce all'ordine. Se invece non si obbedisce si obbedisce. Preferiamo riformulare il paradosso al modo seguente.

Proposizione 10.1. Consideriamo le due asserzioni

$P = \text{"L'asserzione } Q \text{ è vera"}$

$Q = \text{"L'asserzione } P \text{ è falsa"}$.

Allora è possibile provare l'equivalenza $P \Leftrightarrow \neg P$.

Dim. Per provare $P \Rightarrow \neg P$ osserviamo che la verità di P comporta la verità di Q che a sua volta comporta la falsità di P . Per provare $\neg P \Rightarrow P$ osserviamo che se P è falsa allora Q deve essere necessariamente falsa e quindi P vera.

Ora in logica classica sia che si assegni a P il valore di verità 0 o il valore di verità 1, una tale equivalenza risulta impossibile.¹⁸

Come mostra il seguente teorema, nelle logiche che ammettono più valori di verità essa può invece sussistere.

Teorema 10.2. Esistono logiche fuzzy tali che il fatto che la formula

$$P \Leftrightarrow \neg P$$

valga con grado 1 non è una contraddizione ma vale se e solo se il valore di verità di P è 0.5.

Dim. Consideriamo una logica fuzzy in cui

- l'equivalenza $\Leftrightarrow$ sia interpretata da una operazione binaria $\leftrightarrow$ tale che $x \leftrightarrow y = 1$ se e solo se x coincide con y ,

- e la negazione $\neg$ è interpretata con la funzione $1-x$.

Detto allora λ il valore di verità di P , il fatto che $P \Leftrightarrow \neg P$ valga con grado uno equivale a dire che $\lambda = 1-\lambda$. Nell'intervallo U tale equazione ammette come (unica) soluzione il valore 0.5 e quindi l'affermazione dell'equivalenza non è un assurdo.

Da molti fuzzisti tale teorema mostrerebbe che il paradosso del mentitore è di fatto una prova dell'esigenza di ammettere almeno

¹⁸ L'importanza di tale paradosso non è da sottovalutare poiché la sua struttura è alla base dei famosi "teoremi di Gödel".

un terzo valore di verità diverso da 0 e da 1 e quindi una dimostrazione della necessità della logica fuzzy.¹⁹

Tuttavia anche se il paradosso del mentitore come lo abbiamo esposto viene risolto, a me sembra che il meccanismo che ne è alla base, legato all'autoriferimento, continui a generare paradossi che sono sostanzialmente dello stesso tipo. Ad esempio, anche se accettiamo la logica fuzzy non possiamo certo escludere che sia definibile tra i valori di verità l'usuale nozione di identità. Allora possiamo riformulare il paradosso del mentitore scrivendo:

P = "La frase seguente ha valore di verità uguale ad 1"

Q = "La frase precedente ha valore di verità diverso da 1".

Ora, se P avesse valore 1 allora Q avrebbe valore 1 e quindi P avrebbe valore diverso da 1. Se P avesse valore diverso da 1 allora Q risulterebbe vera, cioè con valore 1. Allora P asserirebbe qualcosa di vero e quindi avrebbe valore 1. Si è provato quindi che, indicando con $v(P)$ il valore di verità di P ,

$$v(P) = 1 \Leftrightarrow \neg(v(P) = 1).$$

Tale equivalenza è assurda poiché abbiamo indicato con $=$ la relazione classica di identità.²⁰ Da notare che tale equivalenza è diversa da

$$P \Leftrightarrow \neg P$$

che, potendo P assumere valori in U , non risulta assurda in logica fuzzy.

Ma esistono infiniti modi simili di produrre un paradosso. E' sufficiente coinvolgere l'ambiente "non fuzzy" dei valori di verità. Ad esempio consideriamo le frasi seguenti:

P = il valore di verità di Q è strettamente maggiore del mio valore di verità

Q = il valore di verità di P è maggiore o uguale al mio valore di verità.

Entrambe le asserzioni appartengono alla logica classica poiché coinvolgono l'ordinamento tra valori di verità e tale ordinamento

¹⁹ Un discorso simile si potrebbe forse fare per il paradosso di Russell nel quale si prova che, posto $R = \{X : X \text{ non appartiene ad } X\}$, risulta che $R \in R \Leftrightarrow \neg(R \in R)$. In una logica fuzzy tale conclusione non sarebbe assurda ma comporterebbe solo che il valore di verità dell'asserzione $R \in R$ è 0.5.

²⁰ Argomenti simili relativi al paradosso del mentitore sono considerati ampiamente in letteratura, ad esempio si veda [Reiger 2001].

è crisp. Ora P non può essere vera poiché in questo caso il valore di Q dovrebbe essere strettamente maggiore di 1. Appurata la falsità di P , Q non può essere vera, se lo fosse dovrebbe avere valore di verità minore di 0. Tuttavia Q non può essere falsa poiché in tale caso verificando la condizione di avere valore minore o uguale a quello di P , sarebbe di fatto vera.

In definitiva la logica fuzzy, contrariamente a quanto viene spesso affermato, non sembra affatto in grado di risolvere tale versione del paradosso del mentitore.

11. Non solo paradossi: la scienza ingenua

Un altro paradosso che può essere affrontato dalla logica fuzzy è quello, di Poincaré, legato alla relazione di indiscernibilità. Esporremo tale paradosso nel capitolo 3 quando avremo fornito qualche strumento tecnico in più. Tuttavia, che sia o non sia uno strumento valido per trattare i paradossi, la logica fuzzy non è certo nata per tale scopo come non è nata per essere uno strumento di analisi per il fenomeno della vaghezza nel linguaggio umano. Essa è nata per risolvere alcuni problemi sorti in un ambiente quanto mai lontano dalle speculazioni filosofiche: quello ingegneristico della teoria del controllo. Pertanto per capire il significato della logica fuzzy credo che si debbano capire i motivi del suo successo nel risolvere tale classe di problemi. Riprenderemo questo discorso nel quarto capitolo dopo avere parlato un po' del controllo fuzzy. Qui anticipo solo che tale successo è legato alle potenzialità che si aprono nel dare dignità scientifica e tecnologica al punto di vista qualitativo in una scienza la quale, invece, storicamente è nata dal mettere esclusivamente al proprio centro il punto di vista quantitativo. Più in generale credo che la nascita della logica fuzzy debba essere inquadrata nell'ambito della rivalutazione della cosiddetta "scienza ingenua". Con tale termine viene indicato quell'insieme di "teorie" che l'uomo comune tende naturalmente ad elaborare nel corso della sua quotidiana attività di sopravvivenza in un mondo che non ha né la dimensione delle particelle atomiche né quello delle galassie. Rientrano in tale ambito ricerche apparentemente diverse come quelle sulle ontologie, sulla geometria senza punti ed altro (si vedano ad esempio [Bozzi 1998], [Varzi 1999], [Smith e Casati 1994], [Casati e Varzi 2002] [Gerla 1995]).

In realtà la parola “qualitativo” potrebbe essere considerata discutibile nel caso della logica fuzzy. Infatti, in un certo senso i modelli che si propongono del qualitativo sono di fatto di tipo numerico. La questione è che nella logica classica la scrittura di una asserzione significa per convenzione che essa è vera, e quindi la scrittura della sua negata significa che è falsa. Pertanto il livello linguistico permette di rappresentare anche il livello semantico. Non può avvenire lo stesso per la logica fuzzy che ha bisogno invece di una dimensione numerica.

12. Probabilità e fuzziness: esiste un rapporto ?

Chiudo questo capitolo introduttivo considerando la questione della connessione tra questi due modi di graduare la verità che sembrano avere qualcosa in comune. Diciamo subito che si tratta di cose nettamente diverse sia dal punto di vista filosofico che da quello tecnico. Il fatto che in entrambi i casi ci si riferisca all'intervallo U non deve condurre a confusione.²¹ Per convincerci di questo fatto supponiamo di fare due esperimenti.

Primo esperimento. Nella stanza A ci sono 3 bicchieri d'acqua uguali ed una bottiglietta contenente un veleno incolore. Una o due gocce di tale bottiglietta creano un leggero malessere, ma se si arriva a 3 la morte è sicura. Supponiamo ora che mi dicano che 3 gocce sono state versate in uno dei bicchieri e che successivamente io veda Carlo entrare nella stanza, scegliere un bicchiere a caso e bere. Allora sono portato ad affermare che:

- Mario è avvelenato con probabilità $1/3$.

Secondo esperimento. Supponiamo che nella stanza B ci siano ancora 3 bicchieri uguali ma che io venga a sapere che in questo caso in ogni bicchiere si sia versata una sola goccia del veleno. In questo caso sono portato a pensare che:

- Luigi ha una avvelenamento di grado $1/3$.

²¹ Tuttavia il paradigma per la definizione di un apparato inferenziale per la logica fuzzy si applica perfettamente alla definizione di un apparato inferenziale per la logica probabilistica. Questa idea è esplorata in [Gerla 1994]*b* ed è esposta brevemente alla fine del terzo capitolo.

Si vede allora che lo stesso numero $1/3$ misura cose completamente diverse.²² Nel primo caso la proprietà “essere avvelenato” è considerata classica ed alla fine dell’esperimento Carlo o è avvelenato (e quindi morto) oppure non lo è. Nel secondo caso la stessa proprietà è considerata “graduata” ed alla fine dell’esperimento tale proprietà è verificata con grado $1/3$.

D’altra parte basta osservare il ruolo che ha l’informazione nelle due situazioni. Supponiamo che nell’osservare il primo esperimento io, guardando dal buco della serratura, mi accorga che il bicchiere più lontano dalla porta non sia stato toccato. Allora se vedo Carlo prendere un bicchiere vicino alla porta, la mia valutazione probabilistica circa l’avvelenamento di Carlo passa da $1/3$ a $1/2$. Nelle stesse ipotesi invece nel secondo esperimento continuo a pensare che il grado di avvelenamento di Luigi sia $1/3$. Ciò è in accordo con il fatto che le valutazioni probabilistiche variano a seconda dell’informazione disponibile mentre quelle legate alla vaghezza no.²³

Queste considerazioni sembrano confutare, più in generale, ogni teoria della vaghezza che vede questa come espressione di mancanza di conoscenza. E’ questa ad esempio l’idea di T. Williamson che ritiene che il fenomeno dei casi borderline, che caratterizzerebbe la vaghezza, non appartenga alla realtà ma derivi dalle nostre limitate capacità cognitive (si veda [Williamson 1994] e [Williamson 2000]). Questo punto di vista, almeno nella forma proposta da Williamson, viene criticato in un interessante articolo di Sebastiano Moruzzi [Moruzzi 2002].

Non sorprende che alle differenze di tipo filosofico tra probabilità e vaghezza corrispondono anche notevoli differenze di tipo tecnico. Quella forse più rilevante è che la logica probabilistica non è vero-funzionale mentre quella fuzzy in generale lo è. Ad esempio consideriamo la congiunzione di due asserzioni. Nel caso della logica fuzzy, come vedremo meglio nel

²² Chi non fosse convinto può proporre ad un amico quale dei due esperimenti accetterebbe di fare per 100.000 euro.

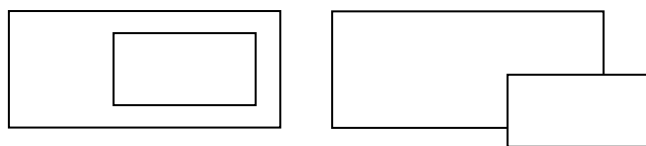
²³ Anche le valutazioni di predicati vaghi variano poiché può variare il sottoinsieme fuzzy utilizzato per modellizzare tale predicato. Questa variabilità non dipende però dall’acquisizione di nuove informazioni ma da “aggiustamenti” di tale modellizzazione provvisoria, aggiustamenti dovuti a motivi di ordine pragmatico o se si vuole dovuti al desiderio di avvicinarsi all’obiettivo che ci si è prefisso.

seguito, il grado di verità di una formula del tipo $\alpha \wedge \beta$ è il composto (tramite una opportuna operazione) del grado di verità di α con il grado di verità di β . Invece una valutazione probabilistica di $\alpha \wedge \beta$ dipende fortemente dal reciproco rapporto di dipendenza tra α e β . Solo se le due formule rappresentano eventi indipendenti allora la probabilità di $\alpha \wedge \beta$ è uguale alla probabilità di α moltiplicata per la probabilità di β .

A dispetto di tali considerazioni non tutti i seguaci della teoria degli insiemi fuzzy concordano con il fatto che fuzzyness e probabilità siano fenomeni totalmente diversi. Ad esempio Kosko sostiene che la teoria della probabilità rientra nella teoria degli insiemi fuzzy e che le probabilità condizionate sarebbero una sorta di relazione fuzzy di inclusione. Citiamo Kosko.

“Che cosa succede se un insieme ne contiene un altro ? Mettiamo un rettangolo piccolo in uno grande. Ecco che il rettangolo grande contiene quello piccolo. Lo contiene al 100%. Può questo contenere assumere misure diverse ? Certo che può. Mettiamo il rettangolo piccolo solo per metà in quello grande ed ecco che quest’ultimo contiene il piccolo solo al 50%. Il rettangolo piccolo è sia dentro che fuori di quello grande. Chiamo questo contenere fuzzy “sottoinsiemità”, vale a dire la misura di quanto un sottoinsieme sia sottoinsieme di un altro insieme”.

Le seguenti figure illustrano quanto detto da Kosko.



Per precisare le cose, riferiamoci per semplicità ad un insieme finito S ed alla probabilità $p : P(S) \rightarrow U$ che si ottiene ponendo $p(X) = |X|/|S|$ (percentuale di elementi in X). La probabilità condizionata di un evento B stante un evento $A \neq \emptyset$ sarà allora definita ponendo $p(B/A) = |A \cap B|/|A|$. L’idea di Kosko è che tale numero rappresenti in realtà la misura del “grado di inclusione” di A in B cioè che posto,

$$Inc(A,B) = |A \cap B|/|A| \quad \text{per ogni } A \neq \emptyset,$$

$$Inc(\emptyset, B) = 1$$

sia poi possibile definire la probabilità condizionata ponendo $p(B/A) = Inc(A, B)$. In altre parole la probabilità condizionata, e quindi la probabilità, sarebbe definibile a partire da una relazione fuzzy *Inc* di inclusione graduata. Poiché l'intuizione della nozione di “inclusione”, anche nella versione graduata, precederebbe quella di probabilità, per Kosko è giusto considerare la probabilità come un capitolo della logica fuzzy.

Detta così, si tratterebbe solo di un fatto nominalistico e quindi il considerare la probabilità come parte della logica non potrebbe essere suscettibile di confutazione. Tuttavia, poiché le parole portano in sé il peso di un significato, possono essere fatte (poche) considerazioni a favore e (molte) considerazioni contro tale scelta. A favore del fatto che *Inc* possa essere considerata una sorta di inclusione si può osservare che,

$$A \subseteq B \Leftrightarrow Inc(A, B) = 1$$

cioè che *Inc* estende l'inclusione usuale. Inoltre

$$A \cap B = \emptyset \Leftrightarrow Inc(A, B) = 0 \text{ (per } A \neq \emptyset \text{)}$$

cioè che insiemi disgiunti non possono essere inclusi tra loro. Infine vale l'implicazione

$$B \subseteq B' \Rightarrow Inc(A, B) \leq Inc(A, B').$$

Contro si può osservare che la relazione fuzzy *Inc* non verifica altre proprietà che sembrano naturali per una relazione di inclusione. Ne elenchiamo tre:

1. *Inc* non verifica l'implicazione

$$A \subseteq A' \Rightarrow Inc(A, B) \geq Inc(A', B).$$

Infatti, supponiamo che $A = \{a, b\}$, $A' = \{a, b, c\}$ e $B = \{b, c\}$. Allora $A \cap B = \{b\}$ e $A' \cap B = \{b, c\}$ e quindi $Inc(A, B) = 1/2$, $Inc(A', B) = 2/3$.

2. *Inc* non verifica la proprietà transitiva.

$$Inc(A, B) \otimes Inc(B, C) \leq Inc(A, C).$$

Infatti, supponiamo che $C = \{c, b\}$, $A = \{a\}$ e $B = \{a, b, c\}$. Allora $Inc(A, B) = 1$, $Inc(B, C) = 2/3$ e $Inc(A, C) = 0$ e quindi la proprietà transitiva potrebbe valere solo se $1 \otimes 2/3 \leq 0$ in contrasto con il fatto che 1 è elemento neutro rispetto a $\otimes$.

3. *Inc* non verifica l'implicazione

$$Inc(B, C) \leq Inc(B - C', C - C').$$

Infatti siano $C' = \{c'\}$, $C = \{c, c'\}$, $B = \{b, c, c'\}$. Allora, poiché $|B| = 3$ e $|B \cap C| = 2$, risulta che $Inc(B, C) = 2/3$. D'altra parte, poiché $|B - C'| = 2$ e $|(B - C') \cap (C - C')| = |\{c\}| = 1$, abbiamo che $Inc(B - C', C - C') = 1/2$.

Nel seguito mostreremo esempi di inclusione graduata, che chiamiamo $\otimes$ -inclusione, che non hanno tali difetti e che per la loro origine logica appaiono più significativi.

CAPITOLO 2

MODELLI MATEMATICI PER LA VAGHEZZA

“The increase of exactness leads to an increase in the amount of information whose relevancy then decreases until a point is reached after which the preciseness and relevancy are mutually excluding characteristics.”
(Principio di incompatibilità di Zadeh)

1. Insiemi fuzzy: un modello matematico per la vaghezza

Nel primo capitolo abbiamo mostrato qualche semplice esempio di ragionamento approssimato e le ragioni che portano a definire una nuova logica. In questo capitolo cominciamo col dare le prime definizioni formali riguardanti gli aspetti “semantici” della logica fuzzy. Ora, poiché la semantica della logica classica si basa sulla nozione di insieme, si pone il problema di come estendere in modo opportuno tale nozione e quindi di proporre un ragionevole modello matematico per la nozione di “insieme vago”. Il discorso può essere introdotto 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 preso un qualunque elemento x in S è possibile dire se x verifica o meno P . Ad esempio, detto N l'insieme dei numeri naturali, la proprietà “essere pari” è ben definita in N . In teoria degli insiemi il seguente principio gioca un ruolo fondamentale.

Principio di comprensione (per meglio dire *assioma di isolamento*). Sia P una proprietà ben definita in un insieme S , allora la classe degli elementi di S che verifica P , cioè l'*estensione* di P , è un insieme.

Ad esempio a proprietà come “essere primo” e “essere pari”, che sono definite in N , corrispondono gli insiemi dei numeri primi e dei numeri pari, rispettivamente.

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 si pone il problema se abbiano senso espressioni come

“l’insieme dei mucchi di grano grandi”,
 “l’insieme degli uomini alti”,
 “l’insieme degli uomini calvi” ?

Una possibile soluzione viene suggerita dalla nozione di funzione caratteristica di un sottoinsieme.

Definizione 1.1. Dato un sottoinsieme X di un insieme S , la *funzione caratteristica* di X è la funzione $c_X : S \rightarrow \{0,1\}$ definita ponendo, per ogni $x \in S$,

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

Proposizione 1.2. La corrispondenza di $P(S)$ in $\{0,1\}^S$ che associa ad ogni elemento $X \in P(S)$ la relativa funzione caratteristica è biettiva. Pertanto è possibile identificare l’insieme $P(S)$ delle parti di S con la classe $\{0,1\}^S$ delle funzioni di S in $\{0,1\}$.

Il fatto che un sottoinsieme di S si possa identificare con la sua funzione caratteristica suggerisce di rappresentare l’estensione di una proprietà vaga con una “*funzione caratteristica generalizzata*”, cioè una funzione caratteristica che può assumere valori nell’intero intervallo U .

Definizione 1.3. Dato un insieme non vuoto S , un *sottoinsieme fuzzy* di S è una funzione $s : S \rightarrow U$ di S nell’intervallo U . Diciamo che s è *crisp* se $s(x) \in \{0,1\}$ per ogni $x \in S$. La classe di tutti i sottoinsiemi fuzzy è indicata con U^S .¹

¹ La definizione di sottoinsieme fuzzy è stata proposta da L. Zadeh in un articolo del 1965 e, sotto il nome di *multivalued set*, contemporaneamente da Dieter Klaua nel contesto dello studio delle logiche a più valori. D’altra parte la nozione di insieme e di relazione fuzzy è implicita in ogni approccio alla logica multivalued del primo ordine. Ad esempio si consideri il libro di Chang e Kislner del 1966 dove si propone una interessantissima estensione di molte nozioni della teoria dei modelli classica a quella che ora viene chiamata logica fuzzy.

Coerentemente con l'uso comune, spesso invece dell'espressione "sottoinsieme fuzzy" useremo quella, meno precisa, di "insieme fuzzy". Fissato x in S , il valore $s(x)$ è interpretato come il *grado di appartenenza* di x ad s . Poiché 0 ed 1 rappresentano il falso ed il vero, nel caso $s(x) = 0$ diremo che x non appartiene all'insieme fuzzy s , nel caso $s(x) = 1$ diremo che x appartiene ad s .

Come nella matematica classica in cui una relazione è definita come un sottoinsieme di un prodotto cartesiano, nella logica fuzzy una relazione fuzzy può essere definita come un sottoinsieme fuzzy di un prodotto cartesiano.

Definizione 1.4. Dati gli insiemi non vuoti $D_1, \dots, D_n$ una *relazione n -aria fuzzy* tra gli elementi di $D_1, \dots, D_n$ è un sottoinsieme fuzzy $r : D_1 \times \dots \times D_n \rightarrow U$ del prodotto cartesiano $D_1 \times \dots \times D_n$.

Esempi di relazioni binarie fuzzy sono "essere amico di", "essere più bravo di" e così via. Nel seguito considereremo due tipi di relazioni fuzzy: gli ordinamenti fuzzy e le equivalenze fuzzy o similarità.

Proposizione 1.5. Possiamo identificare ogni sottoinsieme di S con la sua funzione caratteristica e quindi la classe $P(S)$ dei sottoinsiemi di S con la classe $\{0,1\}^S$ dei sottoinsiemi crisp di S .

In particolare possiamo identificare S con la funzione c_S costantemente uguale ad 1 e l'insieme vuoto $\emptyset$ con la funzione $c_\emptyset$ costantemente uguale a 0. Si noti tuttavia che in questo secondo caso ogni insieme S ammette un "suo" sottoinsieme vuoto. Ciò è in contrasto con il fatto che il principio di ascensionalità in teoria degli insiemi, nell'affermare che due insiemi che hanno gli stessi elementi coincidono, comporta che esiste un solo insieme vuoto.

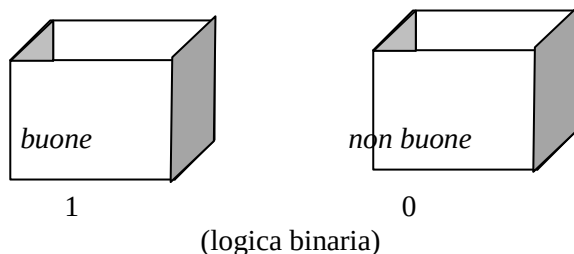
Spesso una proprietà di un insieme fuzzy si riferisce al supporto di tale insieme. Ad esempio viene utilizzata la seguente definizione.

Definizione 1.6. Un insieme fuzzy s si dice *finito* se il suo supporto $Supp(s) = \{x \in S : s(x) \neq 0\}$ è finito.

2. Qualche giustificazione per la definizione proposta

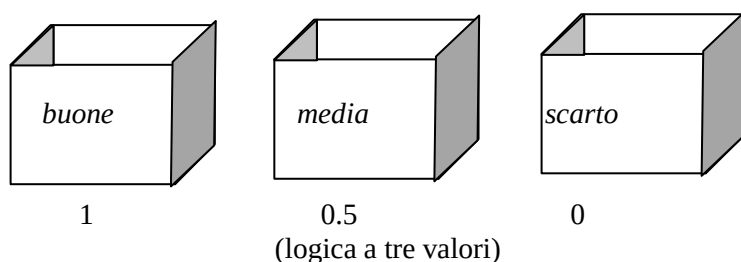
Per giustificare in qualche modo la definizione di insieme fuzzy supponiamo di dovere affrontare il seguente problema:

Problema: *Dato un sacco di mele, mettere quelle buone in una cassetta, quelle non buone in un'altra.*



Nell'ambito di una logica a due valori il problema si risolve in maniera non completamente soddisfacente. Infatti se per molte mele non si avranno problemi nella collocazione, ci saranno anche molte mele sulle quali si avranno forti dubbi se debbano essere messe nella cassetta delle buone o in quella delle non buone. Un modo più adeguato di affrontare un tale compito sembra allora il seguente:

Problema: *Dato un sacco di mele, metti quelle buone in una cassetta, quelle non buone in un'altra e quelle su cui sei incerto ancora in un'altra.*



In questo caso utilizziamo tre valori di verità ma è chiaro che aumentando il numero di valori di verità si riesce ad avere una partizione più rappresentativa della realtà. Ad esempio potremmo volere distinguere all'interno della scatola delle mele buone due livelli diversi con la prospettiva di mettere prezzi diversi. Potremmo addirittura accettare come possibili valori di verità

tutti i numeri reali dell'intervallo U . Ad esempio, potremmo rilevare con qualche strumento il diametro e l'intensità di colore rosso di una mela e poi ricavare con qualche formuletta un valore numerico tra zero ed uno che esprima il grado di bontà della mela in funzione di tali grandezze. In tutti i casi, qualunque sia il modo di fissare i valori di verità, in qualche modo si è definito un "processo" che assegna ad ogni mela un valore che rappresenta il grado di bontà della mela: cioè si è rappresentata la nozione di "essere buono" con un sottoinsieme fuzzy.

Qui nasce la prima perplessità:

pur avendo fissato i valori di verità (numero degli scatoloni con relative "etichette") siamo sicuri che persone diverse non diano diverse partizioni del sacco di mele ?

Immaginiamo che nel centro A di rivendita all'ingrosso di mele esista un operaio dedicato alla classificazione delle mele e che questo operaio sia un extra-comunitario proveniente da un paese povero. Allora mele piccole e con qualche macchia sarebbero considerate comunque non da gettare e quindi a tale mele l'operaio assegnerebbe valore 0.4. Supponiamo che nel centro B l'operaio che smista sia invece un ex stilista (fallito e costretto a fare questo mestiere). In tale caso gli aspetti estetici prendono il sopravvento e le mele macchiate e piccole sarebbero messe nello scatolone con etichetta 0 (da dare ai porci): un pasticcio. Ed allora:

- quale sarà l'insieme fuzzy che rappresenta correttamente la nozione di mela buona ?
- ha senso parlare di un argomento, gli insiemi fuzzy, che è inevitabilmente impregnato di tanta soggettività ?

Un logico, un filosofo o un religioso che abbiano dedicato la propria vita alla Verità non può che cadere nello sconforto perché la Verità è unica e non può dipendere dai gusti o dalla storia di una persona (come è noto il Relativismo che impedisce alla persone di scannarsi a vicenda per la propria verità è uno dei grandi mali di questa epoca). Un imprenditore invece che voglia organizzare un nuovo centro di rivendita avrebbe meno problemi. Confronterebbe i ricavi del centro A e del centro B ed opterebbe per l'assunzione di un extra-comunitario o di uno stilista a seconda di quali dei due centri risulti avere guadagnato di più. In altre parole per l'imprenditore l'interpretazione corretta di "mela buona" è quella più conveniente. La cosa non deve sorprendere. Infatti se ci proponiamo, con la logica fuzzy, di studiare il

comportamento razionale dell'uomo nella vita quotidiana, non ci dobbiamo lamentare troppo del fatto che le persone ragionano più per massimizzare l'utilità che per raggiungere la verità.²

Ma l'interpretazione di un concetto vago non è completamente arbitraria poiché in ogni parola è comunque sedimentato un significato non individuale. Nel caso del paradosso del mucchio di grano, indichiamo con M l'insieme dei mucchi di grano e rappresentiamo il concetto di "essere un mucchio piccolo" tramite una funzione $s : M \rightarrow U$. Allora è vero che esistono molti modi di scegliere s e quindi di modellizzare la nozione di mucchio piccolo, tuttavia tale scelta non è completamente arbitraria. Ad esempio $s(x)$ deve essere decrescente rispetto alla cardinalità di x e tendere a zero quando tale cardinalità tende all'infinito. Inoltre s deve assumere valore 1 almeno per i mucchi con meno di cento chicchi. Questa non completa arbitrarietà consente di dire che un linguaggio vago può trasmettere informazione.

In conclusione, anche se non è affatto scontato che la nozione di sottoinsieme fuzzy sia uno strumento totalmente adeguato a rappresentare i concetti vaghi³, in questo libro accettiamo di rappresentare con tale nozione la vaghezza e proviamo a costruirci un po' di matematica.

3. Inclusione, unione, intersezione e complemento

Avendo definito la classe dei sottoinsiemi fuzzy di un dato insieme, come secondo passo dobbiamo definire la relazione di inclusione e le operazioni di unione, intersezione e complemento.

Definizione 3.1. Dati due sottoinsiemi fuzzy s_1 e s_2 di un insieme S , diciamo che s_1 è *incluso* in s_2 , e scriviamo $s_1 \subseteq s_2$, se $s_1(x) \leq s_2(x)$ per ogni $x \in S$.

D'altra parte questa definizione è coerente con il fatto che se X ed Y sono sottoinsiemi di un insieme non vuoto S allora

² Questo esempio mostra anche che la teoria degli insiemi fuzzy come la teoria delle supervalutazioni accetta che ci siano diverse determinazioni di un concetto vago. Tuttavia, a differenza della teoria delle supervalutazioni, tali determinazioni sono a loro volta insiemi fuzzy.

³ Si vedano ad esempio le interessanti considerazioni di S. Termini nel suo "Su alcuni sviluppi delle scienze dell'informazione".

$$X \subseteq Y \Leftrightarrow c_X(x) \leq c_Y(x).$$

Da notare che l'inclusione definita in questo modo è crisp nel senso che due sottoinsiemi fuzzy o risultano inclusi oppure no. Come vedremo, è anche possibile definire inclusioni graduate considerando il connettivo logico di implicazione. In questo caso l'inclusione dipenderebbe dal tipo di logica fuzzy considerato.

Per definire l'intersezione $s_1 \cap s_2$ tra i due sottoinsiemi fuzzy s_1 e s_2 dobbiamo definire in qualche modo il grado $(s_1 \cap s_2)(x)$ con cui un elemento x appartiene a tale intersezione. Tale grado "dipende" ovviamente dal grado di appartenenza di x ad s_1 e dal grado di appartenenza di x ad s_2 , cioè dipende dai valori $s_1(x)$ e $s_2(x)$. Se indichiamo con $\otimes$ la funzione binaria in U che esprime tale dipendenza, possiamo affermare che il grado di appartenenza di x all'intersezione è $s_1(x) \otimes s_2(x)$, cioè porre $(s_1 \cap s_2)(x) = s_1(x) \otimes s_2(x)$. Da osservare che il valore $s_1(x) \otimes s_2(x)$ è una valutazione nella logica a più valori dell'affermazione

$$x \text{ appartiene ad } s_1 \text{ e } x \text{ appartiene ad } s_2,$$

dove le formule atomiche " x appartiene ad s_1 " e " x appartiene ad s_2 " sono valutate con $s_1(x)$ e $s_2(x)$, rispettivamente. Pertanto l'operazione $\otimes$ deve essere vista come l'interpretazione del connettivo logico e ed esistono quindi tante possibili definizioni dell'intersezione quante sono le possibili definizioni di $\otimes$.

Similmente per potere definire l'unione abbiamo bisogno di definire una operazione $\oplus$ in U in modo che $s_1(x) \oplus s_2(x)$ possa essere considerato come il grado di appartenenza di x a $s_1 \cup s_2$. Si tratta quindi di effettuare una valutazione nella logica a più valori dell'affermazione

$$x \text{ appartiene ad } s \text{ oppure } x \text{ appartiene ad } s',$$

Pertanto l'operazione $\oplus$ deve essere vista come una interpretazione in una logica a più valori del connettivo di disgiunzione. Infine per definire il complemento $\neg s_1$ è necessaria una operazione $\sim$ tale che $\sim s_1(x)$ possa essere considerata una valutazione dell'affermazione

$$\text{non è vero che } x \text{ appartiene ad } s.$$

Pertanto $\sim$ sarà l'interpretazione del connettivo $\neg$ di negazione.

Tali considerazioni mostrano che per potere definire le operazioni di intersezione, unione e complemento nella classe U^S dei sottoinsiemi fuzzy di S abbiamo bisogno di avere definito in U tre operazioni $\otimes$, $\oplus$, $\neg$.

Definizione 3.2. Una *struttura di valutazione* in U è una struttura algebrica del tipo $V = \{U, \leq, \otimes, \oplus, \neg, 0, 1\}$ che estende l'algebra di Boole $B_2 = (\{0, 1\}, \leq, \min, \max, 1-x, 0, 1)$.

Quest'ultima condizione è necessaria se vogliamo che la logica fuzzy sia una estensione di quella classica (cosa che viene sempre fatto in letteratura). Ad esempio possiamo considerare la struttura $V = (U, \wedge, \vee, \neg, 0, 1)$ dove, per ogni $x, y \in U$,

$$x \wedge y = \min\{x, y\}, \quad x \vee y = \max\{x, y\}, \quad \neg(x) = 1-x. \quad (3.1)$$

Più in avanti esamineremo quali proprietà è ragionevole che siano verificate da una struttura di valutazione.

Definizione 3.3. Sia S un insieme e V una struttura di valutazione, allora dati due sottoinsiemi fuzzy s_1 e s_2 , definiamo i sottoinsiemi fuzzy $s_1 \cup s_2$, $s_1 \cap s_2$ e $\neg s_1$ ponendo, per ogni $x \in S$

$$(s_1 \cup s_2)(x) = s_1(x) \oplus s_2(x);$$

$$(s_1 \cap s_2)(x) = s_1(x) \otimes s_2(x);$$

$$(\neg s_1)(x) = \sim s_1(x).$$

E' possibile riformulare le definizioni ora date in termini algebrici utilizzando la nozione di *potenza diretta*.⁴

Definizione 3.4. Data una struttura algebrica A di dominio A ed un insieme non vuoto S chiamiamo *potenza diretta con insieme di indici in S* la struttura algebrica A^S che ha come dominio l'insieme A^S delle funzioni di S in A ed in cui ogni operazione n -aria h in A è estesa ad una operazione n -aria h in A^S definita ponendo

$$h(f_1, \dots, f_n)(x) = h(f_1(x), \dots, f_n(x)).$$

Se in A esiste un elemento c che gioca un ruolo particolare (elemento neutro od altro) allora questo elemento viene identificato in A^S con la funzione costantemente uguale a c .

Ad esempio se A coincide con l'anello $\mathbf{R}$ dei numeri reali ed S con l'insieme N dei numeri naturali allora $\mathbf{R}^N$ è la struttura algebrica il cui dominio è l'insieme delle successioni di numeri reali ed in cui sono definite le operazioni di addizione e

⁴I lettori che non amano i formalismi matematici possono tranquillamente saltare questa parte.

moltiplicazione di due successioni e quella di opposto di una successione. L'elemento neutro 0 dell'addizione in $\mathbf{R}$ viene identificato con la successione costantemente uguale a 0. L'elemento neutro 1 della moltiplicazione viene identificato con la successione costantemente uguale ad 1. In questo modo si definisce l'anello delle successioni di numeri reali. Da osservare che la potenza diretta di $\mathbf{A}$ conserva tutte le proprietà di $\mathbf{A}$ che siano esprimibili tramite equazioni. Ad esempio le proprietà commutativa ed associativa della somma e del prodotto in $\mathbf{R}$ vengono ereditate dall'anello delle successioni. La successione costantemente uguale a 0 è elemento neutro rispetto alla somma di successioni. La successione costantemente uguale ad 1 è elemento neutro rispetto alla moltiplicazione.

Se invece consideriamo la potenza diretta di $\mathbf{B}_2$ con insieme di indici in S otteniamo la struttura algebrica $\mathbf{B}_2^S = (\{0,1\}^S, \cap, \cup, -, c_\emptyset, c_S)$ dove $\cap, \cup$ e $-$ sono le operazioni in $\{0,1\}^S$ definite ponendo, per ogni c_1, c_2 in $\{0,1\}^S$,

$$(c_1 \cap c_2)(x) = \min\{c_1(x), c_2(x)\} ;$$

$$(c_1 \cup c_2)(x) = \max\{c_1(x), c_2(x)\} ;$$

$$(-c_1)(x) = 1 - c_1(x).$$

Abbiamo già visto come sia possibile identificare $P(S)$ con $\{0,1\}^S$. Consideriamo ora $P(S)$ come struttura algebrica $\mathbf{P}(S) = (P(S), \cap, \cup, -, \emptyset, S)$, che, come è noto, viene detta *algebra di Boole dei sottoinsiemi di S*. E' immediato verificare il seguente teorema.

Teorema 3.5. L'applicazione $H : P(S) \rightarrow \{0,1\}^S$ definita ponendo $H(X) = c_X$ è un isomorfismo tra l'algebra di Boole $\mathbf{P}(S)$ e la potenza diretta $\mathbf{B}_2^S$.

Analogamente possiamo provare il seguente teorema.

Teorema 3.6. Sia S un insieme e $\mathbf{V}$ una struttura di valutazione, allora l'algebra dei sottoinsiemi fuzzy di S è la struttura algebrica $\mathbf{V}^S = (U^S, \cup, \cap, -, c_\emptyset, c_S)$ che si ottiene come potenza diretta di $\mathbf{V}$ con insieme di indici S . Tale algebra è una estensione della potenza diretta $\mathbf{B}_2^S$ e quindi dell'algebra di Boole $\mathbf{P}(S)$.

Definizione 3.7. La struttura algebrica $\mathbf{V}^S$ viene chiamata *algebra dei sottoinsiemi fuzzy di S*.

Naturalmente le proprietà dell'algebra dei sottoinsiemi fuzzy di un dato insieme dipendono dalle proprietà della struttura di valutazione considerata, cioè dal tipo di operazioni $\otimes$, $\oplus$ e $\sim$ definite in U . Infatti vale la seguente proposizione.

Teorema 3.8. Tutte le equazioni e disequazioni che valgono in V valgono anche in V^S . Le equazioni e disequazioni soddisfatte da V^S sono soddisfatte anche in B_2 e quindi nella logica classica.

4. Norme, conorme e negazioni

A mio parere tentare di individuare una particolare struttura di valutazione che rappresenti la “vera” logica fuzzy non ha molto senso. Infatti, se ci si riferisce, come è ragionevole, all'intervallo $[0,1]$, allora per questioni di cardinalità, non può valere un teorema di completezza funzionale come invece avviene nella logica proposizionale classica. Quindi è inevitabile che dovendo rappresentare situazioni nuove si debba essere pronti ad usare operazioni nuove non definibili da quelle già fissate.

Nel seguito indichiamo alcune classi di strutture di valutazione che sembrano di maggiore interesse.

Come abbiamo già detto, una condizione irrinunciabile è che le operazioni $\otimes$, $\oplus$ e $\sim$ quando siano ristrette a $\{0,1\}$ coincidano con le usuali interpretazioni dei connettivi logici della logica classica. Inoltre supponiamo ad esempio di volere che l'operazione di intersezione verifichi le proprietà

- $(s_1 \cap s_2) \cap s_3 = s_1 \cap (s_2 \cap s_3)$ (associativa),
- $s_1 \cap s_2 = s_2 \cap s_1$ (commutativa),
- $s \cap c_S = s$ (S è l'elemento neutro),
- $s_1 \subseteq s_2 \Rightarrow s_2 \cap s \subseteq s_1 \cap s$ (monotonia).

Se si tiene conto della proposizione 3.8, appare evidente che è necessario considerare la seguente classe di operazioni.

Definizione 4.1. Un operazione $\otimes$ in U che verifichi le proprietà :

- $(x \otimes y) \otimes z = x \otimes (y \otimes z)$ (associativa),
- $x \otimes y = y \otimes x$ (commutativa),
- $x \otimes 1 = x$ (1 è l'elemento neutro),
- $x \leq y \Rightarrow x \otimes z \leq y \otimes z$ (monotonia),

dove x , y e z sono elementi di U prende il nome di *norma triangolare* (si veda [Schweizer e Sklar 1983]).

In letteratura gli esempi principali di norme triangolari si ottengono ponendo

$$x \otimes y = \min\{x, y\} \text{ (norma di Zadeh),}$$

$$x \otimes y = x \cdot y \text{ (norma della moltiplicazione usuale)}$$

$$x \otimes y = |x+y-1|_n \text{ (norma di Łukasiewicz)}$$

dove con $|\cdot|_n : R \rightarrow U$ denoto la *funzione di normalizzazione* definita ponendo $|x|_n = x$ se $x \in U$, $|x|_n = 0$ se $x < 0$ e $|x|_n = 1$ se $x > 1$.

La seguente proposizione evidenzia alcuni aspetti di tali norme il cui interesse apparirà nel seguito. La nozione di potenza n -esima di un elemento γ di U rispetto ad una norma triangolare $\otimes$ viene definita al solito modo tramite le equazioni

$$\gamma^1 = \gamma, \gamma^{p+1} = \gamma^p \otimes \gamma$$

Proposizione 4.2. Sia $\gamma \in U$ diverso da 0 ed 1, allora

- se $\otimes$ è l'usuale prodotto allora, per ogni $n \geq 1/(1-\gamma)$, $\gamma^n \neq 0$ e $\lim_{n \rightarrow \infty} \gamma^n = 0$;
- se $\otimes$ è la norma di Łukasiewicz allora $\gamma^n = 0$ e $\lim_{n \rightarrow \infty} \gamma^n = 0$;
- se $\otimes$ è il minimo allora $\gamma^n = \gamma \neq 0$ per ogni n e $\lim_{n \rightarrow \infty} \gamma^n \neq 0$.

Una ampia classe di norme triangolari viene data dalla seguente proposizione di cui omettiamo la dimostrazione (si veda [Valverde 1985]). Nel seguito, detto l un elemento di $[0, \infty]$ chiameremo *somma troncata a livello l* l'operazione $+_l$ definita in $[0, l]$ ponendo

$$x+_ly = \min\{x+y, l\}.$$

Da notare che nel seguito consideriamo la semiretta estesa $[0, \infty]$ in cui si assume che ∞ sia maggiore di tutti i numeri in $[0, \infty)$ e che $x+\infty = \infty+x = \infty$ per ogni $x \in [0, \infty]$.

Proposizione 4.3. Sia $f : [0, 1] \rightarrow [0, +\infty]$ una funzione continua, strettamente decrescente e tale che $f(1) = 0$. Posto $l = f(0)$, definiamo l'operazione $\otimes$ ponendo, per ogni x, y in $[0, 1]$,

$$x \otimes y = f^{-1}(f(x)+_l f(y))$$

Allora $\otimes$ è una norma triangolare continua ed f è un isomorfismo tra la struttura $(U, \otimes)$ e la struttura $([0, l], +_l)$.

La funzione f prende il nome di *generatore additivo* di $\otimes$. Se ad esempio $f(x) = 1-x$ allora nel caso se $x+y-1 \geq 0$ risulta che $x \otimes y = 1-$

$(1-x+1-y) = x+y-1$ mentre $x \otimes y = 0$ altrimenti. Pertanto $\otimes$ coincide con la norma di Łukasiewicz. Se $f(x) = -\text{Log}(x)$ allora $x \otimes y = 10^{\text{Log}(x)+\text{Log}(y)} = x \cdot y$ e quindi $\otimes$ coincide con l'usuale prodotto.

Le proprietà che vorremmo fossero verificate dall'unione di due sottoinsiemi fuzzy sono le stesse dell'intersezione salvo il fatto che sarebbe desiderabile che l'insieme vuoto, che noi identifichiamo con $c_\emptyset$, funga da elemento neutro.

$$s \cup c_\emptyset = s$$

Per ottenere questo dobbiamo supporre che 0 sia elemento neutro dell'operazione $\oplus$. Pertanto sembra ragionevole la seguente definizione.

Definizione 4.4. Chiamiamo *co-norma triangolare* una operazione $\oplus : U \times U \rightarrow U$ tale che:

- $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ (associativa),
- $x \oplus y = y \oplus x$ (commutativa),
- $x \oplus 0 = x$ (0 è l'elemento neutro),
- $x \leq y \Rightarrow x \oplus z \leq y \oplus z$ (monotonia),

dove x, y e z sono elementi di U .

In realtà usualmente le co-norme vengono definite a partire da una data norma ed una data negazione. Infatti si vuole conservare la legge di De Morgan

$$s_1 \cup s_2 = \neg(\neg s_1 \cap \neg s_2)$$

e per ottenere questo abbiamo bisogno di interpretare il connettivo di negazione e poi imporre che le operazioni $\otimes$, $\oplus$ e $\sim$ siano legate dalla equazione $x \oplus y = \neg(\neg x \otimes \neg y)$. Ciò significa che è opportuno concentrare la nostra attenzione sulla interpretazione $\sim$ del connettivo logico $\neg$ di negazione.

Ora, se si vuole ottenere la legge di doppia negazione, cioè la proprietà $\neg(\neg(s)) = s$, allora la funzione $\sim$ deve verificare la proprietà

$$\sim(\neg(x)) = x \quad (\text{doppia negazione}).$$

Inoltre è evidente che $\sim$ deve essere decrescente

$$x \leq y \Rightarrow \neg(x) \geq \neg(y) \quad (\text{decrescenza}).$$

Questo perché più un'asserzione A è vicina al vero più la sua negata deve essere vicina al falso. Infine, poiché vogliamo che $\neg$ estenda la negazione classica, dobbiamo richiedere anche che

$$\neg(0) = 1 \quad ; \quad \neg(1) = 0.$$

Definizione 4.5. Chiamiamo *buona negazione* una involuzione decrescente che scambi 0 con 1, cioè una funzione $\sim : U \rightarrow U$ tale che

$$\begin{aligned}\sim(\sim x) &= x && \text{(doppia negazione)} \\ x \leq y &\Rightarrow \sim(x) \geq \sim(y) && \text{(decrecenza)} \\ \sim(0) &= 1 \quad ; \quad \sim(1) = 0. && \text{(estensione)}\end{aligned}$$

Una funzione molto semplice che verifica tutte queste condizioni si ottiene ponendo

$$\sim(x) = 1-x.$$

Tale funzione non è unica, ad esempio le condizioni sono verificate anche dalla funzione

$$\sim(x) = \sqrt{1-x^2}$$

che si ottiene proiettando $1-x$ tramite la funzione radice quadrata. D'altra parte non è difficile caratterizzare tutte le buone negazioni. Infatti vale la seguente proposizione.

Proposizione 4.6. Una funzione $\sim : U \rightarrow U$ è una buona negazione se e solo se è una funzione decrescente e simmetrica rispetto alla bisettrice del primo quadrante. L'unica funzione lineare che verifica tali condizioni è la funzione $\sim(x) = 1-x$.

Nel seguito supporremo sempre che $\sim$ sia lineare, cioè che coincida con $1-x$.

Una volta che si sia definita una norma ed una buona negazione, si perviene alla seguente definizione.

Definizione 4.7. Chiamiamo *co-norma associata alla norma $\otimes$ ed alla negazione $\sim$* , l'operazione $\oplus$ definita dall'equazione

$$x \oplus y = \sim(\sim x \otimes \sim y). \quad (4.1)$$

In altri termini la co-norma associata a $\otimes$ è la proiezione di $\otimes$ tramite $\sim$.

Tale definizione deve essere giustificata dalla seguente proposizione.

Proposizione 4.7. L'operazione $\oplus$ definita tramite (4.1) è una co-norma triangolare.

Ad esempio:

- La co-norma associata al minimo $\wedge$ è il massimo $\vee$.
- La co-norma associata al prodotto è definita ponendo:

$$x \oplus y = x + y - xy.$$

- La co-norma associata alla norma di Łukasiewicz è definita ponendo,

$$x \oplus y = |x + y|_n.$$

Nota. Abbiamo legato le definizioni di intersezione ed unione di due insiemi fuzzy alla interpretazione dei connettivi logici e quindi alle norme ed alle co-norme scelte. Tuttavia nella teoria degli insiemi classici è possibile anche dare una definizione di tipo “reticolare” per tali operazioni. Infatti possiamo definire l’intersezione di due sottoinsiemi X ed Y come il più grande dei sottoinsiemi contenuti sia in X che in Y , cioè come l’estremo inferiore tra X ed Y . Similmente possiamo definire l’unione come il più piccolo dei sottoinsiemi contenenti sia X che Y , cioè come l’estremo superiore tra X ed Y . Se si adotta questo punto di vista e si vogliono definire l’intersezione e l’unione di due sottoinsiemi fuzzy s_1 ed s_2 come estremo inferiore ed estremo superiore nel reticolo $(U^S, \subseteq)$, allora si ottengono due definizioni che non sono legate alle norme triangolari. Più precisamente, si ottengono le seguenti definizioni

$$(s_1 \cap s_2) = \text{Min}\{s_1(x), s_2(x)\} ; (s_1 \cup s_2) = \text{Max}\{s_1(x), s_2(x)\}$$

che si riferiscono quindi alla norma di Zadeh ed alla co-norma associata.

5. E’ l’uso dei connettivi ad imporre la normalizzazione

La considerazione dei connettivi logici e delle relative operazioni insiemistiche permette di porre in evidenza perché sia necessario *normalizzare* l’informazione relativa a predicati vaghi utilizzando per tutti i predicati una stessa struttura di valutazione. Infatti si pone il seguente problema:

Perché considerare “alto”, “ricco” predicati vaghi e non riferirsi direttamente all’altezza misurata in centimetri o alla ricchezza misurata in milioni ?

In fondo l’attività di misurare potrebbe essere vista proprio come una risposta scientifica alla considerazione di proprietà che

possono essere attribuite ad un oggetto con gradi diversi. Per fare ancora un esempio, la nozione di velocità potrebbe essere proprio la precisazione della proprietà “veloce”, la nozione di peso una precisazione della proprietà “pesante” e così via⁵.

Per rispondere a tale questione consideriamo ad esempio l’informazione rappresentata dalla seguente tabella:

matricola	pagato tasse	altezza	ricchezza
23222	0	1.65m	3 ml
32122	1	1.90m	0.3 ml
22223	0	1.55m	0.003 ml
...	...	...	...
32454	1	1.50m	0.8 ml

e supponiamo che tale tabella sia stata “tradotta” nella seguente in cui sono coinvolti solo valori di verità:

matricola	pagato tasse	altezza	ricchezza
23222	0	0.6	1
32122	1	1	0.8
22223	0	0	0
...	...	...	...
32454	1	0.3	0.6

Indiscutibilmente la seconda tabella appare rappresentare l’informazione in modo peggiore della prima. Ciò se non altro per una certa arbitrarietà insita nella normalizzazione. Tuttavia proviamo ad interrogare tali tabelle chiedendo : “*dammi l’elenco di tutti gli studenti alti e non ricchi*”. In tale caso la prima tabella non è affatto utile. Solo la seconda tabella, insieme ad una fissata interpretazione dei connettivi, può fornire una risposta. Ad esempio se $\otimes$ è il minimo e $\sim(x) = 1-x$ avremo la tabella:

⁵ Si veda in proposito il paragrafo 13 del quarto capitolo in cui si parla della teoria medioevale delle qualità graduate.

matricola	pagato tasse	alto e non ricco
23222	0	0
32122	1	0.2
22223	0	0
...	...	...
32454	1	0.3

In altri termini l'informazione fornita da nozioni vaghe può essere elaborata solo dopo che un *processo di normalizzazione* abbia ridotto i diversi pezzi di informazione ad un ambiente comune che nel nostro caso è l'intervallo U . Tale processo di normalizzazione si esprime, nel caso della ricchezza, tramite una opportuna funzione $n_r : N \rightarrow U$, nel caso dell'altezza tramite una opportuna funzione $n_a : [0,2] \rightarrow U$. Naturalmente il fatto che sia necessario normalizzare non vuole dire che sia facile normalizzare. Infatti è evidente che parole come “alto” o “ricco” possono essere interpretate in infiniti modi e questo perché un concetto vago non solo è vago ma è anche ambiguo. D'altra parte se si accetta che la nozione di “alto” può essere rappresentata da diversi insiemi fuzzy, è anche naturale accettare che alcuni insiemi fuzzy possono essere considerati buone rappresentazioni ed altri meno buone. Ma a sua volta la nozione di “buona rappresentazione” è una nozione vaga e quindi si dovrebbe interpretare la nozione di “grande” con una classe fuzzy di insiemi fuzzy.

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 alto che cambia continuamente. La stessa interpretazione sembra potersi evolvere nel tempo in un gioco linguistico difficilmente dominabile dal formalismo matematico. Come abbiamo già osservato, stranamente la traduzione in inglese o in francese della parola “grande” non è poi tanto ambigua. Si può affermare che “big” è la traduzione esatta di “grande”. Si pone allora il problema:

se non esiste una interpretazione univoca di un concetto vago come si può dire che tale concetto è traducibile in modo univoco, come di fatto avviene?

Sembrerebbe che la parola grande corrisponda non ha un preciso “insieme di oggetti grandi” ma piuttosto ad un meccanismo linguistico complesso e flessibile. Quando si traduce la parola grande in realtà si traduce un tale meccanismo linguistico.

6. Altri connettivi: implicazione, equivalenza, negazione

Un altro importante connettivo logico è l'implicazione che ovviamente deve essere interpretato da una opportuna operazione binaria $\rightarrow$ in U . Per capire come deve essere fatta tale operazione teniamo conto del fatto che, più che un significato semantico, il connettivo di implicazione ha un ruolo all'interno dell'apparato dimostrativo con particolare riferimento alla regola del Modus Ponens. Ora, supponiamo di sapere che

- l'asserzione A vale con grado a

- l'asserzione B vale con grado b

allora dobbiamo assegnare alla formula $A \rightarrow B$ un valore x che renda il Modus Ponens una regola di inferenza valida. Dobbiamo tenere conto del fatto che, come abbiamo visto nel esaminare il paradosso del mucchio di grano, tale regola agisce al modo seguente:

ESSENDO A è provata con grado a

E $A \rightarrow B$ è provata con grado x

ALLORA

B è provata con grado $a \otimes x$

Ora, come vedremo, il grado con cui è provata una formula viene interpretato come un limite inferiore al valore di verità della formula. Allora, perché il grado $a \otimes x$ sia corretto deve essere $a \otimes x \leq b$. In altre parole una valutazione x della formula $A \rightarrow B$ rende il Modus Ponens corretto se e solo se $a \otimes x \leq b$. D'altra parte per non perdere informazione, conviene attribuire ad x il massimo valore possibile per cui ciò accada. Si perviene allora alla seguente definizione

$$a \rightarrow b = \text{Sup}\{x \in U : a \otimes x \leq b\}$$

che funzionerebbe bene se tale estremo superiore fosse un massimo, cioè se fosse $a \otimes (a \rightarrow b) \leq b$. Ciò accade se $\otimes$ è una norma triangolare continua. Infatti, essendo la funzione $f(x) = a \otimes x$ continua e crescente, l'insieme $\{x \in U : a \otimes x \leq b\} = f^{-1}([0, b])$ sarà un intervallo chiuso. Perveniamo allora alla seguente definizione.

Definizione 6.1 Data una norma triangolare continua $\otimes$, chiamiamo *implicazione associata a $\otimes$* o *residuo di $\otimes$* , l'operazione binaria definita da:

$$a \rightarrow b = \text{Max}\{x \in U : a \otimes x \leq b\}.$$

Ad esempio, se $\otimes$ coincide con il minimo allora abbiamo che

$$a \rightarrow b = 1 \text{ se } a \leq b$$

$$a \rightarrow b = b \text{ se } a > b.$$

Se $\otimes$ coincide con il prodotto allora

$$a \rightarrow b = 1 \text{ se } a \leq b$$

$$a \rightarrow b = |b/a|_n \text{ se } a > b$$

Se $\otimes$ coincide con la norma di Łukasiewicz allora

$$a \rightarrow b = 1 \text{ se } a \leq b$$

$$a \rightarrow b = b - a + 1 \text{ se } a > b.$$

Si dimostra che il residuo verifica le seguenti proprietà.

Proposizione 6.2. L'implicazione $\rightarrow$ associata ad una norma triangolare continua verifica le seguenti condizioni:

i) $a \rightarrow b = 1$ se e solo se $a \leq b$,

ii) $\rightarrow$ è decrescente rispetto alla prima coordinata e crescente rispetto alla seconda

iii) $a \otimes x \leq b \Leftrightarrow x \leq a \rightarrow b$

iv) $(a \rightarrow b) \otimes (b \rightarrow c) \leq a \rightarrow c$.

Negli approcci più interessanti alla logica fuzzy si parte da una norma triangolare continua e poi si definisce l'implicazione come residuo. Poi si passa a definire la negazione e l'equivalenza al modo seguente.

Definizione 6.3. Prende il nome di *negazione associata alla norma triangolare $\otimes$* , la funzione definita ponendo

$$\sim(a) = a \rightarrow 0 = \text{Max}\{x \in U : a \otimes x = 0\}.$$

Prende il nome di *equivalenza associata a $\otimes$* l'operazione binaria $\leftrightarrow$ definita ponendo $x \leftrightarrow y$ uguale a $(x \rightarrow y) \otimes (y \rightarrow x)$.

Ad esempio, se $\otimes$ coincide con il minimo allora abbiamo che

$$\sim a = 1 \text{ se } a = 0$$

$$\sim a = 0 \text{ se } a > 0$$

$$a \leftrightarrow b = 1 \text{ se } a = b$$

$$a \leftrightarrow b = a \wedge b \text{ se } a \neq b$$

Se $\otimes$ coincide con il prodotto allora

$$\sim a = 1 \quad \text{se } a = 0$$

$$\sim a = 0 \text{ se } a > 0$$

$$a \leftrightarrow b = 1 \text{ se } a = b$$

$$a \leftrightarrow b = \min\{a, b\} / \max\{a, b\} \text{ se } a \neq b$$

Se $\otimes$ coincide con la norma di Łukasiewicz allora

$$\sim a = 1 - a$$

$$a \leftrightarrow b = 1 - |a - b|.$$

La seguente proposizione elenca alcune delle proprietà di tali operazioni.

Proposizione 6.4. La funzione $\sim$ verifica le seguenti condizioni

$$i) \sim a = 1 \text{ se e solo se } a = 0,$$

$$ii) \sim \text{ è decrescente}$$

$$iii) a \otimes x = 0 \Leftrightarrow x \leq \sim a$$

La funzione $\leftrightarrow$ verifica le seguenti:

$$i) a \leftrightarrow b = 1 \text{ se e solo se } a = b,$$

$$ii) a \leftrightarrow b = b \leftrightarrow a,$$

$$iii) (a \leftrightarrow b) \otimes (b \leftrightarrow c) \leq a \leftrightarrow c.$$

7. Interpretazioni fuzzy

Passiamo ora occuparci della semantica, cioè tentare di precisare che cosa significhi interpretare un linguaggio che accetti nozioni vaghe. Per semplificare le cose supporremo che il linguaggio per la logica del primo ordine fuzzy non sia diverso da quello della logica classica.

Definizione 7.1. Si chiama *alfabeto* un qualunque insieme finito A . Chiamiamo *parola in A* ogni n -pla $(a_1, \dots, a_n)$ di elementi di A . Indichiamo con A^+ l'insieme di tutte le parole in A e chiamiamo *linguaggio formale* un sottoinsieme di A^+ .

In accordo con l'uso comune, indicheremo con $a_1 \dots a_n$ la parola $(a_1, \dots, a_n)$. Tipico esempio di alfabeto è l'alfabeto Morse, costituito da punto e linea, o l'insieme delle lettere che sono su di una tastiera del computer.

Definizione 7.2. Un *alfabeto per la logica del primo ordine* è un insieme costituito, oltre che dagli usuali connettivi logici $\wedge, \vee, \neg$,

$\Rightarrow$, dalla virgola, dalle parentesi (,) e dai quantificatori $\forall, \exists$, anche da:

- un insieme non vuoto C di *costanti*,
- un insieme Rel di simboli per *relazioni*,
- un insieme Op di simboli per *operazioni*,
- una funzione $ar : (Rel \cup Op) \rightarrow N$ detta *arità*.
- una successione $x_1, x_2, \dots$ di simboli chiamati *variabili*.

Se $s \in Rel \cup Op$ ed $ar(s) = n$, allora diremo che s è *n-aria*.

Il linguaggio delle formule viene definito al modo seguente.

Definizione 7.3. Chiameremo *termine* ogni parola che si possa costruire secondo le seguenti regole:

- ogni variabile è un termine
- ogni costante è un termine
- se $t_1, \dots, t_n$ sono termini ed h è il nome di una operazione n -aria allora $h(t_1, \dots, t_n)$ è un termine.

In altre parole un termine è la descrizione di un algoritmo che utilizza solo le operazioni e le costanti che appartengono al dato linguaggio. Nella terza regola abbiamo utilizzato la notazione prefissa in cui il nome della funzione viene posto all'inizio del termine. Tuttavia per i simboli di operazioni binarie si preferisce di solito utilizzare la notazione infissa che consiste nell'applicare la seguente regola:

- se t_1, t_2 sono termini e $\diamond$ è un simbolo per una operazione binaria, allora $(t_1)\diamond(t_2)$ è un termine.

Ad esempio consideriamo un linguaggio come quello della teoria degli anelli in cui sono presenti i nomi di operazioni $+$, $\cdot$, $-$ e le costanti $0, 1$. Allora i polinomi di una o più variabili sono esempi tipici di termini. Sono ancora esempi di termini espressioni del tipo,

$$(1+1)\cdot(x_2+-(x_1)), -((x_1+1)\cdot(1+0)).$$

Naturalmente nel linguaggio comune dei matematici si usa un modo più agile di quello della definizione 7.1 per denotare un termine. Ad esempio si utilizzano simboli come x, y, z per denotare le variabili e si omettono le parentesi quando non sembrano necessarie. Con tali semplificazioni i due termini indicati sopra possono essere rappresentati, rispettivamente, con

$2 \cdot (y-x)$, $-(x+1) \cdot (1+0)$. Inoltre si utilizzano scritture del tipo x^2 , x^3 , x^4 ... per denotare $x \cdot x$, $(x \cdot x) \cdot x$, $((x \cdot x) \cdot x) \cdot x$...

Definizione 7.4. Una *formula atomica* è una espressione del tipo $r(t_1, \dots, t_n)$ dove r è un simbolo per una relazione n -aria e $t_1, \dots, t_n$ sono termini.

Per le formule atomiche costruite con un simbolo per relazione binaria i matematici preferiscono di solito la notazione infissa. Pertanto scriveranno $x^2 + 2 = 0$ e $x^3 + y^2 \leq x - 2 \cdot z$ invece che $=(x^2 + 2, 0)$ e $\leq(x^3 + y^2, x - 2 \cdot z)$, rispettivamente.

Definizione 7.5. Un linguaggio del primo ordine è l'insieme F delle parole in un alfabeto del primo ordine definito assumendo che:

- ogni formula atomica è una formula,
- se α_1, α_2 sono formule, allora $(\alpha_1) \wedge (\alpha_2)$, $(\alpha_1) \vee (\alpha_2)$ e $(\alpha_1) \rightarrow (\alpha_2)$ sono formule,
- se α è una formula, allora $\neg(\alpha)$ è una formula,
- se α è una formula e x_i è una variabile, allora $\forall x_i(\alpha)$ e $\exists x_i(\alpha)$ sono formule.

Una formula è *ground* se in essa non compaiono variabili o quantificatori. Un atomo ground viene chiamato *fatto*. E' evidente che una formula ground è ottenuta componendo fatti tramite i connettivi logici. Le nozioni di *occorrenza libera di una variabile* in una formula, è definita come al solito. Una formula è detta *universale* se è del tipo $\forall x_1, \dots, \forall x_n(\alpha)$ con α priva di quantificatori. Scriveremo $t(x_1, \dots, x_n)$ e $\alpha(x_1, \dots, x_n)$ per sottolineare che le variabili che compaiono in t e α sono tra $x_1, \dots, x_n$.

Per interpretare un linguaggio del primo ordine nella logica classica si fissa un insieme non vuoto D chiamato *dominio*. Inoltre ai nomi propri (le costanti) presenti nel linguaggio vengono associati elementi del dominio, ai simboli che denotano proprietà vengono associati sottoinsiemi di D , ai simboli che denotano relazioni vengono associate relazioni in D . Pertanto in logica classica una interpretazione è una coppia (D, I) dove D è un insieme ed I è una funzione che associa:

- ad ogni nome di relazione n -aria r una relazione n -aria $I(r) \subseteq D^n$

- ad ogni nome di operazione n -aria h una operazione n -aria $I(h) : D^n \rightarrow D$
- ad ogni nome di costante c , un elemento $I(c) \in D$.

Non è difficile estendere un tale modo di procedere alla logica fuzzy. E' sufficiente assumere che i simboli rappresentanti proprietà o relazioni siano interpretati con sottoinsiemi fuzzy o con relazioni fuzzy.

Definizione 7.6. Una *interpretazione fuzzy* di un linguaggio del primo ordine è una coppia (D, I) dove D è un insieme non vuoto chiamato *dominio* ed I è una funzione che associa ad ogni costante c un elemento $I(c)$ di D , ad ogni simbolo per operazione n -aria h una operazione n -aria $I(h) : D^n \rightarrow D$ ed ad ogni simbolo di relazione n -aria r una relazione n -aria fuzzy $I(r) : D^n \rightarrow U$.

A volte chiamerò anche *struttura fuzzy* una interpretazione fuzzy e dirò che $I(r)$ è l'*estensione* del predicato r . Ad esempio in una situazione in cui sono presenti alcune persone chiamate *Mario*, *Luigi* e *Maria* ed in cui si considerino dei predicati come *amico*, *giovane*, allora una interpretazione sarà data da un dominio D costituito dall'insieme delle persone di un dato posto, da tre persone particolari $I(\text{Mario})$, $I(\text{Luigi})$ e $I(\text{Maria})$ di cui conosco il nome, da un insieme fuzzy $I(\text{giovane}) : D \rightarrow U$ che è l'estensione del predicato *giovane* e da una relazione fuzzy $I(\text{amico}) : D \times D \rightarrow U$ che è l'estensione del predicato *amico*.

8. Semantica vero-funzionale

Per potere definire il valore di verità di una asserzione rispetto ad una data interpretazione dobbiamo prima definire che cosa sia l'interpretazione di un termine. Procediamo esattamente come nella logica classica.

Definizione 8.1. L'interpretazione di un termine t le cui variabili siano comprese tra $x_1, \dots, x_n$ è la funzione n -aria $I(t) : D^n \rightarrow D$ definita ponendo

- $I(t)(d_1, \dots, d_n) = d_i$ se t coincide con x_i
- $I(t)(d_1, \dots, d_n) = I(c)$ se t coincide con la costante c .
- $I(t)(d_1, \dots, d_n) = I(h)(I(t_1)(d_1, \dots, d_n), \dots, I(t_p)(d_1, \dots, d_n))$ se $t = h(t_1, \dots, t_p)$.

Pertanto se un termine è una descrizione di un algoritmo algebrico, la sua interpretazione è la funzione descritta da tale algoritmo.

Detto questo, per ogni formula α le cui variabili libere sono tra $x_1, \dots, x_n$, e per ogni $d_1, \dots, d_n$ vogliamo definire il valore $I(\alpha, d_1, \dots, d_n)$ che prende il nome di *valore di verità di α nell'interpretazione (D, I) quando le variabili libere in α siano interpretate con $d_1, \dots, d_n$* .

Definizione 8.2. Data una interpretazione (D, I) sia $\alpha(x_1, \dots, x_n)$ una formula le cui variabili libere o vincolate siano tra $x_1, \dots, x_n$. Allora definiamo il valore $I(\alpha, d_1, \dots, d_n)$ ponendo:

- $I(r(t_1, \dots, t_p), d_1, \dots, d_n) = I(r)(I(t_1)(d_1, \dots, d_n), \dots, I(t_p)(d_1, \dots, d_n))$,
- $I((\alpha_1 \wedge \alpha_2), d_1, \dots, d_n) = I(\alpha_1, d_1, \dots, d_n) \otimes I(\alpha_2, d_1, \dots, d_n)$,
- $I((\alpha_1 \vee \alpha_2), d_1, \dots, d_n) = I(\alpha_1, d_1, \dots, d_n) \oplus I(\alpha_2, d_1, \dots, d_n)$,
- $I((\alpha_1 \Rightarrow \alpha_2), d_1, \dots, d_n) = I(\alpha_1, d_1, \dots, d_n) \rightarrow I(\alpha_2, d_1, \dots, d_n)$,
- $I(\neg(\alpha), d_1, \dots, d_n) = \sim(I(\alpha, d_1, \dots, d_n))$,
- $I(\forall x_i(\alpha), d_1, \dots, d_n) = \inf_{d \in D} I(\alpha, d_1, \dots, d_{i-1}, d, d_{i+1}, \dots, d_n)$
- $I(\exists x_i(\alpha), d_1, \dots, d_n) = \sup_{d \in D} I(\alpha, d_1, \dots, d_{i-1}, d, d_{i+1}, \dots, d_n)$.

Se α non ha variabili libere allora $I(\alpha, d_1, \dots, d_n)$ non dipende da $d_1, \dots, d_n$ e quindi possiamo scrivere $I(\alpha)$ al posto di $I(\alpha, d_1, \dots, d_n)$. Se α ha come unica variabile libera x_i allora $I(\alpha, d_1, \dots, d_n)$ dipende solo da d_i e quindi possiamo scrivere $I(\alpha, x_i/d_i)$ al posto di $I(\alpha, d_1, \dots, d_n)$. Se α è una formula e $\forall x_1 \dots \forall x_n(\alpha)$ la sua chiusura universale poniamo

$$I(\alpha) = I(\forall x_1, \dots, \forall x_n(\alpha)).$$

La definizione 2 fornisce un modo di valutare le formule che si dice *vero-funzionale* in quanto il valore di verità delle formule composte è funzione dei valori di verità delle componenti.

Possiamo ora dare la nozione di modello di un sistema fuzzy di assiomi, concetto che sarà esteso e sviluppato nel prossimo capitolo.

Definizione 8.3. Consideriamo un insieme fuzzy $v : F \rightarrow U$ di formule, allora una interpretazione fuzzy (D, I) è chiamata *modello di v* se risulta che $I(\alpha) \geq v$ per ogni formula α . In questo caso scriviamo $m \models v$.

Come avviene nella logica classica, l'insieme fuzzy v viene chiamato *insieme fuzzy delle ipotesi* o *insieme fuzzy degli assiomi*. Nel prossimo capitolo torneremo su tale definizione estendendola alle semantiche non necessariamente vero-funzionali. In particolare, se T è un insieme di formule (che come al solito identifichiamo con la relativa funzione caratteristica), una struttura fuzzy (D, I) si dirà *modello del sistema T di assiomi* se $I(\alpha) = 1$ per ogni formula $\alpha \in T$. Nei prossimi paragrafi considereremo vari esempi di modelli fuzzy di teorie classiche.

9. Esempi: preordini ed ordini fuzzy

Per mostrare che cosa si debba intendere per modello fuzzy di una teoria, in questo e nei prossimi paragrafi considereremo alcune classi interessanti di strutture fuzzy. Cominciamo con strutture di tipo puramente *relazionale* e precisamente con i preordini fuzzy. A tale scopo partiamo da un linguaggio che contiene solo un nome di relazione binaria R .

Definizione 9.1. Chiamiamo $\otimes$ -preordine ogni modello fuzzy dei seguenti due assiomi

- A1 $\forall x R(x, x)$ (riflessività)
 A2 $\forall x \forall y \forall z ((R(x, y) \wedge R(y, z)) \Rightarrow R(x, z))$ ($\otimes$ -transitività)

Poiché nel nostro linguaggio esiste un solo nome di relazione, una interpretazione (D, I) di un tale linguaggio è definita da una coppia (D, r) dove D è il dominio ed $r = I(R)$ è la relazione binaria fuzzy utilizzata per interpretare R .

Proposizione 9.2. Una interpretazione fuzzy (D, r) è un $\otimes$ -preordine se e solo se soddisfa le condizioni:

- a1. $r(x, x) = 1$;
 a2. $r(x, y) \otimes r(y, z) \leq r(x, z)$.

I $\otimes$ -preordini crisp coincidono con i preordini. Esempi interessanti di $\otimes$ -preordine sono le relazioni di preferenza. Infatti è naturale accettare che una relazione di preferenza sia graduata in quanto è sensato considerare anche il grado di preferenza che un individuo può avere per un oggetto rispetto ad un altro.

Definizione 9.3. Chiamiamo $\otimes$ -ordine un modello degli assiomi A1 ed A2 e dell'assioma

$$A3 \quad \forall x \forall y ((R(x,y) \wedge R(y,x)) \Rightarrow x=y).$$

E' immediato verificare che tale assioma corrisponde alla condizione

$$a3 \quad (r(x,y) \otimes r(y,x) \neq 0 \Rightarrow x = y).$$

Tuttavia i "fuzzisti" invece che tale condizione preferiscono la seguente, più debole,

$$a3^* \quad (r(x,y) = 1) \wedge (r(y,x) = 1) \Rightarrow x = y.$$

Una formula del primo ordine che corrisponda ad $a3^*$ può essere scritta, ad esempio, in un linguaggio che abbia un connettivo logico vero che è interpretato dalla funzione $vero : U \rightarrow U$ definito dal porre $vero(x) = 1$ se $x = 1$ e $vero(x) = 0$ altrimenti.

Definizione 9.4. Un $\otimes$ -ordine in senso debole è un $\otimes$ -preordine (D,r) tale che

$$A3^* \quad (\text{vero}(R(x,y)) \wedge \text{vero}(R(y,x))) \Rightarrow x = y.$$

E' evidente che i $\otimes$ -ordini crisp coincidono con gli ordini classici. Le inclusioni graduate costituiscono una interessante classe di $\otimes$ -ordini in senso debole.

Definizione 9.5. Sia S un insieme non vuoto, $\otimes$ una norma triangolare continua e $\rightarrow$ il relativo residuo. Allora la $\otimes$ -inclusione o *inclusione graduata associata a* $\otimes$ è la relazione binaria fuzzy i definita in U^S ponendo, per ogni coppia s, t di sottoinsiemi fuzzy di S ,

$$i(s,t) = \text{Inf}\{s(x) \rightarrow t(x) : x \in S\}.$$

Da notare che la relazione di $\otimes$ -inclusione è di tipo logico poiché si ottiene interpretando in una logica a più valori l'espressione

"per ogni x , se x appartiene ad s allora x appartiene a t ."

In ciò si differenzia dall'inclusione graduata vista come probabilità condizionata proposta da Kosko di cui abbiamo parlato alla fine del primo capitolo.

Proposizione 9.6. La relazione fuzzy di $\otimes$ -inclusione è un $\otimes$ -ordine in senso debole.

Da notare che la relazione di $\otimes$ -inclusione si ottiene interpretando in U l'espressione

“per ogni x , se x appartiene ad s allora x appartiene a t .”

10. Esempi: le similarità

Un'altra classe molto importante di strutture fuzzy sono le similarità che corrispondono alle relazioni di equivalenza nella matematica classica.⁶

Definizione 10.1. Chiamiamo $\otimes$ -similarità ogni $\otimes$ -preordine simmetrico, cioè ogni modello degli assiomi A1, A2 e dell'assioma

$$A4 \quad \forall x \forall y ((R(x,y) \Rightarrow R(y,x)) \quad (\text{simmetria}).$$

Ovviamente se (S, e) è una struttura fuzzy allora A4 è verificata se e solo se

$$a4 \quad e(x,y) = e(y,x).$$

Le similarità crisp coincidono con le (funzioni caratteristiche delle) relazioni di equivalenza e quindi prendono il nome anche di *equivalenze fuzzy*.

Possiamo ottenere esempi di similarità nella classe dei sottoinsiemi fuzzy di un dato insieme S .

Definizione 10.3. Sia S un insieme non vuoto, $\otimes$ una norma triangolare continua e $\leftrightarrow$ la relativa equivalenza. Allora la $\otimes$ -uguaglianza tra insiemi fuzzy è la relazione binaria fuzzy eq definita in U^S ponendo, per ogni coppia $s, t \in U^S$

$$eq(s,t) = \text{Inf}\{s(x) \leftrightarrow t(x) : x \in S\}.$$

La definizione di eq è di tipo logico nel senso che si ottiene interpretando in U l'espressione

“per ogni x , x appartiene ad s se e solo se x appartiene a t .”

Non è difficile provare la seguente proposizione.

⁶ Le prime definizioni di similarità sono state date in Zadeh [1971] ($\otimes$ uguale al minimo), in Menger [1951] ($\otimes$ uguale all'usuale prodotto) e in Ruspini [1997] ($\otimes$ uguale alla congiunzione di Łukasiewicz).

Proposizione 10.4. La relazione di $\otimes$ -uguaglianza tra insiemi fuzzy è una $\otimes$ -similarità.

Un'altra interessantissima nozione di similarità è legata al seguente principio degli indiscernibili di Leibnitz.

Principio degli indiscernibili di Leibnitz. Due oggetti sono uguali se e solo se tutto ciò che si può dire di uno può essere detto anche dell'altro.

In altre parole $x = y$ se e solo se vale $A(x) \Leftrightarrow A(y)$ per ogni proprietà A . Se si prende tale principio come un modo per definire l'uguaglianza, e lo si relativizza ad un dato insieme di proprietà che si ritengono rilevanti, allora si ottiene la definizione di una relazione di equivalenza. Ad esempio, riferendoci ad un dato insieme di persone, possiamo concentrarci sul sesso e sull'essere laureati e quindi chiamare equivalenti due persone che sono dello stesso sesso e che sono entrambe laureate oppure che hanno lo stesso sesso e sono entrambe non laureate. Se indichiamo con $M(x)$ e $L(x)$ le proprietà di essere maschio e di essere laureato, possiamo porre, per definizione, $x \equiv y$ se e solo se vale la formula

$$(M(x) \Leftrightarrow M(y)) \wedge (L(x) \Leftrightarrow L(y)).$$

Riprendendo tali considerazioni in un ambiente fuzzy, otteniamo la seguente definizione.

Teorema 10.5. Sia $(s_i)_{i \in I}$ una famiglia di sottoinsiemi fuzzy di un insieme non vuoto S . Allora la relazione fuzzy $e : S \times S \rightarrow U$ definita ponendo

$$e(x, y) = \inf_{i \in I} s_i(x) \leftrightarrow s_i(y)$$

è una $\otimes$ -similarità.

11. Una dualità tra similarità e distanze

Esiste uno stretto legame tra le similarità e le distanze. Infatti è intuitivo che più due oggetti sono "vicini" più sono da considerare simili. In termini quantitativi, più la distanza tra due oggetti è piccola più il loro grado di similarità è grande. Ad esempio, riferiamoci alla logica di Łukasiewicz in cui risulta che

$$\lambda \leftrightarrow \mu = 1 - |\lambda - \mu|.$$

Allora se S è un insieme non vuoto, una similarità tra due sottoinsiemi fuzzy s e t si può definire ponendo

$$eq(s,t) = \inf\{1-|s(x)-t(x)| : x \in S\} = 1 - \sup\{|s(x) - t(x)| : x \in S\}.$$

e tale similarità è “duale” della famosa distanza tra funzioni definita ponendo

$$d(s,t) = \sup\{|s(x) - t(x)| : x \in S\}.$$

Per la precisione il legame esiste tra le similarità e le pseudo-metriche estese di cui ricordiamo la definizione.

Definizione 11.1. Chiamiamo *pseudo-metrica estesa* su S ogni funzione $d : S \times S \rightarrow [0, \infty]$ tale che

$$d(x,x) = 0; \quad d(x,y) = d(y,x); \quad d(x,y) \leq d(x,z) + d(z,y)$$

per ogni $x, y, z \in S$.

Le usuali distanze si ottengono supponendo che d non assuma mai il valore infinito e che $d(x,y) = 0$ solo nel caso in cui $x = y$.

Ad esempio, data una famiglia $(s_i)_{i \in I}$ di sottoinsiemi fuzzy di S , la funzione d definita ponendo $d(x,y) = \sup_{i \in I} |s_i(x) - s_i(y)|$ è una pseudo-metrica estesa. Risulta inoltre che

$$1 - d(x,y) = 1 - \sup_{i \in I} |s_i(x) - s_i(y)| = \inf_{i \in I} (1 - |s_i(x) - s_i(y)|).$$

Pertanto il duale di tale pseudo-metrica coincide con la similarità associata alla famiglia $(s_i)_{i \in I}$. I due seguenti teoremi mostrano come si presenta la situazione in generale.

Teorema 11.2. Sia f un generatore additivo di una norma continua $\otimes$. Allora, per ogni pseudo-metrica estesa d , la relazione fuzzy $e_d : S \times S \rightarrow U$ definita ponendo

$$e_d(x,y) = f^1(d(x,y) \wedge f(0))$$

è una $\otimes$ -similarità.

Ad esempio se $\otimes$ è il prodotto usuale, nello spazio metrico euclideo posso definire una relazione di $\otimes$ -similarità ponendo

$$e(x,y) = 10^{-d(x,y)}.$$

Viceversa, ad ogni $\otimes$ -similarità, con $\otimes$ norma generata da un generatore continuo, è possibile associare una pseudo-metrica estesa. Omettiamo la verifica.

Teorema 11.3. Sia f un generatore additivo di una norma $\otimes$ ed e una $\otimes$ -similarità. Allora la funzione $d_e: S \times S \rightarrow [0, \infty]$ definita ponendo

$$d_e(x, y) = f(e(x, y))$$

è una pseudo-metrica estesa.

Dai teoremi 11.2 e 11.3 segue che le nozioni di similarità e di distanza sono interscambiabili.⁷

12. Esempi: strutture algebriche fuzzy

Nel definire la nozione di interpretazione fuzzy abbiamo deciso di interpretare le operazioni in modo classico. Pertanto la nozione di struttura algebrica non si presta ad una estensione in ambito fuzzy. Tuttavia tale estensione può essere fatta per alcune nozioni di carattere algebrico. Ad esempio osserviamo che, dato un gruppo $G = (G, \cdot, ^{-1}, 1)$, un sottogruppo di G è individuato da un sottoinsieme G' di G tale che

$$1 \in G', \quad x, y \in G' \Rightarrow x \cdot y \in G', \quad x \in G' \Rightarrow x^{-1} \in G'.$$

Se denotiamo con A il predicato che esprime l'appartenenza a G' , allora tali condizioni possono essere espresse dalle seguenti formule della logica del primo ordine

$$A(1), \quad A(x) \wedge A(y) \Rightarrow A(x \cdot y), \quad A(x) \Rightarrow A(x^{-1}).$$

⁷ Pertanto il fatto che lo studio degli spazi metrici sia enormemente più esteso dello studio delle similarità può avere solo spiegazioni che vanno oltre la matematica. Ne elenco alcune possibili:

- a) Colpa dello strapotere degli analisti e dei geometri che vogliono emarginare i logici (spiegazione politico-accademica fornita dai logici).
- b) Colpa di Nicola (Oresme) che ha contribuito al passaggio dal qualitativo al quantitativo (spiegazione dei filosofi).
- c) Per fare pagare le tasse ai contadini in Egitto era necessario misurare i confini dei terreni utilizzati (spiegazione degli storici-tributaristi).
- c). Nell'evoluzione del bambino le capacità legate alla percezione dello spazio precedono le capacità logiche (spiegazione della psicologia genetica).
- d) E' vero che dobbiamo amare i nostri *simili* ma è molto più importante mantenere le *distanze* (spiegazione del cattolico alto-borghese).
- e) E' l'accettazione della distanza tra sé e la propria madre che segna il momento centrale dell'evoluzione del neonato (spiegazione di uno psicologo).

Ciò suggerisce la seguente definizione che si riferisce ad un linguaggio che si ottiene estendendo il linguaggio per la teoria dei gruppi tramite un predicato monadico A .

Definizione 12.1. Un $\otimes$ -sottogruppo fuzzy⁸, in breve un *gruppo fuzzy* è un modello della teoria che si ottiene aggiungendo agli assiomi della teoria dei gruppi i seguenti assiomi.

$$A1 \ A(0),$$

$$A2 \ \forall x \forall y (A(x) \wedge A(y) \Rightarrow A(x \cdot y)),$$

$$A3 \ \forall x (A(x) \Rightarrow A(x^{-1})).$$

Un modo più esplicito di definire la nozione di gruppo fuzzy è fornito dalla seguente proposizione la cui dimostrazione è evidente.

Proposizione 12.2. Un $\otimes$ -sottogruppo fuzzy è definito da un gruppo G e da un sottoinsieme fuzzy g di G tale che,

$$i) \ g(1) = 1$$

$$ii) \ g(x \cdot y) \geq g(x) \otimes g(y)$$

$$iii) \ g(x) = g(x^{-1}).$$

Forse gli esempi più interessanti di gruppi fuzzy si ottengono supponendo che G sia il gruppo $\mathcal{Z}(S)$ delle trasformazioni di un insieme non vuoto S in sé. In questo caso per ogni $h \in \mathcal{Z}(S)$, possiamo interpretare $g(h)$ come grado di ammissibilità (oppure il costo, oppure il tempo di realizzazione o altro) della trasformazione h .

Esempio 1. Possiamo definire il gruppo fuzzy delle traslazioni “piccole” del piano al modo seguente. Indichiamo con $\mathcal{Z}(E)$ il gruppo simmetrico sull'insieme E dei punti del piano euclideo.⁹ Se indichiamo con $|\tau|$ l'ampiezza di una traslazione, allora la

⁸ Non è difficile estendere tale definizione a tutti i tipi di strutture algebriche e definire, ad esempio, la nozione di sottomonoido fuzzy, sottoanello fuzzy e così via.

⁹ Gli elementi di tale gruppo sono le funzioni invertibili di E in sé, l'operazione di prodotto coincide con la composizione, l'operazione di inverso con l'usuale inversione di una funzione, l'elemento neutro coincide con l'applicazione identica.

nozione di “essere piccola” sarà interpretabile con un insieme fuzzy del tipo $g(\tau) = n(|\tau|)$ dove $n : [0, \infty] \rightarrow U$ è una sorta di “normalizzazione”. E’ ragionevole supporre che n sia una funzione strettamente decrescente tale che $n(0) = 1$ e $n(\infty) = 0$ e che quindi n sia un generatore di norma triangolare. Ad esempio supponiamo che $n(x) = 10^{-x}$ e che quindi $\otimes$ sia l’usuale prodotto. Allora definiamo $g : \mathcal{X}(E) \rightarrow U$ ponendo, per ogni $\tau \in \mathcal{X}(E)$,

$$g(\tau) = \begin{cases} 10^{-|\tau|} & \text{se } \tau \text{ è una traslazione di ampiezza } |\tau| \\ 0 & \text{se } \tau \text{ non è una traslazione.} \end{cases}$$

E’ evidente che se i è l’applicazione identica allora $g(i) = 1$ e che $s(f) = s(f^{-1})$. Inoltre la disuguaglianza $g(\tau\tau') \geq g(\tau) \otimes g(\tau')$ è immediata nel caso in cui τ oppure τ' non siano traslazioni. Altrimenti basta osservare che per la proprietà triangolare risulta che $|\tau\tau'| \leq |\tau| + |\tau'|$, e quindi che $10^{-|\tau\tau'|} \geq 10^{-|\tau|} \cdot 10^{-|\tau'|}$.

Esempio 2. Definiamo l’insieme fuzzy $g : \mathcal{X}(E) \rightarrow U$ di trasformazioni ponendo

$$g(f) = \begin{cases} 1 & \text{se } f \text{ è una identità} \\ 0.9 & \text{se } f \text{ è una traslazione} \\ 0.8 & \text{se } f \text{ è una isometria che non è una traslazione} \\ 0.5 & \text{se } f \text{ è una similitudine che non è una isometria} \\ 0 & \text{altrimenti} \end{cases}$$

Allora, una semplice verifica per casi mostra che g definisce un $\otimes$ -sottogruppo di $\mathcal{X}(E)$ con $\otimes$ norma di Gödel.

13. Similarità e gruppi fuzzy di trasformazioni

Sappiamo che, dato un insieme non vuoto S , è possibile associare ad ogni gruppo di trasformazioni in S una relazione d’equivalenza e, viceversa, associare ad ogni relazione di equivalenza un gruppo di trasformazioni in S . Vale infatti la seguente proposizione la cui dimostrazione è ovvia.

Proposizione 13.1. Sia S un insieme non vuoto e sia G un sottogruppo del gruppo delle trasformazioni di S , allora possiamo definire in S una relazione di equivalenza $\equiv$, ponendo:

$$x \equiv y \Leftrightarrow \exists g \in G : g(x) = y.$$

Viceversa, sia $\equiv$ una relazione d’equivalenza in S , allora l’insieme

$$G = \{h \in \mathcal{Z}(S) : \forall x \, x \equiv h(x)\}$$

è un sottogruppo del gruppo delle trasformazioni dell'insieme S .

Ad esempio sia G il gruppo delle isometrie del piano euclideo che immaginiamo agire sull'insieme dei triangoli. Allora possiamo chiamare *uguali* due triangoli che possono essere trasformati uno nell'altro tramite una isometria. Viceversa, definita la relazione di uguaglianza tra triangoli in termini di uguaglianza di lati, possiamo definire il gruppo delle isometrie come il gruppo delle trasformazioni che portano ogni triangolo in un triangolo uguale.

Detto questo, non è difficile estendere in ambito fuzzy tale connessione tra gruppi ed equivalenze (si veda [Gerla 2006]c).

Teorema 13.2. Sia e una $\otimes$ -similarità in S con $\otimes$ continua e sia $g : \mathcal{Z}(S) \rightarrow U$ l'insieme fuzzy di trasformazioni definita ponendo:

$$g(h) = \text{Inf}\{e(x, h(x)) : x \in S\}.$$

Allora g è un $\otimes$ -sottogruppo fuzzy del gruppo $\mathcal{Z}(S)$.

Volendo, g può essere interpretato come una sorta di fuzzyficazione del sottogruppo identico di $\mathcal{Z}(S)$. Infatti se tale sottogruppo può essere definito come l'insieme delle trasformazioni che portano ogni elemento x in un elemento uguale, g viene ad essere il l'insieme fuzzy delle trasformazioni che portano ogni elemento x in un elemento "simile".

Proposizione 13.3. Dato un $\otimes$ -sottogruppo fuzzy $g : \mathcal{Z}(S) \rightarrow U$ del gruppo simmetrico $\mathcal{Z}(S)$, sia $e : S \times S \rightarrow U$ la relazione fuzzy definita ponendo

$$e(x, y) = \text{Sup}\{g(h) : h \in \mathcal{Z}(S), h(x) = y\}.$$

Allora e è un $\otimes$ -similarità.

Ricordando che il quantificatore esistenziale è interpretato con l'operatore di estremo superiore, la similarità e si ottiene come interpretazione in ambiente fuzzy dell'affermazione "*esiste una trasformazione in g capace di mutare x in y* ".

14. Quanto è elastica una struttura fuzzy ?

Una struttura fuzzy si ottiene interpretando i predicati vaghi tramite insiemi fuzzy e la scelta del profilo di tali insiemi fuzzy è inevitabilmente soggettiva e “poco stabile”. Essa può infatti fluttuare in conseguenza di cambiamento del contesto, di aggiustamenti pragmatici, di cambio della scala dei valori, di processi di apprendimento, e così via. Possiamo tentare di rappresentare una variazione di un'interpretazione tramite una funzione $k : U \rightarrow U$ che rappresenti in un modo uniforme la variazione subita dai valori di verità. Ad esempio: data una similarità $e : S \times S \rightarrow U$, possiamo considerare una nuova relazione fuzzy $e' : S \times S \rightarrow U$ definita dal porre $e'(x,y) = k(e(x,y))$ e porci il problema di quali siano le proprietà che si conservino nel passare da e ad e' :

siamo sicuri che e' sia ancora una similarità, sia cioè ancora riflessiva, transitiva, e simmetrica ?

Il problema di esaminare che cosa si conserva quando avvengono tali tipi di variazione per una struttura fuzzy è di una certa importanza per la logica fuzzy. Infatti per tale logica, diversamente da quanto avviene nell'approccio di Tarski alla semantica, è necessario confrontarsi con nozioni come “flessibilità”, “apprendimento”, “distanza tra modelli”, “negoiazione dei significati” e così via (si veda ad esempio [Belohlavek 2007]). Ad esempio, afferma J. A. Goguen:

“... si potrebbe dire che la scala delle misure (dei gradi di verità) sia determinata a meno di una trasformazione che conservi l'ordinamento (e forse di 0 ed 1). Nello stesso modo la temperatura è misurata in una scala lineare.... Detto in termini più tecnici, i gradi di appartenenza potrebbero essere determinati a meno di trasformazioni appartenenti al gruppo degli automorfismi della struttura di valutazione.[Goguen 68/69]”.

In [Genito et al. 2011] si tenta di precisare la questione lungo le linee seguenti.

Definizione 14.1. Sia (D,I) una interpretazione fuzzy e $k : U \rightarrow U$ una funzione monotona tale che $k(1) = 1$ e $k(0) = 0$. Chiamiamo k -deformazione di (D,I) l'interpretazione (D,I_k) definita ponendo

$$I_k(\underline{r})(d_1, \dots, d_n) = k(I(\underline{r})(d_1, \dots, d_n)) \quad (14.1)$$

per ogni simbolo r di relazione n -aria e $d_1, \dots, d_n$ in D e lasciando immutata l'interpretazione delle costanti e dei nomi di funzioni.

Non è difficile provare che in una k -deformazione la validità delle formule atomiche si conserva. Più precisamente risulta che:

$$I_k(\alpha, d_1, \dots, d_n) = k(I(\alpha, d_1, \dots, d_n)). \quad (14.2)$$

per ogni formula atomica α e $d_1, \dots, d_n$ in D . La questione si presenta più complessa per le formule non atomiche almeno che non si facciano ipotesi su k . Vale ad esempio il seguente teorema di cui omettiamo la semplice dimostrazione che si effettua per ricorsione sulla complessità delle formule.

Teorema 14.2. Sia k un isomorfismo tra due strutture di valutazione V e V' , allora, per ogni formula α e $d_1, \dots, d_n \in D$,

$$k(I(\alpha, d_1, \dots, d_n)) = I_k(\alpha, d_1, \dots, d_n). \quad (14.3)$$

In particolare per ogni formula chiusa α ,

$$k(I(\alpha)) = I_k(\alpha). \quad (14.4)$$

La proprietà (14.3) può essere espressa anche dicendo che (D, I_k) è *elementarmente equivalente* a (D, I) via k . In particolare, se si deforma un modello fuzzy tramite un automorfismo di V , allora in un certo senso si conservano tutte le proprietà del modello (in accordo con quanto supposto da J. A. Goguen). Sfortunatamente il gruppo degli automorfismi di una struttura di valutazione potrebbe essere molto povero. Ad esempio la struttura più famosa, quella proposta da Łukasiewicz, risulta essere rigida¹⁰.

Proposizione 14.3. La struttura di valutazione V definita dalla norma di Łukasiewicz è rigida.

Dalla rigidità di V sembrerebbe seguire la rigidità della logica fuzzy e ciò è in forte contrasto con l'idea che abbiamo di tale logica. Una strategia differente è quella di ammettere la possibilità di trasformazione non solo dei valori di verità ma anche dell'intera struttura di valutazione, cioè delle operazioni utilizzate per l'interpretazione dei connettivi logici.

¹⁰ Una struttura algebrica viene detta *rigida* se il suo gruppo di automorfismi si riduce all'applicazione identica.

Definizione 14.4. Sia $V = (U, I)$ una struttura di valutazione e $k : U \rightarrow U$ una funzione biettiva che conserva l'ordine e tale che $k(0) = 0$ e $k(1) = 1$, allora denotiamo con V_k la struttura di valutazione $(U, \otimes_k, \oplus_k, \neg_k, 0, 1)$ definita ponendo

$$x \otimes_k y = k^{-1}(k(x) \otimes k(y)) ;$$

$$x \oplus_k y = k^{-1}(k(x) \oplus k(y)) ;$$

$$\neg_k x = k^{-1}(\neg k(x)).$$

Definizione 14.5. Sia (D, I) una struttura fuzzy e sia $k : U \rightarrow U$ una funzione biettiva che conserva l'ordine e tale che $k(0) = 0$ e $k(1) = 1$. Allora la *k-deformazione globale* di (D, I) è l'interpretazione fuzzy I_k vista come valutazione in V_k .

Stante tale definizione la dimostrazione della seguente proposizione è immediata.

Teorema 14.6. Sia (D, I) una struttura fuzzy e sia $k : U \rightarrow U$ una funzione biettiva tale che $k(0) = 0$ e $k(1) = 1$. Allora k è un isomorfismo di V in V_k e quindi la *k-deformazione globale* (D, I_k) di (D, I) è elementarmente equivalente a (D, I) via k .

Tale teorema, che è banale dal punto di vista matematico, mostra tuttavia che in un certo senso non esistono difficoltà a modificare la scala di valutazione in modo arbitrario purché si accetti di modificare anche la struttura di valutazione in modo corrispondente.

Ovviamente sono possibili altri tipi di “manipolazione” delle valutazioni. Possiamo cambiare l'insieme dei valori di verità, possiamo identificare valori di verità vicini e così via. Possiamo anche procedere a “tagli” come vedremo nel prossimo paragrafo.

15. La nozione di taglio e le supervalutazioni

Un caso estremo di deformazione è quello del “taglio” con cui si trasforma brutalmente una struttura fuzzy in una struttura classica. L'idea è che sia possibile fissare una soglia $\lambda \in U$ tale che i gradi di verità superiori o uguali a λ siano accettabili per ritenere vera una affermazione. Cominciamo col definire la nozione di taglio per una relazione fuzzy, nozione che è ben nota in letteratura.

Definizione 15.1. Data una relazione fuzzy $r : S^n \rightarrow U$ e $\lambda \in U$, chiamiamo λ -taglio di r la relazione n -aria

$$C(r, \lambda) = \{x \in S^n : r(x) \geq \lambda\}.$$

Una relazione fuzzy r in S è caratterizzata dalla famiglia dei suoi tagli, infatti, per ogni $x \in S^n$,

$$r(x) = \sup\{\lambda \in U : x \in C(r, \lambda)\}.$$

La famiglia $(C(r, \lambda))_{\lambda \in U}$ dei tagli di r verifica una importante proprietà di continuità.

Definizione 15.2. Dato un insieme S , chiamiamo *catena continua di relazioni n -arie* ogni famiglia $(R_\lambda)_{\lambda \in U}$ di relazioni n -arie in S tale che

$$R_\lambda = \bigcap_{\mu < \lambda} R_\mu.$$

Il seguente teorema permette di identificare le relazioni fuzzy n -arie con le catene continue di relazioni n -arie. La semplice dimostrazione può essere trovata in [Gerla 2001]b.

Teorema 15.3. La famiglia $(C(r, \lambda))_{\lambda \in U}$ dei tagli di una relazione fuzzy n -aria è una catena continua di relazioni n -arie. Viceversa se $(R_\lambda)_{\lambda \in U}$ è una catena continua di relazioni n -arie in S , allora ponendo, per ogni $x \in S^n$,

$$r(x) = \sup\{\lambda \in U : x \in R_\lambda\}$$

si ottiene una relazione n -aria fuzzy la cui famiglia di tagli coincide con $(R_\lambda)_{\lambda \in U}$.

La nozione di taglio può essere data anche in termini di funzioni caratteristiche.

Proposizione 15.4. Sia $\lambda \in U$ e sia $c_\lambda : U \rightarrow \{0,1\}$ definita ponendo

$$c_\lambda(x) = 1 \text{ se } x \geq \lambda ; \quad c_\lambda(x) = 0 \text{ altrimenti.}$$

Allora, data una relazione fuzzy r , la composizione $r_\lambda = c_\lambda \circ r$ è la funzione caratteristica di $C(r, \lambda)$.

Tale proposizione suggerisce la seguente definizione di taglio per una struttura fuzzy.

Definizione 15.5. Data una struttura fuzzy (D, I) , per ogni $\lambda \in U$ chiamiamo λ -taglio di (D, I) la c_λ -deformazione di (D, I) ed indichiamo con (D, I_λ) tale struttura.

In altre parole il λ -taglio di una struttura fuzzy si ottiene tagliando ad altezza λ tutte le relazioni fuzzy di tale struttura. Una struttura fuzzy (D, I) è caratterizzata dalla famiglia dei suoi tagli, infatti, dato un nome di relazione n -aria r , risulta che, per ogni $x \in S^n$,

$$I(r)(x) = \text{Sup}\{\lambda \in U : x \in I_\lambda(r)\}.$$

L'estensione della nozione di catena continua alle interpretazioni fuzzy è ovvia.

Definizione 15.6. Una *catena continua di interpretazioni* è una famiglia $(D, I_\lambda)_{\lambda \in U}$ di modelli con lo stesso dominio, con la stessa interpretazione delle operazioni ed in cui la famiglia di interpretazioni $(I_\lambda(r))_{\lambda \in U}$ di un dato predicato r è una catena continua.

Procedendo in modo analogo a quanto fatto per le relazioni fuzzy, è immediato ora dimostrare il seguente teorema.

Teorema 15.7. Data una struttura fuzzy la famiglia dei suoi tagli è una catena continua. Viceversa ogni catena continua $(D, I_\lambda)_{\lambda \in U}$ determina una struttura fuzzy (D, I) ponendo $I(h) = I_\lambda(h)$ per ogni operazione h e

$$I(r)(d_1, \dots, d_n) = \text{Sup}\{\lambda \in U : (d_1, \dots, d_n) \in I_\lambda(r)\}$$

per ogni relazione r . Pertanto possiamo identificare le strutture fuzzy con le catene continue di interpretazioni classiche.

Supponiamo che la nozione di “determinazione” della teoria delle supervalutazioni si possa formalizzare tramite la nozione di taglio. Allora, per quanto detto sopra, una interpretazione fuzzy è caratterizzata dall'insieme delle sue possibili determinazioni e questo fatto parlerebbe a favore della teoria delle supervalutazioni. Purtroppo esistono problemi per quanto riguarda la nozione di verità poiché si dovrebbe accettare che una proprietà di una struttura fuzzy può essere dichiarata vera solo se tutti i tagli di tale struttura verificano tale proprietà. Per esaminare tale questione consideriamo ad esempio le similarità.

Proposizione 15.8. Supponiamo che $\otimes$ sia la norma di Gödel e sia $r : S \times S \rightarrow U$ una relazione fuzzy in un insieme non vuoto S . Allora r è una $\otimes$ -similarità se e solo se tutti i λ -tagli di r sono relazioni di equivalenza. Pertanto le $\otimes$ -similarità possono essere identificate con le catene continue di equivalenze.

Purtroppo questo riferirsi ai tagli non sempre è possibile. Ad esempio nella maggior parte dei casi le similarità possono modellare nozioni significative del linguaggio comune solo se si coinvolgono norme triangolari diverse da quella di Gödel. Ad esempio supponiamo che e sia una $\otimes$ -similarità che vuole modellizzare la nozione spaziale “essere vicino” e che $a_1, \dots, a_n$ sia una successione di punti tali che $e(a_i, a_{i+1}) = 0.9$ per $i = 1, \dots, n-1$. Allora dalla proprietà transitiva seguirebbe che $e(a_1, a_n) \geq e(a_1, a_2) \otimes \dots \otimes e(a_{n-1}, a_n) = 0.9 \otimes \dots \otimes 0.9$. Se $\otimes$ fosse la norma di Gödel ciò comporterebbe che $e(a_1, a_n) \geq 0.9$ mentre è chiaro che se i punti sono disposti opportunamente lungo una retta, allora $e(a_1, a_n)$ dovrebbe essere prossimo a 0.¹¹ Le stesse considerazioni valgono per le similarità che vogliono, ad esempio, modellare la sinonimia tra parole. Il problema naturalmente risiede nella proprietà transitiva:

la proprietà transitiva per una relazione fuzzy non può essere in generale definita in termini della stessa proprietà per i tagli di tale relazione.

Sembrerebbe allora che la teoria delle supervalutazioni non sia adeguata a rappresentare nemmeno una proprietà tanto fondamentale come quella della transitività.

16. Il Cervino di Varzi, oggetti crisp, oggetti vaghi

Fino ad ora abbiamo parlato della vaghezza come di una possibile caratteristica dei predicati. Tuttavia si pone il problema se possano esistere anche “cose” che siano vaghe. Tale questione è stata posta, ad esempio, da Achille Varzi nell’articolo “I confini del Cervino” [Varzi 2011]. Egli osserva che, con riferimento alla “cosa” chiamata Cervino, esistono zolle di terreno che

¹¹Un esempio di $\otimes$ -similarità che potrebbe modellizzare tale tipo di situazione si ottiene ponendo $e(x, y) = 10^{-d(x, y)}$ ed assumendo che $\otimes$ sia l’usuale moltiplicazione.

appartengono senza alcun dubbio a questo monte ed esistono zolle che sicuramente non gli appartengono. Tuttavia per moltissime zolle è alquanto dubbio se sia possibile asserire o negare tale appartenenza. L'esistenza di tali zolle "borderline" mostra indiscutibilmente che quando si parla del Cervino compare in qualche modo il fenomeno della vaghezza. Tuttavia la posizione di Varzi è radicalmente contro l'idea che il termine "Cervino" denoti un oggetto vago e, più in generale, che possano esistere oggetti vaghi. Egli infatti afferma che:

"L'unica nozione di vaghezza che mi è intelligibile è di natura linguistica, o cognitiva in senso lato, e questa corrisponde alla lettura de dicto: la vaghezza risiede cioè nel sistema di rappresentazione, non nel mondo che viene rappresentato. Dire che il referente di un termine non è un oggetto ben determinato — che non ha confini ben definiti — significa dire che il termine designa in modo vago, non che esso designa un oggetto vago. E quale sarebbe quest'oggetto vago? Come avremmo fatto a dargli un nome, se è così evanescente?"

Non sarebbe quindi il Cervino ad essere una entità vaga ma è vago il rapporto tra il termine designante "Cervino" e la cosa o, per meglio dire le cose designate da tale termine. Poiché tali cose sono viste come possibili determinazioni del Cervino, le idee di Varzi si collegano naturalmente alla teoria delle supervalutazioni.

Tale problematica come tutta la problematica legata alla vaghezza è alquanto complessa ed in proposito si rimanda alla letteratura sull'argomento. Qui mi limito ad avanzare alcune considerazioni.

Nell'esempio dl Cervino la vaghezza in effetti appare a due differenti livelli. In uno, in cui il Cervino viene visto come il composto delle proprie zolle, la vaghezza risiede nella nozione "essere una zolla del Cervino" e si manifesta nel fatto che esistono zolle borderline (sul "confine" del Cervino). In tale senso non si può affermare che il Cervino sia vago, ma che è vago il rapporto tra il Cervino e le sue zolle. Equivalentemente, si potrebbe dire che le zolle del Cervino costituiscono un insieme vago e che è in questo che risiede la vaghezza di cui parla Varzi. Tuttavia non mi sembra che possa essere questo a rendere vago il Cervino. Ciò perché il Cervino è qualcosa di diverso dall'insieme

(fuzzy) delle sue zolle e perché non può essere il rapportarsi in modo vago ad altri oggetti (le zolle) che può rendere un oggetto (il Cervino) vago.

In un secondo livello, che mi sembra di maggiore interesse, la vaghezza risiede nel rapporto tra il termine denotante Cervino e le varie montagne che si otterrebbero togliendo o aggiungendo zolle alla montagna che, in un certo istante, ho chiamato Cervino. Differentemente dal primo livello ciò non ha molto a che fare con il fatto che esistono zolle bordeline. Questo tipo di vaghezza può essere riscontrata anche negli oggetti dai confini netti come può essere una montagna rocciosa dalle pareti a piombo.

Ora un modo per inquadrare la questione è tenere conto che nella individuazione di un oggetto è presente sempre anche una “ospite occulta” che non viene mai esplicitata: questa ospite è una qualche nozione di uguaglianza.¹² Se guardiamo la sedia S che ci è davanti individuiamo un oggetto ma se qualcuno sposta la sedia di un metro, anche dopo lo spostamento non esitiamo a ritenere che essa rimanga esattamente “la stessa” sedia e che il termine S possa continuare ad essere utilizzato per denotarla. Nel caso che, come matematici, siamo affezionati al punto di vista estensionale, possiamo dire che il termine S denota la classe completa di equivalenza

$$[s] = \{s : s \equiv s\}$$

dove s è la sedia concreta particolare che in un certo istante ho percepito e $\equiv$ è la relazione di equivalenza indotta dal movimento.¹³ Equivalentemente possiamo dire che S denota la coppia costituita da s più l’equivalenza $\equiv$. In accordo con tale esempio possiamo azzardare la seguente definizione.

*Un oggetto (crisp) è un ccdi (cosa concreta in un dato istante) più un criterio di equivalenza.*¹⁴

¹² La questione di che cosa sia un oggetto fisico non è banale ed esiste una vasta letteratura in proposito. Citiamo ad esempio [Noll 1993], [Borgo, Guarino, Masolo 1996]b, [Borgo, Guarino, Masolo 1996]c, [Bennett 2002]), [Bottani 2001].

¹³ In questo modo il termine S ha una sola interpretazione (la classe $[s]$) e non, contrariamente al punto di vista dei supervalutazionisti, infinite determinazioni (gli elementi della classe $[s]$).

¹⁴ Poiché non si specifica il tipo di equivalenza, tale definizione è abbastanza elastica da essere compatibile con vari punti di vista sulla

Supponiamo ora che, avendo osservato che la sedia che abbiamo davanti non è ben bilanciata, decidiamo di accorciare un pò una sua gamba con una raspa. Otteniamo in questo modo un nuovo *ccdi* che si può ritenere ancora uguale al *ccdi* precedente. Il grado di questa uguaglianza, tuttavia, pur essendo prossimo ad 1 è sicuramente diverso da 1. In caso contrario non sarebbe difficile produrre il paradosso per cui lavorando man mano di raspa si ottiene una sedia-fantasma che, pur non esistendo più, rimane uguale alla sedia originaria. A questo punto o si ritiene di essere davanti ad una nuova sedia, magari cambiandone il nome da *S* a *T*, oppure si ritiene che sia possibile convivere razionalmente con la vaghezza ed accettare un indebolimento della nozione di eguaglianza. Un fenomeno non diverso si presenta per l'oggetto-Cervino. Supponiamo che in una passeggiata su tale montagna (*ccdi* originario) io decida di raccogliere una zolla di terra per poter riempire un vaso. Allora è indiscutibile che quello che rimane è un *ccdi* diverso da quello originario (una diversa montagna).¹⁵ Tuttavia continuo a chiamare Cervino tale nuovo *ccdi* che quindi considero "uguale" alla montagna che avevo prima. Anche in questo caso la parola "uguale" non rappresenta l'identità e neanche una relazione di equivalenza ma quella che abbiamo chiamato una relazione fuzzy di equivalenza o similarità.

In definitiva sembra ragionevole la seguente definizione.

Definizione 16.1. Un *oggetto* è un *ccdi* più un criterio di equivalenza fuzzy o similarità. L'oggetto viene detto *crisp* se tale similarità è *crisp*, *vago* altrimenti.

Naturalmente uno stesso *ccdi* può dare luogo a molti oggetti sia *crisp* che *vaghi*: dipende tutto dal tipo di similarità che si intende scegliere.

natura degli oggetti materiali (si veda il tridimensionalismo, il quadridimensionalismo, il sequenzialismo come esposto ad esempio in [Varzi 2002]).

¹⁵ In caso contrario, molto meglio di quanto siano capaci di fare i fisici, avrei creato materia dal nulla con finanziamenti bassissimi ed utilizzando solo un po' di energia, una zappa ed un po' di filosofia!

Si osservi che una tale definizione di oggetto trova un forte riscontro nell'approccio categoriale alla teoria degli insiemi fuzzy. Ad esempio in [Novák et al. 1999] viene definita come rappresentativa della teoria degli insiemi fuzzy la categoria degli Ω -insiemi. Gli oggetti di tale categoria sono le coppie (X, e) dove e è una relazione fuzzy simmetrica e transitiva (non necessariamente riflessiva). La scelta di tale categoria è fortemente motivata dal fatto che si vuole una categoria che sia un *topos* cioè che abbia proprietà ragionevoli che permettano di riprodurre tutto quanto può essere fatto dalla categoria degli insiemi.

17. Oggetti vaghi, proprietà, gruppi

Nei paragrafi precedenti abbiamo visto che esistono vari modi per ottenere una similarità e questo comporta che la definizione 16.1 si possa particularizzare in modi diversi di definire un oggetto. Ad esempio potremmo riferirci alla similarità intesa come invarianza rispetto ad un dato gruppo fuzzy di trasformazioni e quindi considerare la seguente definizione.

Definizione 17.1. Un *oggetto* è un *ccdi* più un gruppo fuzzy di trasformazioni ammissibili per *ccdi*. Tale oggetto viene detto *crisp* o *vago* a seconda se tale sottogruppo sia *crip* o *vago*.

Nel caso della sedia *crisp* si deve coinvolgere il gruppo dei movimenti rigidi. Nel caso della sedia *vaga* è necessario considerare il gruppo fuzzy delle trasformazioni ottenibili con colpi di raspa (gruppo che non sarebbe difficile definire). La stessa cosa può essere detta per il Cervino *vago*.

Una possibilità più interessante è considerare la similarità definita da un dato insieme di proprietà considerate rilevanti.

Definizione 17.2. Un *oggetto* è un *ccdi* più un insieme di proprietà che si considerano rilevanti. Se le proprietà sono vaghe l'oggetto si dice *vago*, altrimenti si dice *crisp*.

In questo caso un oggetto *vago* rimane se stesso durante una eventuale trasformazione in relazione al conservarsi o meno delle proprietà rilevanti. Nel caso del Cervino potrei considerare rilevanti le proprietà “essere pulito” o “essere silenzioso” o altre

ancora. Queste proprietà sono scalpite solo per un milionesimo dal fatto che manchi a meno qualche zolla. Se invece confronto un Cervino immacolato che ho conosciuto in passato con un Cervino invaso da una busta di plastica che mi è davanti, sono autorizzato ad esclamare “Eh !: purtroppo non esiste più il Cervino di una volta !” o, in maniera più sofisticata “il Cervino di una volta ha un grado di similarità nullo con il Cervino che mi sta davanti”.

Tale modo di definire un oggetto vago permette anche di ridefinire in termini di similarità le idee di Varzi della vaghezza come proprietà tra denotante e denotato. L’idea è che si possano definire “simili” due elementi di D che hanno gli stessi nomi nel linguaggio in considerazione. Per precisare tale idea, dobbiamo abbandonare la scelta fatta dalla logica a più valori in cui una costante è interpretata con un elemento del dominio ed accettare che anche l’interpretazione delle costanti possa essere vaga. Ciò significa che se C denota l’insieme dei termini del linguaggio, una interpretazione dei termini è rappresentata da una relazione vaga $I : C \times D \rightarrow U$ e che il valore $I(c, d)$ è il grado con cui è lecito considerare d come interpretazione di c . Se consideriamo l’insieme fuzzy i_c definito dall’equazione $i_c(x) = I(c, x)$, allora i_c è l’insieme fuzzy degli elementi che sono interpretazione di c o, se si vuole, l’estensione della proprietà, che indichiamo con $A_c(x)$, “ x è una interpretazione di c ”. Ciò suggerisce la seguente definizione.

Definizione 17.3. Sia $C = \{c_1, \dots, c_n\}$ l’insieme dei termini di un linguaggio, D un dominio e $I : C \times D \rightarrow U$ una interpretazione vaga, allora chiamiamo *omonimia* la similarità associata all’insieme $(A_c)_{c \in C}$ di predicati, cioè la fuzzy relazione e ottenuta ponendo

$$e(x, x') = \bigotimes_{c \in C} (i_c(x) \leftrightarrow i_c(x')).$$

Tale equazione definisce appunto $e(x, x')$ come il grado di verità dell’affermazione

“*comunque si scelga un termine c , c è un nome di x se e solo se c è un nome di x'* ”.

Da notare che se chiamiamo “*senza nome*” un elemento $x \in D$ tale che $i(c, x) = 0$ per ogni $c \in C$, con tale definizione due oggetti che non hanno nomi vengono considerati equivalenti.

Definizione 17.4. Sia $C = \{c_1, \dots, c_n\}$ l'insieme dei termini di un linguaggio, D un dominio ed $I : C \times D \rightarrow U$ una interpretazione vaga, allora un *oggetto* è un ccdi più la relazione fuzzy di omonimia.

Se se ne accetta l'adeguatezza, tale definizione mostra che l'idea di Varzi per cui la vaghezza risiede nel rapporto tra denotante e denotato può essere riformulata nei termini proposti in questa nota. Nel fare questo evidenzia che quella di Varzi potrebbe essere solo una delle possibili proposte, quella in cui il linguaggio gioca un ruolo esclusivo.¹⁶ Il riferimento invece ad altre possibili similarità ha invece il pregio di evidenziare che l'idea di oggetto è legata ad una più generale capacità di cogliere la similarità o, se si vuole, l'invarianza, senza che questa capacità sia legata necessariamente al linguaggio. D'altra parte Varzi parla di vaghezza di natura linguistica "*o cognitiva in senso lato*". Si potrebbe infatti argomentare che tale capacità sia una componente essenziale dell'intelligenza di ogni essere vivente che esso abbia o meno capacità linguistiche. Anche un cagnolino si rapporta agli oggetti ogni istante: siano tali oggetti crisp come la ciotola che contiene il proprio cibo, siano essi vaghi come il cibo della ciotola che man mano viene mangiato. In tale senso il riferirsi solo al "de dicto" sembrerebbe troppo limitativo.

18. Oggetti in matematica e supervalutazionismo

E' interessante osservare che la definizione che abbiamo proposto di oggetto è abbastanza usuale perfino in matematica.

Consideriamo un semplice esempio di oggetto matematico: il numero razionale $4/6$. La prima cosa che impariamo a scuola è che tale numero è identico a $2/3$ oppure a $600/900$ oppure ad infinite altre frazioni. Questo fatto è alla base della definizione del campo dei razionali che viene riportata nei libri universitari di matematica, definizione che ricordiamo brevemente¹⁷.

¹⁶ D'altra parte questo ruolo del linguaggio è conseguenza inevitabile di una interpretazione "de dicto" della vaghezza.

¹⁷ Ne approfitto per fare pubblicità al mio bellissimo libro "*Tentativi di Fondare la Matematica*" dove si possono trovare le definizioni degli insiemi numerici. Il libro si può acquistare per modico prezzo sul sito *Ilmiolibro* ed è particolarmente consigliabile ad insegnanti di matematica o di filosofia.

1. Si parte dall'insieme $Z \times (Z - \{0\})$ di coppie (n, m) di numeri interi relativi con $m \neq 0$,
2. si definiscono in $Z \times (Z - \{0\})$ due operazioni $+$ e $\cdot$ ponendo

$$(n, m) + (p, q) = (n \cdot q + m \cdot p, mq) ; (n, m) \cdot (p, q) = (n \cdot p, m \cdot q)$$
3. si definisce nell'insieme delle frazioni $Z \times (Z - \{0\})$ una relazione di equivalenza $\equiv$ ponendo $(n, m) \equiv (n', m')$ se e solo se $n \cdot m' = m \cdot n'$
4. si mostra che tale relazione è una *congruenza*¹⁸ della struttura algebrica $(Z \times (Z - \{0\}), +, \cdot)$, cioè che

$$x \equiv x' \text{ e } y \equiv y' \Rightarrow x + y \equiv x' + y', x \cdot y \equiv x' \cdot y'$$
5. si considera il *quoziente* di $(Z \times (Z - \{0\}), +, \cdot)$ modulo $\equiv$, cioè la struttura algebrica
 - il cui dominio è $Q = \{[(n, m)] : m \neq 0\}$ dove $[(n, m)]$ denota la classe di equivalenza $\{(n', m') : (n', m') \equiv (n, m)\}$
 - in cui le due operazioni $+$ e $\cdot$ sono definite ponendo

In accordo con l'abitudine, possiamo anche indicare con n/m la coppia (n, m) . Tuttavia un modo leggermente diverso di riformulare tale costruzione è di non arrivare al quoziente di cui si dice al punto 5 ma di considerare la struttura $(Z \times (Z - \{0\}), +, \cdot, \equiv)$ dove l'uguaglianza è interpretata questa volta dalla congruenza $\equiv$. In altre parole possiamo definire l'oggetto "numero razionale" n/m come la coppia di interi n/m (il *ccdi*) più la relazione di equivalenza $\equiv$. Questa scelta rappresenta molto meglio il modo di procedere di chi effettua calcoli con le frazioni che non pensa affatto di trasportare, nei suoi calcoli, il "peso" di intere classi di equivalenza. E' anche la scelta fatta da un linguaggio potentissimo come quello di *Mathematica* che tratta i razionali come coppie e traduce l'equivalenza in un sistema di riscrittura e di riduzione a forma normale.

A ben riflettere tuttavia la coppia n/m potrebbe definire anche altri oggetti matematici poiché è possibile scegliere in vari modi l'equivalenza da accostare a tale coppia. Ad esempio, se teniamo conto dei tempi di calcolo non sarebbe assurdo assumere come equivalenza l'identità e quindi distinguere tutte le coppie tra loro.

¹⁸ Tale richiesta è essenziale poiché assicura che le operazioni tra classi non dipendono dagli elementi rappresentativi scelti in ciascuna classe.

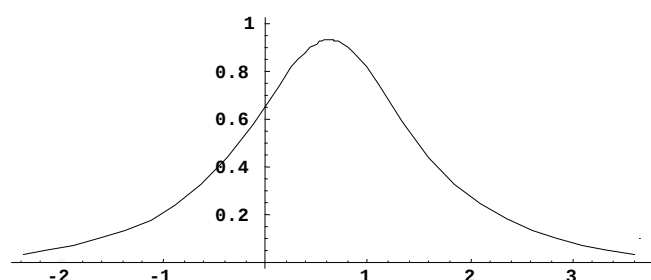
Si tratta di distinguere, il “dividere per 6 e moltiplicare per 4” dal “dividere per 900 e moltiplicare per 600”. In questi casi di solito si parla di “frazione” e non di numero razionale.¹⁹

Ancora, se abbiamo comprato una calcolatrice che non tratta i numeri oltre al quarta cifra decimale, possiamo essere indotti a considerare equivalenti due numeri che coincidono fino alla 4 cifra. Allora con $4/6$ dovremmo denotare una classe che contiene sia la coppia $40001/60000 = 0.666683...$ che la coppia $4666/1000 = 0.4666000.....$ ²⁰

Ma possiamo anche definire un razionale vago se a $4/6$ si affianca una equivalenza vaga. Ad esempio possiamo introdurre nell’insieme delle coppie n/m una eguaglianza “più rilassata” che tenga conto della *vicinanza* tra numeri razionali. Il seguente è un esempio

$$e(n/m, p/q) = 1/((n \cdot q - m \cdot p)/m \cdot q)^2 + 1).$$

In questo caso una coppia come $4/6$ denota un oggetto vago, che



Questo non è il Cervino ma il numero “circa $4/6$ ”

potremmo chiamare “circa $4/6$ ”²¹. Questo oggetto è costituito dall’insieme fuzzy di frazioni il cui grafico, guarda caso, ricorda molto una montagna.

¹⁹ Solo un matematico può pensare la cosa assurda che tagliare una torta in 3 parti è pigliarne 2 sia esattamente la stessa cosa di tagliarla in 900 parti e pigliarne 600 (se non altro per il tempo che ci vuole e per il fenomeno delle briciole entrambi aspetti questi di cui il matematico si disinteressa completamente).

²⁰ La teoria degli *insiemi rozzi* (rough set theory) tratta gli oggetti matematici in questo modo.

²¹ A proposito dei numeri del tipo “circa n ” è interessante il seguente colloquio tratto dal libro di Lewis Carroll, “*Sylvie and Bruno concluded*” e che ho scoperto leggendo un bel libro di Carlo Toffalori [Toffalori 2011]. Bruno sta alla finestra di una casa di campagna,

Numeri di tale tipo sono ampiamente utilizzati dai fuzzisti sotto il nome di “numeri fuzzy” e sono ad esempio alla base della teoria del controllo fuzzy.

Forse l’idea che in un oggetto matematico è implicita una nozione di equivalenza permette di fare entrare in gioco la nozione di verità proposta dalla teoria delle supervalutazioni anche nel preciso mondo della matematica classica. Infatti quando affermiamo, ad esempio, che è vera l’affermazione “il numero $4/6$ è minore del numero $3/4$ ” possiamo intendere che, presa una qualunque determinazione n/m del primo numero ed una qualunque determinazione p/q del secondo numero risulta che è verificata la proprietà $n \cdot q \leq m \cdot p$. In altre parole, tale asserzione è vera se è vera per tutte le possibili determinazioni di $4/6$ e di $3/4$. Supponiamo invece di inventarci una diversa proprietà, ad esempio

“dicesi *strabico* un razionale n/m tale che $n-1 = m+1$ ”.

allora saremmo tentati di affermare che $5/3$ è strabico. Tuttavia un supervalutazionista osserverebbe che una differente determinazione di tale razionale, ad esempio $15/9$ non è più strabica e quindi la proprietà di essere strabico è indeterminata per $5/3$. Naturalmente mentre una affermazione del tipo “ $15/9$ o è strabico oppure non lo è” per un supervalutazionista è vera, per un matematico sarebbe invece non è ben definita in quanto, non essendo ben definita la nozione di “strabico”, è un errore matematico considerare qualunque affermazione che la coinvolga.²²

Concludo questo paragrafo osservando che ovviamente esiste una differenza notevole tra un oggetto vago come il Cervino, che

occupato a stimare il numero dei maiali di un grosso branco che vaga nel prato di fronte. A Sylvie che gli chiede quanti essi siano risponde:

- Circa mille e quattro.
- Intendi dire circa mille – lo corregge Sylvie –. Non è giusto dire “e quattro”, non puoi essere sicuro dei quattro!
- Non potevi fare errore più grosso! – esclamò Bruno trionfante – E’ proprio dei quattro che sono certo; sono qui sotto la finestra! Semmai è dei mille che non sono sicuro!”.

²² Insomma, a dispetto del vecchio metodo delle torte, per insegnare bene le frazioni ai bambini dovrebbe essere fatto loro prima un corso sulla teoria delle congruenze e dei quozienti per strutture algebriche e poi un corso sulla teoria delle supervalutazioni!

abbiamo definito tramite una misteriosa nozione di *ccdi* ed una altrettanto misteriosa e mal definita nozione di similarità, ed un numero vago come “circa 4/6” che invece parte da una precisa definizione di coppia e di similarità. D'altra parte è evidente che esiste sempre una differenza notevole tra qualunque pezzo della realtà ed qualunque modello matematico che pretenda di rappresentarlo.

19. Il paradosso di Evans e quello della nave di Teseo.

L'argomentazione di Evans. Il definire gli oggetti vaghi tramite un insieme di proprietà che si considerano rilevanti può dare forse una risposta all'argomentazione di G. Evans circa l'impossibilità che possa esistere una uguaglianza di tipo vago tra due oggetti [Evans 1978]. Ovviamente se tale argomentazione fosse valida invaliderebbe la definizione di oggetto vago che abbiamo proposto. Il discorso di Evans è il seguente. Indichiamo con $Vago(a = b)$ l'asserzione per cui l'uguaglianza tra a e b è vaga (noi diremmo che vale con valore diverso da 0 ed 1) ed accettiamo la verità di tale asserzione, allora la proprietà $B(x) \equiv Vago(x=b)$ si applica ad a . D'altra parte è evidente che $B(x)$ non è applicabile a b . Pertanto a e b verificano proprietà diverse. Per il principio degli indiscernibili di Leibnitz, ciò comporta che a e b sono diversi tra loro. Ma tale conclusione comporta che $Vago(a=b)$ è falsa. Tale argomentazione costituisce una dimostrazione per assurdo di $\neg Vago(a=b)$ e quindi che non può esistere una uguaglianza vaga. Ora, supponiamo di riferirci ad una “uguaglianza approssimata” e definita dalle proprietà $A_1, \dots, A_n$ e supponiamo che $e(a,b) = 0.5$, allora se denotiamo con $B(x)$ la proprietà $Vago(E(x,b))$ la formula $B(x)$ risulta è totalmente falsa per b mentre è totalmente vera per a . Pertanto se tra le proprietà A_i ci fosse $B(x)$, ciò comporterebbe che $e(a,b) = 0$ in contrasto con le ipotesi. Tutto ciò non è affatto assurdo ma proverebbe solo che $B(x)$ non è tra le proprietà con cui si è definita e , cosa questa abbastanza prevedibile. Esiste una ampia ed interessante letteratura sull'argomentazione di Evans ma non rientra nelle ambizioni di questo libro affrontare in modo dettagliato tale questione (si veda comunque [Pelletier 1989]).

La nave di Teseo. La definizione proposta di oggetto vago forse si presta anche ad affrontare il famoso “paradosso della nave di

Teseo”, esposto da Plutarco nelle *Vite parallele*. Plutarco narra che la nave di legno sulla quale viaggiò Teseo fino a Creta per uccidere il Minotauro fu conservata intatta nel corso dei secoli dagli Ateniesi, i quali sostituivano man mano le parti che si venivano a deteriorare. Dopo un po’ di tempo risultò che, benché la nave conservasse esattamente la sua forma iniziale, tutte le parti usate in origine per costruire la nave erano ormai rimpiazzate.

A questo punto la questione che si pone è se dopo la sostituzione di tutte le parti la nave sia ancora la nave di Teseo da cui si era partiti.²³

Possiamo tentare di fornire una risposta a tale paradosso affermando che il termine “La nave di Teseo” denota l’oggetto vago individuato da un *ccdi* (nave di partenza) più una relazione di similarità tra artefatti di legno. Naturalmente la scelta del tipo di similarità è alquanto arbitraria e dipende dagli scopi che uno si pone. Un curatore di museo farebbe decrescere la similarità abbastanza rapidamente con l’aumentare dei pezzi sostituiti, fino ad assegnare similarità 0 ad una nave che risulti essere solo una riproduzione della nave originaria. Un capitano che si dovesse preoccupare solo dell’efficienza della nave, assegnerebbe grado di similarità 1 a tale riproduzione. D’altra parte in un negozio di navi di Teseo costruite in serie una nave a cui è stato sostituito un pezzo rimane uguale a se stessa con grado 1.

20. Qualche considerazione finale sugli oggetti vaghi

La definizione proposta di oggetto vago non costituisce ovviamente una risposta alla domanda di Varzi circa l’esistenza di oggetti vaghi. Comunque evidenzia che tale questione non appare molto diversa da quella dell’esistenza degli oggetti *tout court* ed entrambe forse non sono troppo diverse dalla questione dell’esistenza degli oggetti matematici. Tuttavia contro l’ipotesi di una esistenza oggettiva gioca proprio il fatto che la scelta del tipo di uguaglianza appare come un intervento fortemente soggettivo dell’uomo. Possiamo decidere che *S* denoti una sedia *crisp* (se si assume l’uguaglianza determinata dai movimenti

²³ Possiamo riscrivere tale paradosso in modo più inquietante riferendoci alla nostra identità personale, dato che ciascuno di noi, sebbene sembri rimanere la stessa persona, cambia continuamente.

rigidi) oppure che denoti una sedia vaga (se si assume la similarità legata all'azione di "lavorare di raspa"). Due sedie a teatro possono essere anche perfettamente identiche ma saranno considerate diverse se collocate diversamente rispetto al palcoscenico. Ancora, possiamo decidere che $4/6$ rappresenti un numero razionale, come fanno i matematici, oppure una frazione, oppure un numero fuzzy oppure un numero approssimato.

Per quanto riguarda gli oggetti matematici la parte intuizionista che è in me mi suggerirebbe che essi siano una invenzione-costruzione della mente umana e quindi che essi non esistano più di quanto possa esistere Topolino. Tuttavia, come tutti i matematici, sono di fatto un platonista il quale agisce come se tali oggetti avessero una esistenza propria. Come scusa si può dire che si tratta di un platonismo logico o psicologico e non necessariamente ontologico. Non mi sembra che per gli oggetti non matematici, siano essi vaghi oppure no, le cose vadano diversamente. In questo senso si può forse affermare che esistono nel mondo sia gli oggetti crisp che quelli vaghi o perlomeno che la mente umana ha bisogno di fare finta che esistano per poter sopravvivere.²⁴

²⁴ Da napoletano non posso fare a meno di ricordare la celebre teoria del *verum = factum* del mio compaesano Gianbattista Vico. Secondo tale teoria si possono conoscere veramente solo le cose che si fanno. Quindi non possiamo conoscere la natura, che ha fatto il buon Dio, ma solo le cose che l'uomo produce, siano esse le poesie, i personaggi dei fumetti, la storia, i numeri razionali, il modello matematico della relatività generale, gli oggetti crisp e, ... gli oggetti vaghi.

CAPITOLO 3

RAGIONAMENTI APPROSSIMATI

1. La logica come strumento per trattare informazione

Nel capitolo precedente ci siamo mantenuti ad un livello semantico esaminando che cosa si debba intendere per interpretazione di un linguaggio quando in tale linguaggio siano ammessi anche predicati vaghi. Naturalmente la logica fuzzy è qualche cosa di più perché in ogni logica è centrale la nozione di dimostrazione. Ora i matematici, da Euclide in poi, vedono le dimostrazioni come strumenti per ricavare teoremi da un dato sistema di assiomi, sistema al quale si assegna, per così dire, un valore assoluto (almeno all'interno di una data teoria). Invece i ragionamenti che si fanno quotidianamente hanno un significato ed un ruolo diverso. Essi costituiscono un modo per estrarre (per meglio dire esplicitare) dall'informazione disponibile nuova informazione. Se si accetta questo punto di vista, un passo fondamentale per definire una logica fuzzy è quello di decidere che cosa si debba intendere per "informazione" su una struttura fuzzy. Ora "pezzi di informazione" potrebbero essere relative ai valori di verità assunti dalle singole formule, ad esempio

"il valore di verità di α è maggiore o uguale di 0.7",

"il valore di verità di α è maggiore di 0.7",

"il valore di verità di α oscilla tra 0.4 e 0.6",

"il valore di verità di α è minore di 0.2",

"il valore di verità di α è esattamente 0.1".

Chiamiamo *constraint sul valore di α* informazioni di tale tipo.⁴⁸ Ogni constraint in questo elenco si riferisce ad un sottoinsieme, per la precisione ad un intervallo in U . Ad esempio la prima delle espressioni indicate, coinvolge l'intervallo $[0.7, 1]$, la seconda l'intervallo $(0.7, 1]$ e così via. In logica fuzzy usualmente ci si

⁴⁸ Non sono queste le sole possibili forme di informazione. Potrebbero ad esempio essere messi in relazione tra loro in qualche modo i valori di verità di due o più asserzioni. Ad esempio se dico "Maria è di gran lunga più fortunata di quanto meriti" metto in relazione i gradi di verità λ e μ di due asserzioni vaghe: "Maria è fortunata" e "Maria merita". Precisamente affermo che vale la relazione fuzzy " λ è di gran lunga maggiore di μ ". Non esploreremo tali possibilità che d'altra parte non mi sembrano mai essere state esaminate in letteratura.

riferisce solo ad intervalli chiusi. Probabilmente il motivo risiede nel fatto che gli intervalli chiusi costituiscono un sistema di chiusura (si veda il paragrafo 14) e questo permette di fondere ogni insieme di constraint sulla stessa formula in un unico constraint. Per la precisione usualmente ci si riduce ad intervalli chiusi del tipo $[\lambda, 1]$ ed il motivo è dato dalla seguente proposizione.

Proposizione 1.1. Indichiamo con $\langle \alpha, \lambda \rangle$ l'informazione "il valore di verità di α è maggiore o uguale di λ ". Allora se la negazione è interpretata con una buona negazione, ogni insieme T di constraint definiti da intervalli chiusi può essere rappresentato da un insieme s_T di tali coppie in modo che ogni formula α compare una ed una sola volta in s_T . In altri termini può essere rappresentato da un insieme fuzzy $s_T : F \rightarrow U$ di formule.

Dim. Basta osservare che se $\sim$ è l'interpretazione della negazione allora T può essere modificato senza alterarne il contenuto informativo applicando prima la regola:

- ogni espressione del tipo "il valore di verità di α è compreso tra λ e μ " può essere sostituita dalle due coppie $\langle \neg \alpha, \sim \mu \rangle$ e $\langle \alpha, \lambda \rangle$ ⁴⁹; e poi la regola:

- fissata $\alpha \in F$, l'insieme $\{\langle \alpha, \lambda \rangle : \langle \alpha, \lambda \rangle \in T\}$ può essere sostituito da $\langle \alpha, \mu \rangle$ dove $\mu = \text{Sup}\{\lambda : \langle \alpha, \lambda \rangle \in T\}$.⁵⁰ $\square$

In definitiva si perviene alla seguente definizione.

Definizione 1.2. Chiamiamo *insieme fuzzy di ipotesi* (o *insieme fuzzy di assiomi* o *valutazione*) qualunque insieme fuzzy $v : F \rightarrow U$ di formule.

E' importante porre in rilievo che se v viene visto come insieme fuzzy di ipotesi, allora, data una formula α , il numero $v(\alpha)$ non è da considerare come valore di verità di α ma come una

⁴⁹ In particolare, possiamo tradurre una informazione del tipo "il valore di verità di α è esattamente λ " con $\langle \neg \alpha, 1-\lambda \rangle$ e $\langle \alpha, \lambda \rangle$.

⁵⁰ In U l'estremo superiore dell'insieme vuoto è uguale a 0. Pertanto nel caso in cui per α non esiste nessuna coppia $\langle \alpha, \lambda \rangle \in T$, si deve semplicemente aggiungere la coppia $\langle \alpha, 0 \rangle$.

informazione sul possibile valore di verità di α . Come abbiamo già detto, rappresenta l'informazione “il valore di verità di α è maggiore o uguale a $v(\alpha)$ ”.

2. L'operatore di conseguenza logica

La semantica proposta nel capitolo 2 si riferisce alla logica del primo ordine ed è vero-funzionale poiché il grado di verità di una proposizione composta è funzione del grado di verità delle proposizioni che la compongono. Per potere procedere in modo più generale, in modo da coinvolgere logiche di ordine diverso da uno e da potere trattare anche semantiche che non sono vero-funzionali⁵¹, è preferibile un modo più astratto di definire la semantica. Osserviamo in proposito che ad ogni interpretazione fuzzy (D, I) è associato un insieme fuzzy di formule e precisamente l'insieme fuzzy $\underline{I} : F \rightarrow U$ che associa ad ogni formula il relativo grado di verità. Possiamo vedere $\underline{I}$ come l'insieme fuzzy delle formule vere in (D, I) . Quindi la semantica vero-funzionale che abbiamo considerato può essere identificata con la classe $M \subseteq U^F$ di tali valutazioni. Discorso analogo potrebbe essere fatto per le logiche del secondo ordine. D'altra parte, anche nel calcolo proposizionale della logica classica o di una logica a più valori la semantica è definita dall'insieme M delle valutazioni vero-funzionali del calcolo proposizionale. Ciò suggerisce la seguente definizione astratta.

Definizione 2.1. Dato un insieme F i cui elementi vengono chiamati *asserzioni* o *formule*, chiamiamo *semantica fuzzy astratta* una classe M di sottoinsiemi fuzzy di F . Gli elementi di M vengono chiamati *modelli*.

Coerentemente con l'interpretazione che abbiamo dato di insieme fuzzy di ipotesi, possiamo estendere la nozione di modello di un insieme fuzzy di ipotesi data nel capitolo 2 alla semantica astratta.

⁵¹ Come vedremo nel seguito, ne esistono di interessanti in cui la vaghezza è a livello metalinguistico.

Definizione 2.2. Data una semantica fuzzy M ed un insieme fuzzy di ipotesi v , diciamo che $m \in M$ è un *modello* di v e scriviamo $m \models v$ se risulta che $m(\alpha) \geq v(\alpha)$ per ogni formula α .

In logica classica la nozione di *conseguenza logica* $\models$ si definisce ponendo, dato un sistema di assiomi T ed $\alpha \in F$,

$$T \models \alpha \Leftrightarrow \text{ogni modello di } T \text{ è anche un modello di } \alpha.$$

Per capire il significato di tale relazione supponiamo che T sia l'insieme dei fatti che si sono venuti a sapere di un "mondo" sconosciuto. Allora, data una asserzione α , se risulta $T \models \alpha$ si è sicuri che in questo mondo sconosciuto α risulta valida e questo anche se le informazioni disponibili non permettono di individuare tale mondo con precisione. Se poniamo

$$Lc(T) = \{\alpha : T \models \alpha\}$$

definiamo un importante operatore detto *operatore di conseguenza logica*. $Lc(T)$ rappresenta l'informazione che è comune a tutti i modelli di T e quindi tutta l'informazione sul mondo sconosciuto che possiamo sperare di ottenere per via logica (indipendentemente dalla nostra effettiva capacità di ottenerla).

Se in logica fuzzy vogliamo definire un operatore analogo, data una valutazione v dobbiamo riferirci all'informazione che è comune a tutti i modelli di v . Ora, data una formula α , l'unica cosa che sappiamo è che il valore di verità di α appartiene all'insieme $\{m(\alpha) : m \models v\}$ e che quindi è maggiore o uguale a $\lambda = \inf\{m(\alpha) : m \in M, m \models v\}$. Tenendo conto che abbiamo deciso di rappresentare tale tipo di informazione con un insieme fuzzy, si perviene alla seguente definizione.

Definizione 2.3. Sia M una semantica fuzzy e sia $v : F \rightarrow L$ una valutazione, allora l'insieme fuzzy $Lc(v)$ di *conseguenze logiche* di v è definito ponendo:

$$Lc(v)(\alpha) = \inf\{m(\alpha) : m \in M, m \models v\}. \quad (2.1)$$

L'operatore $Lc : U^F \rightarrow U^F$ definito in tale modo è chiamato *operatore di conseguenza logica fuzzy*.

Per capire la natura diversa della logica fuzzy rispetto alla logica classica, è interessante definire la nozione di insieme fuzzy delle tautologie.

Definizione 2.4. Data una semantica M , insieme fuzzy delle tautologie è l'insieme fuzzy di formule $Tau : F \rightarrow U$ definito ponendo $Tau = Lc(\emptyset)$ e quindi ponendo, per ogni $\alpha \in F$,

$$Tau(\alpha) = \text{Inf}\{m(\alpha) : m \in M\}. \quad (2.2)$$

Per ogni formula α , il numero $Tau(\alpha)$ rappresenta l'informazione che esiste a priori su α , informazione che dipende solo dalla struttura logica di α . Naturalmente mentre nella logica classica una formula o è una tautologia oppure non lo è, nella logica fuzzy ogni formula è una tautologia con un certo grado.

Esempio 1. Consideriamo la logica fuzzy in cui la disgiunzione è interpretata con la co-norma di Zadeh (cioè con il massimo) e la negazione con la funzione $1-x$. Allora, ad esempio, la tautologia classica $\beta \vee \neg \beta$ assumerà come valore di verità $\lambda \vee (1-\lambda)$ se λ è il valore di verità di β . E' facile allora vedere che il minimo dei valori che può assumere tale formula è 0.5 e che quindi $Tau(\beta \vee \neg \beta) = 0.5$. Ciò significa che anche se non disponiamo di nessuna informazione sullo stato delle cose, in tale logica sappiamo comunque che la formula $\beta \vee \neg \beta$ assume almeno il valore 0.5. Ovviamente la formula $\beta \wedge \neg \beta$ è invece una tautologia con grado zero.

Esempio 2. Differente è il caso della logica in cui la disgiunzione è definita come co-norma della norma di Łukasiewicz. Infatti in tale caso abbiamo che il valore di verità di $\beta \vee \neg \beta$ è $\text{Min}\{x+1-x, 1\} = 1$ e quindi che $\beta \vee \neg \beta$ è una tautologia con grado 1.

Esempio 3. Consideriamo infine la logica che ha come norma il prodotto usuale e quindi come co-norma la funzione $x+y-x \cdot y$. In tale caso la formula $\beta \vee \neg \beta$ viene valutata con $x+(1-x)-x \cdot (1-x) = 1-x+x^2$. Il minimo di tale funzione è uguale a 0.75 e pertanto $Tau(\beta \vee \neg \beta) = 0.75$.

3. Caratteristiche di una dimostrazione classica

Nella logica classica la relazione di conseguenza logica è cruciale e pone un problema fondamentale:

come possiamo verificare il fatto che α sia una conseguenza logica di T ?

Se ci riferiamo direttamente alla definizione, dovremmo prendere tutti i possibili modelli di T ed andare a verificare per ogni modello se esso verifica α . Una impresa che appare chiaramente impossibile. Fortunatamente si è provata l'esistenza di apparati inferenziali che permettono di rendere praticabile tale impresa.⁵² Precisamente viene proposta una nozione di dimostrazione sotto ipotesi e viene definita una relazione $\vdash$ il cui significato è che la relazione $T \vdash \alpha$ sussiste se esiste una dimostrazione di α che utilizza come ipotesi formule in T . Una volta data tale definizione, viene poi dimostrato che

$$T \models \alpha \text{ se e solo se } T \vdash \alpha$$

Tale equivalenza, che è nota sotto il nome di teorema di completezza, permette di ridurre il problema apparentemente impossibile di verificare che α è vera in tutti i modelli di T a quello abbordabile di trovare una opportuna dimostrazione di α . In termini di operatori, viene definito l'operatore di deduzione D ponendo

$$D(T) = \{\alpha \in F : T \vdash \alpha\}$$

ed il teorema di completezza si può esprimere dicendo che $D = Lc$.

Se vogliamo fare qualche cosa di analogo per la logica fuzzy, dobbiamo definire una nozione di “ragionamento approssimato” che permetta di definire un operatore di deduzione che coincida con Lc . Nel primo capitolo abbiamo già visto esempi di “ragionamenti approssimati” quando abbiamo esposto la soluzione di alcuni paradossi proposta dalla logica fuzzy. Ora passiamo a dare definizioni più rigorose ed a tale scopo torniamo alla logica classica esaminando che cosa intendiamo usualmente per dimostrazione. Ci accorgiamo che una dimostrazione di una asserzione α è una successione finita di asserzioni $\alpha_1, \dots, \alpha_n$ che termina con α e tale che ciascuna asserzione α_i deve essere in qualche modo giustificata. Le giustificazioni possono essere del tipo:

“nella proposizione XY del paragrafo ... abbiamo già provato che α_i ”,

⁵² E' questo un teorema provato da Gödel la cui importanza non è abbastanza sottolineata. Stranamente in genere si insiste sui teoremi limitativi di Gödel che dicono ciò che non può fare la logica formale e non su tale teorema che mostra una cosa straordinaria che essa può fare.

“ α_i è un assioma”,

“ α_i è banalmente vero”

“ α_i segue immediatamente dalle formule α e β che ora abbiamo dimostrato”.

Ora la prima giustificazione può essere comunque evitata semplicemente ripetendo la dimostrazione della proposizione XY. La giustificazione “è banalmente vero” significa “è una nota tautologia” e quindi significa appartenente ad un certo elenco Al di tautologie. La giustificazione “segue immediatamente” può essere precisata tramite la seguente nozione di regola di inferenza.

Definizione 3.1. Sia F un insieme i cui elementi chiamiamo *formule*, allora dato $n \geq 1$, una *regola di inferenza n-aria* è una operazione parziale n -aria $r : F^n \rightarrow F$ in F .

Ad esempio il Modus Ponens è una regola che si applica solo a coppie di formule del tipo α e $\alpha \rightarrow \beta$ e restituisce la formula β . Pertanto il suo dominio è l'insieme $Dom(r) = \{(\alpha, \alpha \rightarrow \beta) \in F \times F : \alpha \in F \text{ e } \beta \in F\}$ ed inoltre $r(\alpha, \alpha \rightarrow \beta) = \beta$.

Definizione 3.2. Un *sistema inferenziale* è una coppia (Al, RI) tale che Al è un sottoinsieme di F , detto *l'insieme degli assiomi logici*, ed RI è un insieme di regole di inferenza.

La nozione di sistema inferenziale ci permette poi di definire quella di dimostrazione.

Definizione 3.3. Dato un insieme T di formule, una *dimostrazione π di una formula α sotto ipotesi T* è una successione $\alpha_1, \dots, \alpha_m$ di formule tale che $\alpha_m = \alpha$ e, per ogni $i = 1, \dots, m$, α_i sia giustificata da una delle seguenti condizioni:

- $\alpha_i \in T$ (α è un'ipotesi)
- $\alpha_i \in Al$ (α è un assioma logico)
- $\alpha_i = r(\alpha_{s(1)}, \dots, \alpha_{s(n)})$, dove $r \in RI$, $s(1) < i, \dots, s(n) < i$

In altre parole in una dimostrazione si richiede che ogni formula o sia stata assunta in quanto ipotesi, o in quanto assioma logico oppure si sia ottenuta da formule precedentemente dimostrate

$\alpha_{s(1)}, \dots, \alpha_{s(n)}$ tramite una regola di inferenza r . Indichiamo con $l(\pi)$ la lunghezza di π .

Definizione 3.4. Scriveremo $X \vdash \alpha$ per denotare che esiste una dimostrazione di α sotto ipotesi X . L'operatore di deduzione $D : P(F) \rightarrow P(F)$ associato ad (Al, RI) è definito ponendo, per ogni insieme X di formule,

$$D(X) = \{ \alpha \in F : X \vdash \alpha \}. \quad (3.1)$$

Osservazione. Il termine “esiste” nella definizione di D comporta che nell'approccio usuale alla deduzione, una volta che si sia trovata una dimostrazione non ha nessun interesse la ricerca di dimostrazioni “migliori”. Come vedremo nel definire i ragionamenti approssimati questo non è più vero per la logica fuzzy dove dimostrazioni diverse di una formula α forniscono informazioni su α che possono essere diverse.

4. Dimostrazioni che coinvolgono nozioni vaghe

Estendiamo ora le definizioni precedenti al caso fuzzy e, per prima cosa, estendiamo la nozione di regola di inferenza. A tale scopo se si esamina l'estensione del Modus Ponens proposta per la soluzione del paradosso del mucchio di grano, ci accorgiamo che tale estensione ha una parte sintattica non diversa da quella considerata nelle usuali logiche. Tuttavia a questa parte sintattica è affiancata anche una parte *valutativa* che mostra come si possano ottenere informazioni sul grado di verità del conseguente avendo informazioni sul grado di verità delle premesse.

Definizione 4.1. Una *regola di inferenza fuzzy* è una coppia $r = (r', r'')$, dove

- r' è una operazione parziale n -aria nell'insieme F delle formule il cui dominio denotiamo con $Dom(r)$;
- r'' è una operazione n -aria sull'intervallo U che conserva gli estremi superiori, cioè:

$$r''(x_1, \dots, \sup_{i \in I} y_i, \dots, x_n) = \sup_{i \in I} r''(x_1, \dots, y_i, \dots, x_n). \quad (4.1)$$

Indicheremo una applicazione di una regola di inferenza al modo seguente:

$$\frac{\alpha_1, \dots, \alpha_n}{r'(\alpha_1, \dots, \alpha_n)} \quad ; \quad \frac{\lambda_1, \dots, \lambda_n}{r''(\lambda_1, \dots, \lambda_n)}$$

che si può leggere dicendo:

SE le formule $\alpha_1, \dots, \alpha_n$ sono state provate con gradi (almeno) $\lambda_1, \dots, \lambda_n$, rispettivamente

ALLORA la formula $r'(\alpha_1, \dots, \alpha_n)$ si può ritenere provata con grado (almeno) $r''(\lambda_1, \dots, \lambda_n)$.

Naturalmente, data una semantica M , le regole di inferenza che interessano devono essere “corrette”, cioè deve risultare che, dato un modello $m \in M$, se $m(\alpha_1), \dots, m(\alpha_n)$ soddisfano i constraint $\lambda_1, \dots, \lambda_n$ allora $m(r'(\alpha_1, \dots, \alpha_n))$ soddisfa il constraint $r''(\lambda_1, \dots, \lambda_n)$. Questo significa che se $m(\alpha_1) \geq \lambda_1, \dots, m(\alpha_n) \geq \lambda_n$ allora $m(r'(\alpha_1, \dots, \alpha_n)) \geq r''(\lambda_1, \dots, \lambda_n)$. Stante la crescenza di r'' , ciò equivale ad affermare che

$$m(r'(\alpha_1, \dots, \alpha_n)) \geq r''(m(\alpha_1), \dots, m(\alpha_n)).$$

Un modo semplice per ottenere regole di inferenza fuzzy è il seguente.

Definizione 4.2. Data una regola di inferenza classica r' n -aria ed una norma triangolare continua $\otimes$, ponendo $r''(\lambda_1, \dots, \lambda_n) = \lambda_1 \otimes \dots \otimes \lambda_n$ se $n \neq 1$ e $r''(\lambda_1) = \lambda_1$, otteniamo una regola fuzzy $r = (r', r'')$ che prende il nome di $\otimes$ -estensione canonica di r' .

Ad esempio la $\otimes$ -estensione canonica del Modus Ponens è la regola r tale che:

$$Dom(r) = \{(\alpha, \alpha \Rightarrow \beta) : \alpha \in F, \beta \in F\},$$

$$r'(\alpha, \alpha \Rightarrow \beta) = \beta,$$

$$r''(\lambda_1, \lambda_2) = \lambda_1 \otimes \lambda_2.$$

La $\otimes$ -estensione canonica della regola di generalizzazione è la regola ottenuta ponendo

$$Dom(r) = F,$$

$$r'(\alpha) = \forall x \alpha,$$

$$r''(\lambda) = \lambda.$$

La $\otimes$ -estensione canonica della regola di particolarizzazione è data da

$$Dom(r) = \{\forall x \alpha : \alpha \in F\},$$

$$r'(\forall x \alpha) = \alpha(x/t),$$

$$r''(\lambda) = \lambda.$$

La $\otimes$ -estensione canonica della regola di introduzione di $\wedge$ è la regola r tale che:

$$\begin{aligned} \text{Dom}(r) &= \{(\alpha, \beta) : \alpha \in F, \beta \in F\}, \\ r'(\alpha, \beta) &= \alpha \wedge \beta, \\ r''(\lambda_1, \lambda_2) &= \lambda_1 \otimes \lambda_2. \end{aligned}$$

Un ulteriore strumento da specificare per definire un apparato deduttivo fuzzy è l'insieme fuzzy $a : F \rightarrow U$ di assiomi logici. La scelta di a deve essere tale che $a \subseteq m$ per ogni modello $m \in M$ e quindi tale che $a \subseteq \text{Tau}$.

Definizione 4.3. Chiamiamo *sistema deduttivo fuzzy* una coppia $\Sigma = (a, RI)$ dove a è un sottoinsieme fuzzy di F , detto l'*insieme fuzzy degli assiomi logici*, ed RI è un insieme di regole di inferenza fuzzy.

Definizione 4.4. Una *dimostrazione fuzzy* π di una formula α è una successione $\alpha_1, \dots, \alpha_m$ di formule tale che $\alpha_m = \alpha$, insieme ad una successione di relative *giustificazioni*. Ciò significa che per ogni formula α_i , deve essere specificato se:

- (i) se α_i è stata assunta come assioma logico, oppure
- (ii) se α_i è stata assunta come ipotesi, oppure
- (iii) se α_i è stata ottenuta tramite una regola di inferenza (in questo caso deve essere anche indicata quale regola si è utilizzata ed a quali delle formule "già dimostrate" $\alpha_1, \dots, \alpha_{i-1}$ si è applicata).

Indichiamo con $l(\pi)$ la lunghezza della dimostrazione π . Si osservi che se $l(\pi) = 1$, allora sono possibili solo due casi. Il primo consiste nell'assumere α come ipotesi, il secondo nell'assumere α come assioma logico. Come nel caso classico, per ogni $i \leq m$, il segmento iniziale $\alpha_1, \dots, \alpha_i$ è una dimostrazione di α_i che noi denotiamo con π_i . Differentemente dal caso classico, in una dimostrazione è necessario specificare le giustificazioni in quanto differenti giustificazioni della stessa formula, come vedremo, possono dare luogo a differenti valutazioni. Chiamiamo *insieme fuzzy di ipotesi* un qualunque insieme fuzzy di formule $v : F \rightarrow U$.

Definizione 4.5. Sia $v : F \rightarrow U$ un insieme fuzzy di ipotesi e π una dimostrazione, allora l'*informazione* $I(\pi, v)$ fornita da π dato

v è il numero definito per induzione sulla lunghezza m di π al modo seguente:

Se $l(\pi) = 1$, allora poniamo

$I(\pi, v) = a(\alpha_m)$ se α_m è stata assunta come assioma logico;

$I(\pi, v) = v(\alpha_m)$ se α_m è stata assunta come ipotesi.

Se $l(\pi) > 1$ allora poniamo

$$I(\pi, v) = \begin{cases} a(\alpha_m) & \text{se } \alpha_m \text{ è stata assunta come assioma logico,} \\ v(\alpha_m) & \text{se } \alpha_m \text{ è stata assunta come ipotesi,} \\ r''(I(\pi_{s(1)}, v), \dots, I(\pi_{s(n)}, v)) & \text{se } \alpha_m = r'(\alpha_{s(1)}, \dots, \alpha_{s(n)}) \\ & \text{dove } 1 \leq s(1) < m, \dots, 1 \leq s(n) < m. \end{cases}$$

Detta α la formula provata da π , il significato di $I(\pi, v)$ è:

π assicura che, data l'informazione v , α vale almeno con grado $I(\pi, v)$.

Per esempio, nel primo capitolo v è l'insieme fuzzy tale che $v(\text{Small}(1)) = 1$ e $(\text{Small}(n) \Rightarrow \text{Small}(n+1)) = 0.9$, per ogni $n \in N$. Una dimostrazione π_n della formula $\text{Small}(n)$ è data dalla successione

$$\begin{array}{ll} \text{Small}(1), & \text{(per ipotesi)} \\ \text{Small}(1) \Rightarrow \text{Small}(2), & \text{(per ipotesi)} \\ \text{Small}(2), & \text{(per MP)} \\ \text{Small}(2) \Rightarrow \text{Small}(3), & \text{(per ipotesi)} \\ \text{Small}(3), & \text{(per MP)} \\ \dots, & \\ \text{Small}(n), & \text{(per MP)}. \end{array}$$

L'informazione fornita da una tale dimostrazione è ottenuta per ricorsione tramite le equazioni

$$I(\pi_1, v) = 1 \text{ e } I(\pi_i, v) = 0.9 \times I(\pi_{i-2}, v).$$

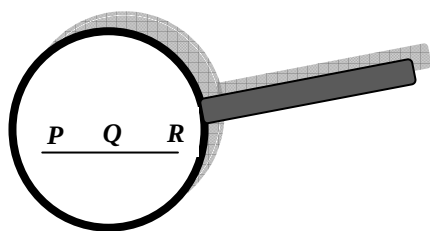
5. Un esempio: risolvere il paradosso della indiscernibilità

Abbiamo già visto esempi di dimostrazioni “approssimate” quando abbiamo esposto il modo di affrontare i paradossi del tipo mucchio di grano nella logica fuzzy. Facciamo ulteriori esempi in relazione al cosiddetto “paradosso” di Poincaré. Tale paradosso si riferisce alla relazione di indistinguibilità sottolineando il fatto che, a dispetto della intuizione comune, questa relazione non può essere transitiva (si veda [Gerla 2008]). Esponiamo una forma tipica di tale paradosso.

Teorema 5.1. Non è possibile percepire la differenza di peso di un corpo di 1 grammo da un corpo di 100 grammi.

Dim. Supponiamo di avere una successione di corpi $c_1, \dots, c_m$ di pesi crescenti da 1 a 100 grammi, ma la cui crescita è tanto impercettibilmente da non permetterci di distinguere la differenza tra un peso ed il successivo. Questo significa che valgono asserzioni del tipo $Ind(c_1, c_2), Ind(c_2, c_3), \dots, Ind(c_{m-1}, c_m)$. Se si ammettesse la proprietà transitiva, come sembra naturale ammettere, allora dalle prime due asserzioni si potrebbe dimostrare l'asserzione $Ind(c_1, c_3)$. Da tale asserzione e da $Ind(c_3, c_4)$ sarebbe possibile dimostrare $Ind(c_1, c_4)$. Proseguendo in tale modo si arriverebbe a dimostrare $Ind(c_1, c_m)$, cioè che 1 grammo non è distinguibile da cento grammi. $\square$

Ora una soluzione immediata di paradossi di questo tipo è semplicemente asserire che la nostra intuizione circa la transitività della relazione di indiscernibilità è sbagliata. A conferma di un tale punto di vista si potrebbe argomentare al modo seguente. E' naturale considerare indiscernibili due punti la cui distanza sia minore di un dato reale positivo ε , e quindi rappresentare l' indiscernibilità tramite la relazione $ind = \{(P, Q) : d(P, Q) \leq \varepsilon\}$.



Per tale relazione è evidente che non vale la proprietà transitiva. Infatti P, Q ed R sono tre punti allineati e successivi tali che la distanza tra il primo ed il secondo e quella tra il secondo ed il terzo sono uguali ad ε allora P è indiscernibile da Q , Q è indiscernibile da R ma P non risulta essere indiscernibile da R . Tuttavia questa soluzione “drastica” mi sembra discutibile poiché non è rispettosa della nostra intuizione la quale suggerisce che l' indiscernibilità dovrebbe pur avere una qualche forma debole di transitività.

Una soluzione differente la possiamo trovare nell'ambito della logica fuzzy assumendo che la indiscernibilità sia una similarità.

Teorema 5.2. Nella logica fuzzy la nozione di similarità permette di rappresentare l'argomentazione di Poincaré conservandone il contenuto intuitivo ma evitando il suo carattere paradossale.

Dim. Riferendoci al linguaggio utilizzato nella dimostrazione del paradosso, consideriamo la teoria fuzzy ottenuta aggiungendo l'assioma che dice che Ind è transitiva (accettato con grado 1) assiomi del tipo $Ind(c_{i-1}, c_i)$ accettati con grado $\lambda \neq 1$ (la fuzziness entra in questo punto). Supponiamo inoltre che la norma triangolare sia tale che $\lambda^{m-1} = 0$. Sotto tali condizioni possiamo formalizzare l'argomentazione di Poincaré nel modo seguente dove si è utilizzata la regola di introduzione della congiunzione ed il Modus Ponens:

Passo 1.

- Poiché
 $Ind(c_1, c_2)$ [con grado λ]
 - e
 $Ind(c_2, c_3)$ [con grado λ]
 possiamo asserire
 $Ind(c_1, c_2) \wedge Ind(c_2, c_3)$. [con grado $\lambda \otimes \lambda$]
 - Conseguentemente, poiché
 $Ind(c_1, c_2) \wedge Ind(c_2, c_3) \Rightarrow Ind(c_1, c_3)$ [con grado 1]
 possiamo asserire
 $Ind(c_1, c_3)$ [con grado $\lambda \otimes \lambda$]

Passo 2.

- Poiché $Ind(c_1, c_3)$ [con grado λ^2]
 - e
 $Ind(c_3, c_4)$ [con grado λ]
 possiamo asserire
 $Ind(c_1, c_3) \wedge Ind(c_3, c_4)$. [con grado λ^3]
 - Conseguentemente, poiché
 $Ind(c_1, c_3) \wedge Ind(c_3, c_4) \Rightarrow Ind(c_1, c_4)$ [con grado 1]
 possiamo asserire
 $Ind(c_1, c_4)$ [con grado λ^3]
 ...

Passo m-1.

- Poiché

$$Ind(c_1, c_{m-1}) \quad [\text{con grado } \lambda^{m-2}]$$

- e

$$Ind(c_{m-1}, c_m) \quad [\text{con grado } \lambda]$$

possiamo asserire

$$Ind(c_1, c_{m-1}) \wedge Ind(c_{m-1}, c_m) \quad [\text{con grado } \lambda^{m-1}]$$

- Conseguentemente, poiché

$$Ind(c_1, c_{m-1}) \wedge Ind(c_{m-1}, c_m) \Rightarrow Ind(c_1, c_m) \quad [\text{con grado } 1]$$

possiamo asserire

$$Ind(c_1, c_m) \quad [\text{con grado } \lambda^{m-1}].$$

Abbiamo pertanto dimostrato la formula $Ind(c_1, c_m)$ ma la dimostrazione permette di concludere solo che tale formula è vera almeno con grado $\lambda^{m-1} = 0$. Ciò non è affatto paradossale poiché equivale a dire che tale dimostrazione non è capace di fornire nessuna informazione sulla asserzione $Ind(c_1, c_m)$. $\square$

6. Ma meglio indebolire la proprietà transitiva

La soluzione ora proposta non sembra completamente convincente in quanto, come posto ad esempio in rilievo in [De Cock e Kerre 2003], l'ipotesi $\lambda \neq 1$ non è giustificata in una attenta formulazione del paradosso. Infatti che io non sia in grado di distinguere il peso di c_i dal peso di c_{i+1} è un fatto completamente vero e può magari essere provato da una serie di esperimenti ripetuti. Per fornire una risposta a tale obiezione, in [Gerla 2008] propongo di indebolire la nozione di equivalenza fuzzy ammettendo una "transitività debole". Ciò può essere fatto in due modi sostanzialmente equivalenti. Cominciamo dal primo che consiste semplicemente nell'assegnare a tale transitività un grado minore di 1 considerando pertanto un sistema di assiomi del tipo seguente:

$$A1 \quad \forall x Ind(x, x), \quad [1]$$

$$A2 \quad \forall x \forall y (Ind(x, y) \Rightarrow Ind(y, x)), \quad [1]$$

$$A3.1 \quad \forall x \forall y \forall z (Ind(x, z) \wedge Ind(y, z) \Rightarrow Ind(x, y)) \quad [\lambda]$$

con $\lambda \neq 1$.

Proposizione 6.1. Supponiamo che la relazione di indiscernibilità verifichi A1, A2, ed A3.1 con $\lambda \neq 1$. Allora è possibile ripetere l'argomentazione di Poincaré evitandone il carattere paradossale.

Dim. Supponiamo di muoverci nell'ambito della logica di Łukasiewicz e che $\lambda^{m-2} = 0$. Allora il ragionamento paradossale di Poincaré può essere riformulato al modo seguente.

Passo 1

Poiché

$Ind(c_1, c_2)$ [con grado 1]

- e

$Ind(c_2, c_3)$ [con grado 1]

possiamo asserire

$Ind(c_1, c_2) \wedge Ind(c_2, c_3)$ [con grado $1 \otimes 1$]

- Pertanto, poiché

$Ind(c_1, c_2) \wedge Ind(c_2, c_3) \Rightarrow Ind(c_1, c_3)$ [con grado λ]

possiamo asserire

$Ind(c_1, c_3)$ [con grado λ^1]

Passo 2.

- Poiché

$Ind(c_1, c_3)$ [con grado λ^1]

- e

$Ind(c_3, c_4)$ [con grado 1]

possiamo asserire

$Ind(c_1, c_3) \wedge Ind(c_3, c_4)$ [con grado λ^1]

- Pertanto, poiché

$Ind(c_1, c_3) \wedge Ind(c_3, c_4) \Rightarrow Ind(c_1, c_4)$ [con grado λ]

possiamo asserire

$Ind(c_1, c_4)$ [con grado λ^2]

...

Passo m-2.

- Poiché

$Ind(c_1, c_{m-1})$ [con grado λ^{m-3}]

- e

$Ind(c_{m-1}, c_m)$ [con grado 1]

possiamo asserire

$Ind(c_1, c_{m-1}) \wedge Ind(c_{m-1}, c_m)$ [con grado λ^{m-2}]

- Pertanto, poiché

$Ind(c_1, c_{m-1}) \wedge Ind(c_{m-1}, c_m) \Rightarrow Ind(c_1, c_m)$ [con grado λ]

possiamo concludere

$Ind(c_1, c_m)$ [con grado λ^{m-2}].

Essendo $\lambda^{m-2} = 0$, tale dimostrazione, pur essendo corretta, non fornisce nessuna informazione sul valore di verità di $Ind(c_1, c_m)$. Pertanto la contraddizione non sussiste. $\square$

Una seconda, a mio parere più interessante idea è che nel parlare di transitività di una relazione si debba tenere conto della capacità di un elemento di potere essere distinto dai rimanenti. Infatti, supponiamo che Ind non sia transitiva, allora possiamo considerare una valutazione della formula

$$Dis(z) \equiv \forall x \forall y (Ind(x, z) \wedge Ind(y, z) \Rightarrow Ind(x, y)).$$

La proprietà espressa da tale formula è che due qualunque oggetti che sono indiscernibili da z sono anche indiscernibili tra loro.⁵³ In altre parole $Dis(z)$ asserisce che z è adeguato come elemento di confronto. Se dis è l'insieme fuzzy che interpreta Dis , allora $dis(z)$ è una misura del buon comportamento di z rispetto alla proprietà transitiva. Ovviamente,

$$dis(z) = Inf\{e(x, z) \otimes e(y, z) \rightarrow e(x, y) : x, y \in S\} \quad (6.1)$$

Ora, la definizione data di Dis permette di provare la formula

$$\forall x \forall y \forall z (Ind(x, z) \wedge Ind(y, z) \wedge Dis(z) \Rightarrow Ind(x, y)), \quad (6.2)$$

o, se si vuole, che

$$dis(x, z) \otimes dis(y, z) \otimes dis(z) \leq dis(x, y) \quad (6.3)$$

e ciò suggerisce la considerazione di un sistema di assiomi del tipo

$$A1 \quad \forall x Ind(x, x),$$

$$A2 \quad \forall x \forall y (Ind(x, y) \Rightarrow Ind(y, x)),$$

$$A3.2 \quad \forall x \forall y \forall z (Ind(x, z) \wedge Ind(y, z) \wedge P(z) \Rightarrow Ind(x, y)).$$

Definizione 6.2. Sia $\lambda \neq 1$, sia m tale che $\lambda^{m-2} = 0$ e supponiamo che la relazione di indiscernibilità verifichi $A1$, $A2$, ed $A3.2$ insieme ad una interpretazione di P con valori inferiori a λ . Allora è possibile ripetere l'argomentazione di Poincaré evitandone il carattere paradossale.

Dim. Consideriamo una relazione di indiscernibilità soddisfacente $A1$, $A2$, $A3.2$ insieme ad una interpretazione di P a valori inferiori a λ , supponiamo che le formule $Ind(c_n, c_{n+1})$ valgano con grado 1 e che le formule $P(c_i)$ valgano con un grado $\lambda_i \leq \lambda$. Inoltre

⁵³ Da notare che tale proprietà coincide con la prima delle "nozioni comuni" esposte nel libro primo degli Elementi di Euclide.

sia m tale che $\lambda^{m-1} = 0$. In queste ipotesi possiamo formalizzare l'argomentazione di Poincaré al modo seguente:

Passo 1.

- Poiché
 $Ind(c_1, c_2)$ [con grado 1]
- e
 $Ind(c_2, c_3)$ [con grado 1]
- e
 $P(c_2)$ [con grado λ_2]
- possiamo asserire
 $Ind(c_1, c_2) \wedge Ind(c_2, c_3) \wedge P(c_2).$ [con grado $1 \otimes 1 \otimes \lambda_2$]
- Pertanto, poiché
 $Ind(c_1, c_2) \wedge Ind(c_2, c_3) \wedge P(c_2) \Rightarrow Ind(c_1, c_3)$ [con grado 1]
- possiamo asserire
 $Ind(c_1, c_3)$ [con grado λ_2]

Passo 2.

- Poiché
 $Ind(c_1, c_3)$ [con grado λ_2]
- e
 $Ind(c_3, c_4)$ [con grado 1]
- e
 $P(c_3)$ [con grado λ_3]
- possiamo asserire
 $Ind(c_1, c_3) \wedge Ind(c_3, c_4) \wedge P(c_3).$ [con grado $\lambda_2 \otimes \lambda_3$]
- Pertanto, poiché
 $Ind(c_1, c_3) \wedge Ind(c_3, c_4) \wedge P(c_3) \Rightarrow Ind(c_1, c_4)$ [con grado 1]
- possiamo asserire
 $Ind(c_1, c_4)$ [con grado $\lambda_2 \otimes \lambda_3$]
- ...

Passo $m-2$.

- Poiché
 $Ind(c_1, c_{m-1})$ [con grado $\lambda_2 \otimes \dots \otimes \lambda_{m-2}$]
- e
 $Ind(c_{m-1}, c_m)$ [con grado 1]
- e
 $P(c_{m-1})$ [con grado λ_{m-1}]
- possiamo asserire
 $Ind(c_1, c_{m-1}) \wedge Ind(c_{m-1}, c_m) \wedge P(c_{m-1})$ [con grado $\lambda_2 \otimes \dots \otimes \lambda_{m-1}$]

- Pertanto, poiché

$Ind(c_1, c_{m-1}) \wedge Ind(c_{m-1}, c_m) \wedge P(c_{m-1}) \Rightarrow Ind(c_1, c_m)$ [con grado 1]
possiamo concludere
 $Ind(c_1, c_m)$ [con grado $\lambda_2 \otimes \dots \otimes \lambda_{m-1}$].

Poiché $\lambda_2 \otimes \dots \otimes \lambda_{m-1} \leq \lambda^{m-2} = 0$, tale dimostrazione di $Ind(c_1, c_m)$, pur essendo corretta, non fornisce nessuna informazione sul valore di verità di $Ind(c_1, c_m)$. Pertanto la contraddizione non sussiste. $\square$

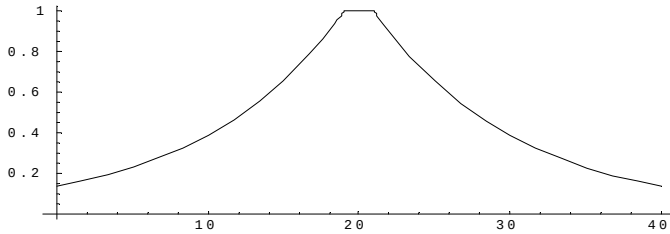
Da notare che ogni modello del sistema di assiomi A1, A2, A3.1 determina un modello del sistema di assiomi A1, A2, A3.2 con p costantemente uguale a λ e viceversa. Infatti dire che l'assioma

$$\forall x \forall y \forall z (Ind(x, z) \wedge Ind(y, z) \Rightarrow Ind(x, y))$$

è verificato con grado λ equivale a dire che $ind(x, z) \otimes ind(y, z) \rightarrow ind(x, y) \geq \lambda$ e quindi, per la proprietà del residuo, che $ind(x, z) \otimes ind(y, z) \leq ind(x, y)$.

La seguente proposizione fornisce un semplice esempio di modello del sistemi di assiomi considerati in questo paragrafo.

Proposizione 6.3. Fissato $\varepsilon \geq 0$, $\lambda \in U$ e $b = \lambda^{1/\varepsilon}$, definiamo nell'insieme dei reali R la relazione fuzzy $ind : R \times R \rightarrow U$ ottenuta ponendo $ind(x, y) = 1$ se $|x - y| \leq \varepsilon$ e $ind(x, y) = b^{|x - y| - \varepsilon}$ altrimenti. Allora, (R, ind) è un modello del sistema A1, A2, A3.1 di assiomi quando si considera come norma l'usuale prodotto.



Insieme fuzzy dei numeri indiscernibili da 20. Si noti che i numeri tra 19 e 21 non sono assolutamente discernibili da 20.

Notiamo esplicitamente che il fatto che *ind* non sia in grado di distinguere due numeri con distanza inferiore ad ε mentre registra distanze superiori comporta che la proprietà transitiva non può essere verificata. Ad esempio risulta che $ind(0,1) = ind(1,2) = 1$ mentre $ind(0,2) = 0.9$.

In [Gerla 2008] si prova che i modelli del sistema $A1, A2, A3.2$ di assiomi sono legati strettamente ad un approccio alla geometria senza punti in cui la nozione di diametro di una regione gioca un ruolo fondamentale.

7. Fondere l'informazione fornita da dimostrazioni diverse

Nella logica classica avere una dimostrazione di una asserzione non è poi tanto differente dall'averne due o dieci. Almeno se si guarda solo alla questione se tale asserzione possa essere considerata un teorema o meno. Nel caso della logica fuzzy le cose stanno differentemente e questo fatto a mio parere è una delle caratteristiche più affascinanti di tale logica.

Per fare un esempio torniamo alla dimostrazione che abbiamo esposto nel primo capitolo come soluzione della logica fuzzy del paradosso del mucchio di grano. Tale dimostrazione utilizza come unico fatto la formula $Small(1)$ considerata vera con grado 1. Tuttavia è anche ragionevole, ad esempio, assumere con grado 1 anche le formule $Small(2), \dots, Small(10)$. In questo caso oltre la dimostrazione π_1 già fornita nel primo capitolo di $Small(3)$, (in cui $I(\pi_1, v) = 0.81$), noi potremmo considerare anche la dimostrazione π_2 costituita dai passi:

$P(2)$ in quanto ipotesi

$P(2) \Rightarrow P(3)$ in quanto ipotesi

$P(3)$ per *MP* dalle due formule precedenti.

In tale caso $I(\pi_2, v) = 0.9$. D'altra parte, se π_3 è la dimostrazione (di lunghezza 1) che consiste nell'assumere $Small(3)$ come ipotesi, allora $I(\pi_3, v) = 1$. In definitiva tre diverse dimostrazioni della stessa formula forniscono tre diverse risposte !

Come ormai dovrebbe essere chiaro, la cosa è solo apparentemente contraddittoria. Infatti le informazioni fornite dalle tre dimostrazioni sono rispettivamente:

- $Small(3)$ vale almeno con grado 0.81,
- $Small(3)$ vale almeno con grado 0.9,
- $Small(3)$ vale almeno con grado 1.

Messe insieme tali informazioni equivalgono ad affermare che tale formula è vera con grado 1.

Appare allora naturale proporre la seguente definizione.

Definizione 7.1. Dato un sistema deduttivo fuzzy Σ , chiamiamo *operatore di deduzione associato a Σ* l'operatore $D : U^F \rightarrow U^F$ definito ponendo, per ogni $v \in U^F$ ed $\alpha \in F$,

$$D(v)(\alpha) = \text{Sup}\{I(\pi, v) : \pi \text{ dimostrazione di } \alpha\}. \quad (7.1)$$

Il significato di $D(v)(\alpha)$ è ancora

data l'informazione v , α vale almeno con grado $D(v)(\alpha)$,

ma ora possiamo anche affermare che:

$D(v)(\alpha)$ è la valutazione migliore possibile che il nostro sistema inferenziale ci permette di ottenere sul valore di verità di α data l'informazione v .

L'operatore $D : U^F \rightarrow U^F$ deve essere visto come un operatore che associa ad ogni insieme fuzzy v di ipotesi l'insieme fuzzy $D(v)$ dei relativi teoremi.

Giungiamo ora alla definizione principale della logica fuzzy.

Definizione 7.3. Data una semantica fuzzy M ed un apparato deduttivo Σ , diremo che la coppia (M, Σ) costituisce una *logica fuzzy* se l'operatore di conseguenza logica definito da M coincide con l'operatore di deduzione definito da Σ .

Quando sono verificate le condizioni date da tale definizione diciamo anche che vale un *teorema di completezza*.

Data una semantica, non è detto che esista sempre un apparato deduttivo per cui valga il teorema di completezza. Il viceversa invece vale.

Proposizione 7.4. Dato un apparato deduttivo fuzzy Σ , esiste sempre una semantica rispetto alla quale vale il teorema di completezza. Detto D il relativo operatore di deduzione, tale semantica coincide con l'insieme M dei punti fissi di D .

Dim. Anticipando i contenuti dell'ultimo paragrafo di questo capitolo, osserviamo che D è un operatore di chiusura e quindi

$D(v)$ è il minimo punto fisso di D contenente v . In altre parole,
 $D(v) = \cap \{m \in M : m \supseteq v\} = Lc(v)$. $\square$

Coerentemente con questo fatto, nel seguito considereremo apparati deduttivi di cui non diamo esplicitamente una semantica. Naturalmente un tale tipo di semantica non è molto soddisfacente in quanto abbiamo una idea di modello che, visto tramite l'insieme fuzzy delle formule in esso vere, corrisponde alla completezza dell'informazione. Ad esempio sarebbe ragionevole avere che tutti gli insiemi fuzzy in M siano massimali in M .

8. Oltre l'intervallo dei reali

In questo libro abbiamo scelto di riferirci all'intervallo U come all'insieme dei valori di verità. D'altra parte questo viene fatto dalla maggioranza degli studiosi della teoria degli insiemi fuzzy, Esistono infatti molte considerazioni a favore di tale scelta, ragioni che si possono riassumere nel fatto che la natura varia in modo continuo e che la nozione di numero reale sembra lo strumento principale per la rappresentazione del continuo. Esaminando la cosa più da vicino, cominciamo con l'osservare che nello scegliere un insieme V di valori di verità dovremo sicuramente fare in modo che esista un elemento 0, che corrisponde all'essere totalmente falso, ed un elemento 1, che corrisponde all'essere totalmente vero. Inoltre deve esistere un ordinamento in V per potere affermare che una asserzione è "più vera" di un'altra (cioè per potere rappresentare la natura comparativa della vaghezza). Ad esempio se "Mario è alto" e "Carlo è alto" sono due asserzioni valutate con valori in V allora la conoscenza di tali valori dovrebbe permettere di asserire se Mario è più alto di Carlo o viceversa. Se vogliamo che tale confronto sia sempre possibile, V deve essere un insieme totalmente ordinato (ovviamente con 0 come minimo ed 1 come massimo). E' ragionevole inoltre supporre che V sia denso se vogliamo ad esempio applicare i nostri formalismi al caso che dati due oggetti grandi se ne possa trovare uno nuovo di grandezza intermedia. Infine l'ordinamento in V deve essere completo se non altro per motivi tecnici. Infatti la completezza interviene sia nella interpretazione dei due quantificatori, sia nel processo di "fusione" dell'informazione fornita dall'insieme delle possibili dimostrazione di una formula.

Appare allora evidente perché U sia un ottimo candidato per essere l'insieme dei valori di verità di una logica fuzzy. Tuttavia quella di U non è la sola scelta ragionevole. Ad esempio esistono proprietà vaghe per cui non sempre è possibile una comparazione. Per illustrare tale possibilità consideriamo la proprietà “essere bravo” riferita a studenti di una scuola. Allora una valutazione in U di tale predicato comporterebbe che, dati due studenti che non abbiano lo stesso livello di bravura, sia sempre possibile dire chi è più bravo di un altro. D'altra parte se, come di fatto avviene, il giudizio di un insegnante viene espresso da un voto da 0 a dieci, allora non è difficile tradurre i diversi voti con corrispondenti gradi di verità in U (ad esempio considerando l'insieme $\{0, 1/10, 2/10, \dots, 10/10\}$).

E' completamente soddisfacente questo modo di procedere ?

Per certi aspetti direi di sì: ad esempio il Preside potrebbe decidere di premiare in ciascuna classe l'alunno che ha il voto maggiore. Tuttavia questo modo di rappresentare la nozione di bravo non è completamente adeguato in quanto uno studente potrebbe essere bravo per taluni aspetti e carente per altri. Ad esempio si potrebbe rappresentare il grado di bravura di uno studente con la coppia $(0.4, 0.7)$ dove 0.4 rappresenta il grado di conoscenza delle materie insegnate e 0.7 il grado di creatività ed autonomia di giudizio. Oppure, dovendo esprimere un giudizio globale, si potrebbe tenere presente il grado di bravura nelle singole materie. Più in generale, potrebbe essere significativo “misurare” il grado di bravura con una n -pla $(x_1, \dots, x_n)$ di valori in U dove x_i rappresenta il grado di bravura rispetto al criterio i -esimo. In tale caso la struttura di valutazione dovrebbe essere il reticolo non totalmente ordinato U^n e non il reticolo totalmente ordinato U .

9. Altre forme di informazione: constraints e bireticoli

Abbiamo già messo in rilievo che l'apparato deduttivo di un sistema inferenziale fuzzy non elabora valori di verità ma informazioni sui valori di verità delle formule considerate. La confusione che a volte si manifesta in letteratura deriva dal fatto che si sono sempre considerate informazioni del tipo “constraint inferiore” cioè del tipo “il valore di verità di α è maggiore o uguale a λ ” e si è portati quindi ad identificare un tale constraint con il valore λ e quindi la classe dei constraint con U . Tuttavia

non esistono motivi per non considerare altri tipi di informazione, ad esempio

“il valore di verità di α è minore di $3/4$ ”

oppure

“il valore di verità di α si colloca tra $1/3$ e $3/4$ ”.

Di fatto, la forma più generale di informazione che potrebbe essere considerata è del tipo

“il valore di verità di α appartiene ad X ”

dove X è un sottoinsieme qualunque di U . Possiamo pertanto identificare il nostro sistema di informazione con una classe $Cons$ di sottoinsiemi di U i cui elementi chiamiamo *pezzi di informazione* oppure *constraint*. In $Cons$ si considera un ordinamento $\leq_i$ che si definisce come il duale della relazione di inclusione

$$X \leq_i Y \Leftrightarrow X \supseteq Y$$

e che tiene conto del livello di informazione. Infatti se X ed Y sono due informazione su un valore di verità λ e $X \supseteq Y$, allora dobbiamo ritenere che Y fornisca più informazioni su λ di quanto faccia X . Ancora, se ho una famiglia $(X)_{i \in I}$ di constraints su λ allora la “fusione” di tale informazione è data dall’intersezione $\bigcap_{i \in I} X_i$ che d’altra parte coincide con l’estremo superiore rispetto $\leq_i$. Pertanto, poiché tale possibilità di fusione è essenziale, è utile ipotizzare che $Cons$ sia chiuso rispetto all’intersezione.

Definizione 9.1. Chiamiamo *sistema di informazione su U* un qualunque sistema di chiusura su U .

Ad esempio si potrebbero considerare come sistema di informazione la classe di tutti gli intervalli chiusi in U (compreso l’insieme vuoto) oppure la classe di tutti gli intervalli del tipo $[\lambda, 1]$. Dato un sistema di informazione, una regola di inferenza dovrebbe stabilire come le informazioni sui possibili valori di verità delle premesse consentano di calcolare informazioni sul valore di verità della conclusione. Dovremmo allora considerare regole di inferenze del tipo

$$\frac{\alpha_1, \dots, \alpha_n}{r'(\alpha_1, \dots, \alpha_n)}, \quad \frac{C_1, \dots, C_n}{r''(C_1, \dots, C_n)}$$

con $C_1, \dots, C_n$ e $r''(C_1, \dots, C_n)$ elementi di $Cons$. L’interpretazione di una tale regola sarebbe che:

- se si è provato che le formule $\alpha_1, \dots, \alpha_n$ hanno valori di verità verificanti $C_1, \dots, C_n$
- allora la formula $r'(\alpha_1, \dots, \alpha_n)$ ha valore di verità verificante $r''(C_1, \dots, C_n)$.

L'ipotesi di continuità sulla parte valutativa di una regola deve allora riferirsi all'ordinamento $\leq_i$ e non all'ordinamento usuale in U . Inoltre un sistema di assiomi logici dovrebbe essere una funzione $la : For \rightarrow Cons$ il cui significato è che, anche se non si ha nessuna informazione, per ogni formula α , il valore di verità di α appartiene ad $la(\alpha)$ (cioè verifica il constraint $la(\alpha)$). Supponiamo che M sia una semantica fuzzy, allora la nozione di tautologia si ottiene definendo $Tau : F \rightarrow Cons$ tramite l'equazione

$$Tau(\alpha) = \langle \{m(\alpha) : m \in M\} \rangle$$

dove, come viene fatto usualmente nella teoria dei sistemi di chiusura, con $\langle X \rangle$ si intende l'intersezione di tutti gli elementi di $Cons$ contenenti X . Ad esempio, se $Cons$ è l'insieme degli intervalli chiusi, allora nella logica di Zadeh $Tau(\alpha \vee \neg \alpha) = [0.5, 1]$ e $Tau(\alpha \wedge \neg \alpha) = [0, 0.5]$. Infine, ogni dimostrazione d di α fornisce una informazione $I(d, v)$ e quindi, la massima informazione sul valore di verità di α si ottiene ponendo

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

Un tale modo di vedere la logica fuzzy è ancora tutto da sviluppare.

Un altro possibile approccio alla logica fuzzy è quello di valutare una formula α con una coppia $(\lambda, \mu) \in U \times U$ dove λ è una misura dei motivi a favore della verità di α e μ una misura dei motivi contro la verità di α . In questo caso si vengono a definire due possibili relazioni d'ordine. La prima, che si riferisce alla verità, si ottiene ponendo

$$(\lambda, \mu) \leq_t (\lambda', \mu') \Leftrightarrow \lambda \leq \lambda', \mu \geq \mu'.$$

Questo significa che una formula α viene considerata più vera di una formula α' se esistono più motivi a favore e meno motivi contro. La seconda relazione, che si riferisce al grado di informazione, si ottiene ponendo

$$(\lambda, \mu) \leq_k (\lambda', \mu') \Leftrightarrow \lambda \leq \lambda', \mu \geq \mu'.$$

Tale definizione è coerente con il fatto che un aumento di informazione può aumentare sia i motivi a favore che i motivi contro una certa asserzione. Le coppie $(1, 0)$ e $(0, 1)$, che risultano

il massimo ed il minimo rispetto $\leq_i$ rappresenteranno il vero ed il falso, rispettivamente. Le coppie (0,0) e (1,1), che risultano essere il minimo ed il massimo rispetto a $\leq_k$, rappresentano la completa mancanza di informazione e l'informazione completamente contraddittoria, rispettivamente. In questo modo a $U \times U$ viene fornita di una doppia struttura di reticolo limitato.

Strutture di tale tipo in cui sono definiti due ordinamenti sono studiate in letteratura sotto il nome di *bi reticoli*. Anche in questo caso si dovrebbe precisare come dare una adeguata definizione di apparato inferenziale per una logica fuzzy basata sui bi reticoli.

10. Perché il principio di induzione è falso

Avendo definito in maniera un po' più precisa le nozioni di semantica e di apparato deduttivo per una logica fuzzy, è possibile ora riguardare sia il paradosso del mucchio di grano che la questione del principio di induzione matematica.

Per quanto riguarda il primo mostriamo che in effetti qualunque predicato vago genera un paradosso di tale tipo.

Proposizione 10.1. È possibile generare un paradosso simile a quello del mucchio di grano in corrispondenza di ogni predicato vago capace di variare gradualmente dal vero al falso.

Dim. Sia P il nome di una proprietà vaga in un insieme non vuoto D e supponiamo che P sia interpretata da un sottoinsieme fuzzy s di D che varia gradualmente tra 0 ed 1. Questa condizione la possiamo precisare supponendo che, fissato $\varepsilon > 0$, esista una successione $(d_n)_{n \in \mathbb{N}}$ di elementi di D tale che

$$i) \quad s(d_1) = 1$$

$$ii) \quad s(d_n) - s(d_{n+1}) \leq \varepsilon$$

$$iii) \quad s(d_n)_{n \in \mathbb{N}} \text{ è decrescente e tende a zero.}$$

Allora, riferendoci alla logica di Łukasiewicz,

$$|P(d_1)| = 1 \quad \text{e} \quad |P(d_n) \Rightarrow P(d_{n+1})| = 1 - s(d_n) + s(d_{n+1}) \geq 1 - \varepsilon.$$

Ora se ε è molto piccolo volendo inquadrare la situazione nell'ambito della logica classica siamo portati ad assumere che il valore $|P(d_n) \Rightarrow P(d_{n+1})|$ sia approssimativamente uguale ad 1. Ma in tale caso, sempre tramite la logica classica, potremmo riprodurre il paradosso del mucchio di grano riuscendo a dimostrare che, per ogni fissato n , vale $P(d_n)$ in contrasto con

l'ipotesi per cui la valutazione $s(x_n)$ tende a zero e quindi, da un certo punto in poi, $P(n)$ venga percepita come falsa.

Da tale dimostrazione segue che 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,
- due punti qualunque dello spazio sono sempre *vicini*,
- se una persona è giovane rimarrà sempre giovane (Corollario: non esistono persone vecchie) ...

Passando al principio di induzione, ci accorgiamo che il suo rapporto con la logica fuzzy è alquanto problematico.

Teorema 10.2. La logica fuzzy non permette di risolvere il paradosso del mucchio di grano nella forma che utilizza il principio di induzione !

Dim. Nel primo capitolo, quando abbiamo esaminato il punto di vista finitista sul paradosso del mucchio di grano, abbiamo già osservato come il principio di induzione rende possibile una riformulazione del paradosso del mucchio di grano. Infatti proviamo ad accettare, insieme alle due formule

i) $Pic(1)$

ii) $\forall n(Pic(n) \Rightarrow Pic(n+1))$

anche il principio di induzione applicato al predicato Pic

iii) $Pic(1) \Rightarrow ((\forall n(Pic(n) \Rightarrow Pic(n+1))) \Rightarrow \forall n Pic(n))$.

Allora due semplici applicazioni del *MP* permettono di provare la formula $\forall n Pic(n)$ e quindi, per particolarizzazione, $Pic(\underline{m})$ per ogni prefissato m . Proviamo ora a risolvere tale forma del paradosso con la logica fuzzy supponendo di assegnare al principio di induzione un valore di verità 1 ed attribuendo, come al solito, ad i) e ii) i valori 1 e 0.9, rispettivamente. Allora la conclusione $\forall n Pic(n)$ può essere provata con grado $1 \otimes 0.9 \otimes 1 = 0.9$ in contrasto con il fatto che $\forall n Pic(n)$ sia falsa. $\square$

Naturalmente è ragionevole accettare che il principio di induzione non possa essere esteso ai predicati fuzzy.

Teorema 10.3. In ambito fuzzy il principio di induzione in N non è valido.

Dim. Una prima argomentazione a favore di tale asserzione è di carattere sintattico ed è una semplice estensione di quella utilizzata per dimostrare la proposizione 8.2. Supponiamo di assegnare al principio di induzione un valore di verità λ da stabilirsi. Attribuiamo ad $i)$ e $ii)$ i valori 1 e $\mu \neq 0$, rispettivamente. Allora la conclusione $\forall n \text{ Pic}(n)$ può essere provata con grado $\mu \otimes \lambda$. Poiché sappiamo che $\forall n \text{ Pic}(n)$ è falsa, deve essere $\mu \otimes \lambda = 0$. Se $\otimes$ è la norma del prodotto o del minimo o una qualunque norma priva di divisori dello zero, ciò comporta che $\lambda=0$. Negli altri casi, indichiamo con $(\mu_n)_{n \in N}$ una successione crescente e convergente ad 1, allora per ogni n possiamo riformulare il paradosso con una interpretazione del predicato Pic che comporti che $\mu = \mu_n$. Pertanto, per la continuità di $\otimes$,

$$\lambda = 1 \otimes \lambda = (\lim_{n \rightarrow \infty} \mu_n) \otimes \lambda = \lim_{n \rightarrow \infty} (\mu_n \otimes \lambda) = 0 \otimes \lambda = 0.$$

Una seconda dimostrazione della non validità del principio di induzione è di carattere semantico e consiste nel valutare direttamente il grado di verità del principio di induzione

$$\forall P(P(1) \Rightarrow (\forall n(P(n) \Rightarrow P(n+1)) \Rightarrow \forall n P(n)))$$

espresso come una formula nella logica del secondo ordine. Supponiamo che P sia interpretato dall'insieme fuzzy $s : N \rightarrow U$ di N definito ponendo:

$$s(n) = |1 - (n-1)/k|_n$$

con k fissato in N . In altre parole $s(n) = 1 - (n-1)/k$ per ogni $n < k+1$ e $s(n) = 0$ altrimenti. Supponendo ad esempio di riferirci alla logica di Łukasiewicz,

$$s(n) \rightarrow s(n+1) = |s(n+1) - s(n) + 1|_n = 1 - 1/k \text{ se } n < k+1.$$

Abbiamo pertanto che $|\forall n(P(n) \Rightarrow P(n+1))| = 1 - 1/k$. D'altra parte, essendo $|P(1)| = 1$ e $|\forall n(P(n))| = 0$, possiamo calcolare che

$$\begin{aligned} &|P(1) \Rightarrow (\forall n(P(n) \Rightarrow P(n+1)) \Rightarrow \forall n(P(n)))| \\ &= |\forall n(P(n) \Rightarrow P(n+1)) \Rightarrow \forall n(P(n))| \\ &= 1 - |\forall n(P(n) \Rightarrow P(n+1))| + |\forall n(P(n))| = 1 - 1 + 1/k = 1/k. \end{aligned}$$

Allora abbiamo provato che per ogni $k > 0$ esiste un modo di interpretare la variabile P che rende vera con grado $1/k$ la formula

$$P(1) \Rightarrow (\forall n(P(n) \Rightarrow P(n+1)) \Rightarrow \forall n(P(n)))$$

Poiché la quantificazione universale comporta il considerare l'estremo inferiore di tali valori, il principio di induzione è valido con grado zero, cioè è completamente falso.⁵⁴ $\square$

Stante il noto legame tra il principio di induzione e la nozione di buon ordine, non deve sorprendere che accanto al principio di induzione si debba negare anche che l'ordinamento usuale in N sia un buon ordinamento.⁵⁵

Teorema 10.4. In ambito fuzzy l'ordine usuale in N non è di buon ordine.

Dim. Ripigliando un esempio fatto dal filosofo James Cargile, supponiamo che in una pozza d'acqua vi sia un girino e che tale girino segua la sua normale evoluzione fino a diventare una rana. Se tale evoluzione è stata filmata continuamente, tale fenomeno da origine ad una sequenza di fotogrammi, $F_1, \dots, F_n$. Ora sicuramente nei primi fotogrammi compare un girino e negli ultimi compare una rana. Si pone allora il seguente problema:

qual è il primo fotogramma in cui compare una rana ?

Se l'ordinamento usuale di N fosse un buon ordine tale fotogramma F_m esisterebbe in quanto sarebbe il minimo dell'insieme $\{n \in N : f_n \text{ rappresenta una rana}\}$. Allo stesso tempo l'esistenza di un tale m appare assurda. Infatti in F_{m-1} compare un girino ed in F_m compare (improvvisamente) una rana. Pertanto N non è ben ordinato. $\square$

⁵⁴ È da notare che, in accordo con il significato che assume il quantificatore universale nelle logiche a più valori, non è stata dimostrata l'esistenza di un particolare predicato che rende falso il principio di induzione. Piuttosto si è visto che, comunque si consideri k in N è possibile trovare un predicato (vago) per cui il principio di induzione vale con grado minore o uguale a $1/k$.

⁵⁵ Ricordiamo che una relazione d'ordine $\leq$ in un insieme S è detta di *buon ordinamento* se ogni sottoinsieme non vuoto di S ammette elemento minimo (naturalmente ci si riferisce ai sottoinsiemi classici e non a quelli fuzzy). Il tipico esempio di buon ordinamento è quello che si considera usualmente nell'insieme dei numeri naturali.

11. La logica come studio degli operatori di chiusura

Se D è l'operatore di deduzione della logica classica, allora è immediato dimostrare che sono verificate le seguenti proprietà:

- i) $D(X) \supseteq X$ (inclusione)
- ii) $X \subseteq Y \Rightarrow D(X) \subseteq D(Y)$ (monotonia)
- iii) $D(X) = D(D(X))$ (chiusura).

Il significato di i) è che il meccanismo di deduzione è *conservativo*, cioè le ipotesi da cui si parte non sono messe in discussione durante il processo deduttivo. La proprietà ii) è particolarmente importante poiché esprime la monotonia dei sistemi inferenziali considerati.⁵⁶ Essa asserisce che se si è pervenuti ad una conclusione α a partire da una informazione X , allora l'acquisizione di ulteriore informazione non può farci cambiare opinione circa α . Infine iii) significa che il processo dimostrativo rappresentato da D è saturo, cioè che non è possibile provare, a partire da $D(X)$ niente che non sia già in $D(X)$.

Le tre proprietà considerate possono essere espresse dicendo che D è un operatore di chiusura. La loro importanza ha spinto Tarski ed i logici della sua scuola a proporre un approccio astratto alla logica che viene appunto vista come studio di operatori di chiusura.

In particolare è possibile considerare il reticolo U^F ed estendere l'approccio di Tarski alla logica fuzzy (si veda ad esempio [Gerla 1997], [Gerla 1998], [Gerla 2001]b). Qui ci limitiamo a dimostrare solo i seguenti teoremi.

Teorema 11.1. L'operatore di conseguenza logica Lc associato ad una semantica fuzzy è un operatore di chiusura nel reticolo U^F .

Teorema 11.2. L'operatore di deduzione $D : U^F \rightarrow U^F$ associato ad un sistema deduttivo fuzzy è un operatore di chiusura nel reticolo U^F .

12. Il problema della compattezza

L'operatore di deduzione della logica classica oltre ad essere un operatore di chiusura verifica un'altra fondamentale proprietà: è compatto. Ricordiamo la nozione di compattezza.

⁵⁶ Come è noto, esistono anche logiche non monotone.

Definizione 12.1. Dato un insieme non vuoto S , un operatore monotono $H : P(S) \rightarrow P(S)$ è detto *compatto* se, per ogni sottoinsieme X di S ,

$$H(X) = \cup \{H(X_f) : X_f \text{ sottoinsieme finito di } X\}.$$

In altre parole H è compatto se per ogni $x \in H(X)$ esiste un sottoinsieme finito X_f di X tale che $x \in H(X_f)$.⁵⁷ La compattezza dell'operatore di deduzione D esprime la finitezza della nozione di dimostrazione. Infatti se $\alpha \in D(T)$ allora esiste una dimostrazione π di α sotto ipotesi T ed in tale dimostrazione è utilizzato solo un insieme finito T_f di elementi di T . La proprietà di compattezza è di notevole importanza: da essa (e dal teorema di completezza) segue ad esempio che se ogni parte finita di una teoria T ammette un modello allora l'intera teoria T ammette un modello.⁵⁸ Inoltre, come abbiamo già osservato, esso permette (insieme al teorema di deduzione) di ridurre le dimostrazioni sotto ipotesi a dimostrazioni di tautologie. Tuttavia la vera ragione alla base dell'importanza della compattezza è il fatto che essa è un presupposto della effettività di ogni processo deduttivo. Non affronto tale questione poiché richiederebbe conoscenze di teoria della computabilità un po' più avanzate ed, in particolare, la teoria degli "*enumeration operator*" (si veda [Rogers 1967]).

In ogni caso si pone il problema di come si possa estendere la definizione di compattezza alla logica fuzzy e se la logica fuzzy sia compatta. Ad esempio Pavelka propone la seguente definizione [Pavelka 1979]a.

Definizione 12.2. Diciamo che un operatore $H : U^F \rightarrow U^F$ è *Pavelka-compatto*, se per ogni formula α ed ogni insieme fuzzy v di formule esiste un sottoinsieme fuzzy finito v_f di v tale che:

⁵⁷ Questa proprietà caratterizza in un certo senso i processi finitari di costruzione. Infatti se si interpreta $H(X)$ come l'insieme delle cose che possono essere in qualche modo costruite utilizzando gli elementi di X allora tale condizione equivale a dire che tutto ciò che può essere costruito a partire da X in realtà può essere costruito utilizzando un numero finito di oggetti in X .

⁵⁸ Infatti la teoria T è consistente in quanto una eventuale contraddizione sarebbe derivabile da una parte finita T_f di T in contrasto con il fatto che per ipotesi T_f ammette un modello. Essendo T consistente deve ammettere anche un modello.

$$H(v)(\alpha) = H(v_f)(\alpha).$$

Purtroppo tale definizione non si dimostra adeguata.

Proposizione 12.3. La logica fuzzy non è in generale Pavelka-compatta.

Dim. Sia $G(n)$ è un predicato che possiamo interpretare come “ n è un numero grande” e sia τ la teoria fuzzy che pone $\tau(G(n)) = 1 - 1/n$. Supponendo di avere assunto la formula $G(n) \Rightarrow \exists x(G(x))$ come assioma logico di grado 1, la dimostrazione

$$G(n) \quad [\text{con grado } 1-1/n]$$

$$G(n) \Rightarrow \exists x(G(x)) \quad [\text{con grado } 1]$$

$$\exists x(G(x)) \quad [\text{con grado } 1-1/n]$$

dimostra la formula $\exists x(G(x))$ con grado $1-1/n$. Fondendo tutte le dimostrazioni ottenute al variare di n , otteniamo che $D(\tau)(\exists x G(x)) \geq \sup_{n \in \mathbb{N}} 1-1/n = 1$. Tuttavia è evidente che nessuna parte finita di τ consente di provare $\exists x G(x)$ con grado 1. $\square$

La questione è che nel determinare il grado di dimostrabilità di una formula α dobbiamo tenere conto anche del processo di fusione delle informazioni fornite dalle varie dimostrazioni di α e tale processo è di carattere infinitario. Come risposta a tale questione in una serie di lavori (si veda ad esempio [Biacino e Gerla 2002] e Gerla [2006]b), si propone una nozione diversa di compattezza che viene identificata con la nozione di continuità in un reticolo.

Definizione 12.4. Sia L un reticolo completo, allora un insieme C di elementi di L è chiamato *diretto* se

$$x \in C, y \in C \Rightarrow \exists z \in C, x \leq z, y \leq z.$$

Chiamiamo *limite induttivo* di C l'elemento $z = \sup(C)$. Un operatore crescente $J : L \rightarrow L$ è detto *continuo* se conserva i limiti induttivi.

Se scriviamo $z = \lim C$ per indicare il limite induttivo di C possiamo anche dire che $J : L \rightarrow L$ è continuo se, per ogni insieme diretto C ,

$$J(\lim C) = \lim J(C). \quad (12.1)$$

E' evidente che la composizione di due operatori continui è un operatore continuo.

Proposizione 12.5. Sia L un reticolo completo e $J : L \rightarrow L$ un operatore continuo con la proprietà di inclusione. Posto allora

$$H(x) = \text{Sup}_n J^n(x), \quad (12.2)$$

$H(x)$ è il minimo punto fisso di J maggiore o uguale ad x . L'operatore $H : L \rightarrow L$ definito in questo modo è un operatore di chiusura continuo che ha gli stessi punti fissi di J .

13. La compattezza come continuità

Possiamo ora proporre la seguente nozione di compattezza.

Definizione 13.1. Dato un insieme non vuoto S ed il reticolo dei sottoinsiemi fuzzy U^S di S , chiamiamo *compatto* un operatore in U^S che sia continuo.

Si pone il problema se una tale identificazione della compattezza con la continuità sia giustificabile. In proposito possiamo dire che in questo modo si ottiene una definizione che estende quella classica. Infatti se L è il reticolo dei sottoinsiemi di un dato insieme, allora gli operatori continui coincidono con quelli compatti. Inoltre nel caso, che a noi interessa, del reticolo dei sottoinsiemi fuzzy di un dato insieme possiamo caratterizzare gli operatori continui in termini di finitezza [Murali 1991]. Per farlo, utilizziamo la relazione $\leq$ definita ponendo, per ogni coppia s e t di insiemi fuzzy, $s \leq t$ se, per ogni $x \in \text{Supp}(s)$, risulta che $s(x) \leq t(x)$.

Proposizione 13.2. Un operatore $J : U^S \rightarrow U^S$ è continuo se e solo se, per ogni sottoinsieme fuzzy s di S ,

$$J(s) = \cup \{J(s_f) \mid s_f \text{ è finite ed } s_f \leq s\} \quad (13.1)$$

Un altro motivo a favore della definizione 13.1 è che abbastanza ampia da permettere di provare che tutti gli operatori di deduzione delle logiche fuzzy sono compatti. A tale scopo cominciamo a considerare l'operatore di conseguenze immediate.

Definizione 13.3. Dato un sistema deduttivo fuzzy (V, al) , poniamo

$$K(v)(\alpha) = \text{Sup}\{r''(v(\alpha_1), \dots, v(\alpha_n)) : (r', r'') \in V, r'(\alpha_1, \dots, \alpha_n) = \alpha\}.$$

Definiamo allora l'operatore di conseguenze immediate $H : U^F \rightarrow U^F$ ponendo, per ogni $v \in U^F$,

$$H(v) = v \cup al \cup K(v). \quad (13.3)$$

$K(v)$ rappresenta l'insieme fuzzy delle formule che possiamo derivare da v tramite l'applicazione di una regola di inferenza. Interpretiamo $H(v)$ come l'insieme fuzzy delle formule che possiamo derivare in un passo, cioè con una dimostrazione di lunghezza uno. Dire che v è un punto fisso per H equivale a dire che v è chiuso rispetto alle regole di inferenza e contiene l'insieme fuzzy degli assiomi logici.

Proposizione 13.4. L'operatore di conseguenza immediata H è compatto, monotono e con la proprietà di inclusione.

Per la proposizione 12.5, essendo H continuo e con la proprietà di inclusione, per ogni v in U^F l'insieme fuzzy $\cup_n H^n(v)$ è il minimo punto fisso di H contenente v . Vale inoltre il seguente teorema.

Teorema 13.5. Detto D l'operatore di deduzione, risulta che

$$D(v) = \cup_n H^n(v). \quad (13.4)$$

Pertanto l'operatore di deduzione è un operatore di chiusura compatto.

$D(v)$ risulta pertanto essere il più piccolo insieme fuzzy di formule contenente v , chiuso per regole di inferenza e contenente l'insieme fuzzy degli assiomi logici.

Concludiamo dicendo che il motivo principale a favore della identificazione della compattezza con la continuità è che la continuità è una condizione necessaria perché l'apparato inferenziale della logica fuzzy abbia i caratteri della effettività. Non vuole essere compito di questo libro entrare nel merito di una tale questione che richiederebbe un po' di tempo. Tuttavia il discorso è analogo a quello per cui per le funzioni di variabile reale la continuità è un presupposto necessario alla computabilità.

14. Paradigma della logica fuzzy e logica probabilistica

Abbiamo già visto che probabilità e vaghezza sono due fenomeni totalmente diversi. Tuttavia il formalismo proposto per la logica fuzzy si dimostra essere interessante anche per la logica probabilistica (si veda [Gerla 1994]b, [Gerla 1994]c, [Gerla

1997]). Basta prendere in considerazione la semantica consistente nella classe M_p delle probabilità finitamente additive definite nell'insieme F delle formule del calcolo proposizionale. Nel seguito denoteremo con **0** la contraddizione $\alpha \wedge \neg \alpha$ e con **1** la tautologia $\alpha \vee \neg \alpha$.

Definizione 14.1. Chiamiamo *semantica della logica probabilistica* la classe M_p delle funzioni $m : F \rightarrow U$ tali che

- a) se α è logicamente equivalente ad α' allora $m(\alpha) = m(\alpha')$
- b) $m(\mathbf{0}) = 0$
- c) $m(\mathbf{1}) = 1$
- d) $m(\alpha \vee \beta) = m(\alpha) + m(\beta) - m(\alpha \wedge \beta)$.

L'interpretazione di un modello $m \in M_p$ è che per ogni formula α il numero $m(\alpha)$ rappresenta la probabilità che, nel *mondo random* m , l'asserzione α sia vera. Accettando il paradigma entro cui si muove la logica fuzzy, la logica probabilistica che dobbiamo definire deve elaborare informazioni parziali su un modello m . Ad esempio in tale logica possiamo elaborare informazione del tipo

- “la probabilità di α oscilla tra 0.6 e 0.9”,
- “la probabilità di β è minore o uguale a 0.1”,
- “la probabilità di γ è maggiore di 0.7”.

Poiché $1 = m(\alpha \vee \neg \alpha) = m(\alpha) + m(\neg \alpha) - m(\alpha \wedge \neg \alpha) = m(\alpha) + m(\neg \alpha)$, risulta che $m(\neg \alpha) = 1 - m(\alpha)$. Pertanto, ripetendo quando detto nel paragrafo 1, è ragionevole rappresentare l'informazione disponibile tramite una funzione $v : F \rightarrow U$ e dire che m è un modello di v se, per ogni formula α , “la probabilità che si verifichi α è almeno $v(\alpha)$ ”.

Per quanto riguarda l'operatore di conseguenza logica, $Lc(v)$ viene ad essere l'estremo inferiore di tutte le probabilità che sono maggiori o uguali a v . In letteratura una funzione di questo tipo ha già un nome.

Definizione 14.2. Una funzione che si possa ottenere come estremo inferiore di una famiglia di probabilità finitamente additive prende il nome di *inviluppo inferiore* o, semplicemente, *inviluppo*.

Pertanto $Lc(v)$ è l'involuppo generato da v e gli involuppi sono i punti fissi di Lc . Essi rivestono una notevole importanza per rappresentare situazioni di informazione non completa in situazioni aventi carattere probabilistico (si veda ad esempio Papamarcou and Fine [1986]).

Una volta definita la semantica rimane la questione di cercare adeguati apparati inferenziali che permettano di calcolare $Lc(v)$. Tale questione è interessante per la costruzione di sistemi diagnostici ed è esaminata, ad esempio in [Weichselberger and Pöhlmann 1990] anche se in casi molto particolari. Nei miei lavori affronto tale questione proponendo un opportuno sistema di regole di inferenza che “funziona” (nel senso che permette di provare un teorema di completezza) ma che probabilmente potrebbe essere migliorato. Tale sistema si basa sui seguenti connettivi.

Definizione 14.3. Dati h e k , indichiamo con C_h^k il connettivo logico il cui significato è che $C_h^k(\alpha_1, \dots, \alpha_h)$ assume valore 1 solo se esistono almeno k formule vere nella sequenza $\alpha_1, \dots, \alpha_h$.

Chi si occupa di reti neurali può interpretare questo connettivo come un neurone con h linee di ingresso che “spara” un impulso se la somma degli impulsi in entrata supera k . Per il teorema di completezza funzionale naturalmente potrei definire tale connettivo e considerare $C_h^k(\alpha_1, \dots, \alpha_h)$ come un modo per indicare la formula $\bigvee \{ \alpha_{i(1)} \wedge \dots \wedge \alpha_{i(k)} : 1 \leq i(1) < \dots < i(k) \leq h \}$.

Definiamo la funzione M ponendo

$$M(\alpha_1, \dots, \alpha_h) = \max \{ k \in N : C_h^k(\alpha_1, \dots, \alpha_h) \text{ consistente} \}$$

per ogni $\alpha_1, \dots, \alpha_h$ in F . Allora la soddisfacibilità di v può essere caratterizzata tramite M .

Teorema 14.4. Un insieme fuzzy v di formule è soddisfacibile se e solo se per ogni $\alpha_1, \dots, \alpha_h$ in F ,

$$v(\alpha_1) + \dots + v(\alpha_h) \leq M(\alpha_1, \dots, \alpha_h). \quad (14.1)$$

L'apparato inferenziale si ottiene aggiungendo ad un apparato inferenziale per il calcolo proposizionale classico le seguenti regole.

Definizione 14.5. Siano h, m, k numeri naturali tali che $h \geq m \geq k$, e sia

$$D(h, m) = \{(\alpha_1, \dots, \alpha_h) : M(\alpha_1, \dots, \alpha_h) = m\},$$

allora chiamiamo h - m - k -regola di inferenza la regola $r = (r', r'')$ il cui dominio è $D(h, m)$ ed inoltre

$$r'(\alpha_1, \dots, \alpha_h) = C^k(\alpha_1, \dots, \alpha_h); \quad r''(x_1, \dots, x_h) = \left\lfloor \frac{x_1 + \dots + x_h - k + 1}{m - k + 1} \right\rfloor_n.$$

Il significato di una h - m - k -regola è chiaro solo per alcuni casi particolari. Per esempio la h -1-1-regola è la regola di introduzione della disgiunzione definita ponendo:

$$\frac{\alpha_1, \dots, \alpha_h}{\alpha_1 \vee \dots \vee \alpha_h} \quad ; \quad \frac{\lambda_1, \dots, \lambda_h}{\lambda_1 \oplus \dots \oplus \lambda_h}$$

dove $\alpha_1, \dots, \alpha_h$ devono essere mutualmente esclusivi è $\oplus$ è la conorma di Łukasiewicz. Ponendo $h = m = k$ otteniamo la regola di introduzione della congiunzione

$$\frac{\alpha_1, \dots, \alpha_h}{\alpha_1 \wedge \dots \wedge \alpha_h} \quad ; \quad \frac{\lambda_1, \dots, \lambda_h}{\lambda_1 \otimes \dots \otimes \lambda_h}$$

dove $\alpha_1, \dots, \alpha_h$ devono essere consistenti tra loro e $\otimes$ è la norma di Łukasiewicz.

In aggiunta alle regole ora date, dobbiamo anche considerare, per ogni h ed m , $h \geq m$, la h - m -regola di collapsing $c = (c', c'')$ definita in $D(h, m)$ ponendo $c'(\alpha_1, \dots, \alpha_h) = \mathbf{0}$ per ogni $\alpha_1, \dots, \alpha_h$ e

$$c''(\lambda_1, \dots, \lambda_h) = \begin{cases} 1 & \text{se } \lambda_1 + \dots + \lambda_h > m, \\ 0 & \text{altrimenti.} \end{cases}$$

Tale regola afferma che se le formule $\alpha_1, \dots, \alpha_h$ sono state provate con gradi $\lambda_1, \dots, \lambda_h$ che violano la condizione di consistenza (14.1) allora è possibile dimostrare la contraddizione $\mathbf{0}$ e quindi qualunque formula. Ad esempio la 2-1-regola di collapsing dice che se sono state dimostrate due formule α_1 e α_2 tra loro contraddittorie con gradi λ_1 e λ_2 con $\lambda_1 + \lambda_2 > 1$, allora possiamo asserire anche la contraddizione $\mathbf{0}$.

Non procedo ulteriormente nella descrizione di tale sistema, dicendo solo che esso permette di ottenere un teorema di completezza. Ciò permette di calcolare l'involuppo generato da una data funzione $v : F \rightarrow U$.

CAPITOLO 4

ESEMPI DI APPARATI INFERENZIALI FUZZY

1. L'approccio "graded" alla logica di Łukasiewicz

Si pone il problema se si possa trovare una logica fuzzy per cui valga un teorema di completezza. La risposta è positiva ed in questo paragrafo esponiamo l'esempio principale: la logica fuzzy di Łukasiewicz del primo ordine. Stante il carattere elementare che ho voluto dare a questo libro, non riporto le dimostrazioni rimandando chi fosse interessato ad una trattazione adeguata (si veda ad esempio [Novák 2007]). Si assume che il linguaggio contenga, per ogni $\lambda \in \bar{U}$ una costante proposizionale $\underline{\lambda}$ dove con $\bar{U}$ denotiamo l'insieme dei numeri razionali in U . Questo significa che nella formazione delle formule di tale linguaggio $\underline{\lambda}$ viene considerato come fosse una formula atomica e nella interpretazione delle formule il valore assegnato a $\underline{\lambda}$ è λ .⁵⁹ L'importanza di tale allargamento consiste nel fatto che i constraint chiusi possono essere espressi asserendo la verità (con grado 1) di alcune formule di tale linguaggio. Precisamente il fatto che il valore di verità di una formula α appartiene all'intervallo $[\lambda, \mu]$ può essere espresso con l'asserire che le due formule $\alpha \Rightarrow \underline{\mu}$ e $\underline{\lambda} \Rightarrow \alpha$ assumono il valore 1. In particolare, l'informazione rappresentata da un insieme fuzzy di ipotesi v può essere rappresentata dell'insieme (crisp) di formule $\{v(\alpha) \Rightarrow \alpha : \alpha \in F\}$. La struttura di valutazione considerata è l'algebra di Łukasiewicz $(U, \otimes, \rightarrow)$ dove ricordo che $\otimes$ e $\rightarrow$ sono definite ponendo

$$x \otimes y = |x + y - 1|_n ; x \rightarrow y = |y - x + 1|_n$$

I connettivi logici corrispondono al tipo di operazioni dell'algebra di Łukasiewicz e saranno $\&$, $\Rightarrow$. Come al solito altri

⁵⁹ Questo "trucco", anche se risulta efficiente, non è certo molto naturale. Nessuno ritiene che $\frac{3}{4}$ sia una asserzione e tantomeno che sia un'asserzione una espressione del tipo $\neg(1/4 \Rightarrow 4/5)$. Se quindi si vuole creare un modello che sia in qualche modo rappresentativo delle forme di ragionamento che coinvolgono i predicati vaghi, allora la soluzione proposta in questo paragrafo non mi sembra molto soddisfacente. D'altra parte, non sono a conoscenza di dimostrazioni del teorema di completezza che non utilizzino tali costanti proposizionali.

connettivi sono invece definiti. Ad esempio assumiamo che $\neg A$ sia una abbreviazione della formula $A \Rightarrow 0$ e corrisponde alla funzione $1-\lambda$. Inoltre $A \vee B$ è abbreviazioni della formula $(B \Rightarrow A) \Rightarrow A$ e $A \wedge B$ della formula $\neg(\neg A \vee \neg B)$. Tali connettivi corrispondono ad una interpretazione in U come massimo e minimo. Le regole di inferenza sono la $\otimes$ -estensione canonica del Modus Ponens

$$\frac{A, A \Rightarrow B}{B} \quad \frac{\lambda, \mu}{\lambda \otimes \mu}$$

e l'estensione canonica della regola di generalizzazione

$$\frac{A}{\forall x(A)} \quad \frac{\lambda}{\lambda}$$

Inoltre si accetta la seguente regola di introduzione delle costanti logiche

$$\frac{A}{\underline{\lambda} \Rightarrow A} \quad \frac{\lambda}{1}$$

Questa ultima regola è particolarmente significativa. Infatti permette in ogni istante di trasformare una affermazione di dimostrabilità con grado λ con una affermazione di dimostrabilità classica, cioè con grado 1.

Passiamo ora a definire l'insieme fuzzy degli assiomi logici. Tale insieme è la funzione $l_a : F \rightarrow U$ tale che $l(A) = \lambda$ se $A = \underline{\lambda}$ e $L(A) = 1$ se A coincide con uno degli assiomi della seguente lista

- R1** $A \Rightarrow (B \Rightarrow A)$
- R2** $(A \Rightarrow B) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$
- R3** $(\neg B \Rightarrow \neg A) \Rightarrow (A \Rightarrow B)$
- R4** $((A \Rightarrow B) \Rightarrow B) \Rightarrow ((B \Rightarrow A) \Rightarrow A)$
- T1** $\forall x(A) \Rightarrow A(x/t)$ (per ogni termine sostituibile t)
- T2** $\forall x(A \Rightarrow B) \Rightarrow (A \Rightarrow \forall x(B))$ con x non libera in B
- B** $(\underline{\lambda} \Rightarrow \underline{\mu}) \Leftrightarrow (\lambda \Rightarrow \mu)$.

Nei casi non specificati l_a assume valore 0.

Possiamo ora enunciare il teorema di completezza.

Teorema 1.1. Per la logica di Łukasiewicz del primo ordine vale il teorema di completezza.

Osserviamo che per tale logica vale anche la regola derivata della introduzione della congiunzione (osservazione che sarà utile quando parleremo di programmazione logica).

Proposizione 1.2. Nella logica di Łukasiewicz la regola di introduzione della congiunzione

$$\frac{A, B}{A \wedge B} \quad \frac{\lambda, \mu}{\lambda \otimes \mu}$$

è una regola derivata.

2. Metalogica fuzzy per logiche crisp

Le logiche fuzzy la cui semantica è di tipo vero-funzionale derivano dall'esigenza di trattare predicati vaghi, cioè predicati che possono essere attribuiti ad oggetti con gradi diversi. Un altro modo per ottenere logiche fuzzy è quello di riferirsi a predicati "crisp" ma di accettare meta-predicati fuzzy. Ad esempio indichiamo con TAU l'insieme delle tautologie della logica classica e supponiamo che T sia un insieme di ipotesi (o, se si vuole, di assiomi propri). Per potere applicare il teorema di deduzione, supponiamo che in T ci siano solo formule senza variabili libere. Allora nella logica classica la relazione $\vdash$ di "deduzione da ipotesi" può essere definita ponendo $T \vdash \alpha$ se $\alpha \in TAU$ oppure se

"esistono $\alpha_1 \in T, \dots, \alpha_n \in T$ tali che $\alpha_1 \wedge \dots \wedge \alpha_n \Rightarrow \alpha \in TAU$ ".

Ora è possibile ammettere che il meta-predicato "essere una ipotesi" sia vago cioè che ci sia una situazione in cui ipotesi diverse siano accettabili con gradi diversi. Ciò significa che dobbiamo riferirci ad insieme fuzzy v di ipotesi invece che ad un insieme crisp T e che, conseguentemente, dobbiamo valutare in U affermazioni del tipo

"esistono $\alpha_1 \in v, \dots, \alpha_n \in v$ tali che $\alpha_1 \wedge \dots \wedge \alpha_n \Rightarrow \alpha \in TAU$ ".

Tenendo conto che l'asserzione " $\alpha_1 \in v, \dots, \alpha_n \in v$ ", in quanto congiunzione di asserzioni atomiche, è valutata con $v(\alpha_1) \otimes \dots \otimes v(\alpha_n)$ e tenendo conto che il connettivo "esiste" viene interpretato in logica fuzzy tramite l'estremo superiore, la valutazione sarà data dal seguente numero

$$\text{Sup}\{v(\alpha_1) \otimes \dots \otimes v(\alpha_n) : \alpha_1 \wedge \dots \wedge \alpha_n \rightarrow \alpha \in TAU\}.$$

Queste considerazioni giustificano la seguente definizione.

Definizione 2.1. Data una norma triangolare $\otimes$ ed un insieme fuzzy di assiomi v , il *grado di dimostrabilità* di una formula α è il numero $D^\otimes(v)(\alpha)$ ottenuto ponendo $D^\otimes(v)(\alpha) = 1$ se $\alpha \in TAU$ e ponendo

$D^\otimes(v)(\alpha) = \text{Sup}\{v(\alpha_1) \otimes \dots \otimes v(\alpha_n) : \alpha_1 \wedge \dots \wedge \alpha_n \rightarrow \alpha \in TAU\}$ altrimenti.

Non è difficile provare che l'operatore $D^\otimes: U^F \rightarrow U^F$ definito in questo modo è un operatore di chiusura. Un modo più elegante di definire $D^\otimes$ è tramite una opportuna nozione di inclusione graduata.

Definizione 2.2. Sia $\otimes$ una norma triangolare ed S un insieme non vuoto, allora per ogni sottoinsieme finito o vuoto X ed ogni insieme fuzzy $s: S \rightarrow U$, chiamiamo $\otimes$ -grado di inclusione di X in v il numero $\text{Incl}_\otimes(X, s)$ definito ponendo

$$\text{Incl}_\otimes(X, s) = \begin{cases} 1 & \text{se } X = \emptyset \\ \text{Incl}_\otimes(X, s) = s(x_1) \otimes \dots \otimes s(x_n) & \text{se } X = \{x_1, \dots, x_n\}. \end{cases}$$

Tale definizione deriva dal fatto che l'inclusione di $\{x_1, \dots, x_n\}$ in s equivale alla congiunzione delle affermazioni

“ x_1 appartiene ad s ”, ..., “ x_n appartiene ad s ”.

Tale nozione graduata di inclusione estende quella classica e verifica le seguenti proprietà.

- i) $X \subseteq Y \Rightarrow \text{Incl}_\otimes(X, s) \geq \text{Incl}_\otimes(Y, s)$
- ii) $s \subseteq s' \Rightarrow \text{Incl}_\otimes(X, s) \leq \text{Incl}_\otimes(X, s')$
- iii) $\text{Incl}_\otimes(X \cup Y, s) \geq \text{Incl}_\otimes(X, s) \otimes \text{Incl}_\otimes(Y, s)$
- iv) $\text{Incl}_\otimes(X, s) \geq \text{Incl}_\otimes(X, s)$.

Ovviamente risulta che

$$D^\otimes(v)(\alpha) = \text{Sup}\{\text{Incl}_\otimes(X, s) : X \in \Pi_f(F), X \vdash \alpha\} \quad (2.1)$$

dove $\Pi_f(F)$ indica l'insieme delle parti finite o vuote di F .

L'operatore $D^\otimes$ definito in questo modo risulta essere l'operatore di deduzione del seguente semplicissimo sistema inferenziale fuzzy.

Definizione 2.3. Data una norma triangolare continua $\otimes$, chiamiamo $\otimes$ -estensione canonica della logica classica la logica fuzzy che si ottiene considerando le $\otimes$ -estensioni del Modus Ponens e della regola di particolarezzazione ed assumendo come

insieme di assiomi logici un insieme adeguato per la logica classica.

Osserviamo che una tale definizione di estensione può essere applicata anche a logiche crisp diverse da quella classica. Basta solo cambiare la relazione di deducibilità $\vdash$.

Teorema 2.4. L'estensione canonica del calcolo dei predicati della logica classica rispetto ad una data norma triangolare continua $\otimes$ ha come operatore di deduzione l'operatore $D^\otimes$.

3. Ogni situazione ha la sua norma: idempotenza

*Ma il discorrere è come il correre, e
non come il portare, ed un cavallo
barbero solo correrà più che cento
frisoni (Galileo Galilei: Il Saggiatore)*

Abbiamo visto che ogni norma triangolare permette di costruire un semplice apparato inferenziale fuzzy che estende la logica classica. Si pone allora la domanda:

quale norma deve essere scelta fra tutte le possibili ?

La mia opinione è che non sia possibile scegliere una volta per tutte ma che in contesti diversi siano necessarie norme diverse. Una distinzione fondamentale tra norme è se vale o meno l'idempotenza $x \otimes x = x$ (tenendo conto che l'unica norma idempotente è il minimo $\wedge$).

Cominciamo a considerare il primo caso. La logica corrispondente certamente non è adatta a spiegare il paradosso del mucchio di grano. Infatti in tale caso, poiché $(0.9)^{n-1} = 0.9$, l'affermazione "un mucchio con n chicchi è piccolo" sarebbe provata con grado 0.9 qualunque sia il valore fissato di n . Mostriamo invece due contesti in cui il minimo sembra particolarmente adatto.

Esempio 1: l'assioma della scelta. Nell'assiomatizzare le fondamentali teorie matematiche è spesso accaduto che alcuni assiomi sono sembrati meno affidabili di altri e che quindi si sia cercato o di dimostrarli o, comunque, di evitare il loro uso nelle dimostrazioni. Il primo celebre esempio è dato dall'assioma delle parallele nella geometria di Euclide. Come è noto, Euclide spesso

portava avanti dimostrazioni complicate pur di non utilizzare l'assioma delle parallele. Altri altrettanto celebri e più recenti esempi riguardano la teoria degli insiemi. Tipico a tale proposito è l'assioma della scelta. E' un fatto che tutti i matematici evitano se è possibile l'uso di tale assioma e che se sono costretti a farlo si sentono poi in dovere di avvisare il lettore e quasi di scusarsene. Allo stesso tempo nessun matematico, al di là della dimostrazione di indipendenza dell'assioma della scelta, dichiarerebbe che tale assioma è più falso o più vero degli altri assiomi. Piuttosto si tratta di una certa "diffidenza" nei suoi confronti difficilmente esprimibile in termini della logica classica. La logica fuzzy che utilizza il minimo riesce invece a rappresentare abbastanza bene un tale fenomeno. Infatti, ad esempio, indichiamo con ZF un sistema di assiomi per la teoria degli insiemi non contenente l'assioma della scelta ed indichiamo con A_S l'assioma della scelta. Allora la nostra diffidenza (ma non completo rifiuto) verso l'assioma della scelta potrebbe essere rappresentata dal seguente "insieme fuzzy degli assiomi che si accettano":

$$v(\alpha) = \begin{cases} 1 & \text{se } \alpha \in ZF-A_S \\ 0.8 & \text{se } \alpha \text{ coincide con } A_S \\ 0 & \text{altrimenti.} \end{cases}$$

Allora è evidente che, riferendoci alla $\wedge$ -estensione canonica della logica classica:

$$D^\wedge(v)(\alpha) = \begin{cases} 1 & \text{se è possibile dimostrare } \alpha \text{ senza } A_S \\ 0.8 & \text{se } A_S \text{ è essenziale per provare } \alpha \\ 0 & \text{se } \alpha \text{ non è conseguenza di } ZF+A_S. \end{cases}$$

Da notare che è la proprietà di idempotenza della norma del minimo che permette di trascurare il numero di volte che si è utilizzato l'assioma della scelta registrando solo se si è utilizzato o meno.

Naturalmente potremmo anche considerare un insieme fuzzy v di assiomi che assegni valori non nulli anche all'ipotesi del continuo o ad altri possibili assiomi che la comunità dei matematici ritiene meno "indiscutibili" di altri. In ambito geometrico è possibile fare un discorso simile in rapporto all'assioma delle parallele.

Esempio 2: percorsi facili. Supponiamo di volere costruire un sistema esperto riferito a possibili percorsi per escursioni in montagna ed alla opportunità di effettuare il percorso per una persona con maggiore o minore esperienza. Se scriviamo $Facile(p)$ per indicare che il percorso p è facile, un modo ragionevole di valutare la formula $Facile(p)$ potrebbe essere il seguente:

- i) valutiamo ogni possibile tipo di difficoltà con un numero tra 0 ed 1 (ad esempio un salto di un metro e mezzo potrebbe essere valutato 0.9),
- 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 $Facile(p)$ sia la negazione del grado di difficoltà di p .

Naturalmente ciò presuppone che nel dichiarare un percorso più o meno facile non si tiene conto ad esempio della sua lunghezza. Se indichiamo con x e y due percorsi successivi e con $x+y$ l'unione di tali percorsi, allora appare naturale accettare l'assioma

$$Facile(x) \wedge Facile(y) \Rightarrow Facile(x+y) \quad (\text{con grado } 1)$$

o, per meglio dire,

$$Facile(x) \Rightarrow (Facile(y) \Rightarrow Facile(x+y)) \quad (\text{con grado } 1).$$

Supponiamo ora di avere una successione di piccoli percorsi $p_1, p_2, \dots$ ciascuno successivo all'altro ed indichiamo con q_i l'unione dei percorsi $p_1, p_2, \dots, p_i$. Supponiamo inoltre che tutti i percorsi siano abbastanza facili e che quindi tutte le formule $Facile(p_i)$ siano, ad esempio, valutate 0.9. Allora:

- usando $Facile(q_1) \Rightarrow (Facile(p_2) \Rightarrow Facile(q_2))$ ed applicando due volte *MP* riusciamo a provare $Facile(q_2)$ con grado $0.9 \otimes 0.9$.
- usando $Facile(q_2) \Rightarrow (Facile(p_3) \Rightarrow Facile(q_3))$ riusciamo a provare $Facile(q_3)$ con grado $0.9 \otimes 0.9 \otimes 0.9$

...

- usando $Facile(q_{99}) \Rightarrow (Facile(p_{100}) \Rightarrow Facile(q_{100}))$ riusciamo a provare $Facile(q_{100})$ con grado $(0.9)^{100}$.

Ora nel caso della norma del prodotto o di quella di Łukasiewicz il valore di $(0.9)^{100}$ è o uguale o quasi uguale a zero. Questo significa che la dimostrazione ora esposta non fornisce nessun tipo di informazione su $Facile(q_{100})$. Tenendo invece conto del fatto che il grado di difficoltà di un percorso non dipende dalla

sua lunghezza, la logica che sembrerebbe fornire migliore informazione per questo tipo di fenomeno sembrerebbe essere quella in cui $\otimes$ è il minimo e quindi $0.9^{100} = 0.9$.

4. Ogni situazione ha la sua norma: non idempotenza

Differentemente dal caso in cui la norma è idempotente, i ragionamenti che sono formalizzati tramite norme non idempotenti hanno la seguente caratteristica:

più volte si utilizzano ipotesi poco affidabili 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". Ciò è 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 è molto vicina al Municipio,

allora cominciamo ad avere qualche esitazione a concludere che il bar è molto vicino al Municipio.

Dal punto di vista tecnico, l'opportunità di non utilizzare troppe volte ipotesi non completamente vere è conseguenza del fatto che, se x ed y sono valori di verità diversi da zero ed 1 allora $x \otimes y$ risulta essere strettamente minore sia di x che di y , cioè strettamente minore del minimo $x \wedge y$.

Tuttavia la norma del prodotto e quella di Łukasiewicz presentano notevoli differenze.

Esempio 1 (per l'uomo calvo è meglio Łukasiewicz).

Mentre la norma del prodotto si presta bene per il paradosso del mucchio di grano non sembra tuttavia molto adeguata per la soluzione del paradosso dell'uomo calvo. Supponiamo infatti di assegnare alla formula $C(10.000)$ grado di verità 1 ed alle formule $C(n) \Rightarrow C(n-1)$ grado di verità γ diverso da 1 e da 0, allora

procedendo in modo simile a quanto fatto per il paradosso del mucchio di grano riusciamo a provare $C(0)$, cioè che “un uomo con zero capelli è capelluto” ma la validità di tale dimostrazione è $\gamma^{10.000-1}$. Ora nel caso del prodotto, per quanto possa essere basso il valore $\gamma^{10.000-1}$ è comunque diverso da zero e ciò contrasta con il fatto che un uomo con zero capelli può ritenersi capelluto solo con grado zero. Invece, nel caso della norma di Łukasiewicz, se assegniamo all’implicazione $C(n) \Rightarrow C(n-1)$ grado di verità $\gamma = 1 - \delta$ con $\delta = 1/(10.000-1)$, allora, perveniamo alla conclusione $C(0)$ con grado $(\gamma)^{10.000-1} = 0$, cosa che sembra più corretta.

Esempio 2 (il paradosso dei test uguali). 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 in tali test 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 $0.5 \otimes 0.5$. D’altra parte è lecito attribuire alla implicazione

$$Supera_test(A) \wedge \neg(Supera_test(B)) \Rightarrow Preferibile(A,B)$$

grado 1. Da ciò segue che è possibile dimostrare $Preferibile(A,B)$ con grado $0.5 \otimes 0.5$. Allo stesso tempo, poiché le capacità dei candidati si sono mostrate uguali, l’asserzione $Preferibile(A,B)$ deve essere considerata completamente falsa (se si vuole, completamente ingiusta). Ne segue che necessariamente deve essere $0.5 \otimes 0.5 = 0$. Una tale conclusione in letteratura viene vista come un paradosso poiché se ci si riferisce alla norma del minimo oppure a quella del prodotto è effettivamente un paradosso. Tuttavia se si accetta la norma di Łukasiewicz accade proprio che $0.5 \otimes 0.5 = 0.5 + 0.5 - 1 = 0$.

Quanto detto in questo e nel precedente paragrafo pone in rilievo che non è ragionevole riferirsi ad una unica norma triangolare ma che situazioni diverse possono portare a logiche in cui la congiunzione è definita da norme diverse. Il seguente esempio mostra che potrebbe essere necessario considerare allo stesso tempo (nella stessa logica) norme diverse.

Esempio 3 (percorsi leggeri e percorsi sicuri). Arricchiamo l'esempio 2 del paragrafo precedente aggiungendo al predicato *Facile* anche il predicato *Leggero*. Questo secondo predicato non si tiene conto del grado di difficoltà ma della fatica maggiore o minore per effettuare un percorso. Dipende quindi fortemente dalla lunghezza dello stesso. Con le stesse notazioni di tale esempio supponiamo che ciascun segmento p_i sia leggero, cioè che sia breve e che non richieda sforzi fisici particolari. Possiamo rappresentare tale tipo di situazione con gli assiomi già visti per *Facile* mentre per *Leggero* si assumono gli assiomi:

$Leggero(p_i)$ (con grado 1),

$Leggero(x) \wedge Leggero(y) \Rightarrow Leggero(x+y)$ (con grado 0.9).

Allora, procedendo come al solito, possiamo provare $Facile(q_n)$ con grado 0.9^n , e $Leggero(q_n)$ con grado $= 0.9^{n-1}$.

Ora, poiché la facilità di un percorso non dipende dalla sua lunghezza, è ragionevole assumere che $0.9^n = 0.9$ per ogni intero n . Poiché la leggerezza di un percorso dipende in maniera determinante dalla sua lunghezza, ci aspettiamo che esista un intero m tale che, per ogni $n \geq m$, $0.9^{n-1} = 0$.

Possiamo ottenere entrambe le condizioni solo ammettendo che le norme triangolari utilizzate per *Facile* e *Leggero* siano differenti e precisamente la norma del minimo e quella di Łukasiewicz, rispettivamente. In tale caso abbiamo bisogno delle due diverse estensioni del Modus Ponens tramite tali norme. Per fare ciò abbiamo bisogno di due implicazioni $\Rightarrow_{\otimes}$ e $\Rightarrow_{\wedge}$ nel nostro linguaggio capaci di esprimere i due tipi di assiomi tramite le formule

$Facile(q_i) \Rightarrow_{\wedge} Facile(q_{i+1})$; $Leggero(q_i) \Rightarrow_{\otimes} Leggero(q_{i+1})$.

D'altra parte non è sorprendente che non ci si possa limitare nemmeno ad un numero finito di connettivi logici quando vogliamo trattare valori di verità in un continuo come U .

La situazione si complica ulteriormente se si osserva che lo stesso predicato *Facile* potrebbe a volte richiedere una interpretazione di $\otimes$ come prodotto. Infatti, indichiamo con $p_1, p_2, \dots$ una successione di percorsi tali che

- la pendenza di p_1 è nulla
- la pendenza di p_n è lievemente superiore alla pendenza di p_{n-1}
- a parte l'eventuale pendenza, i percorsi p_i non presentano ulteriori difficoltà.

In tali ipotesi allora è lecito assumere

$Facile(p_1) ; Facile(p_n) \Rightarrow Facile(p_{n+1})$.

Senza entrare nei particolari, è evidente che siamo in una situazione simile a quella del paradosso del mucchio di grano o dell'uomo calvo e quindi $\otimes$ non può più essere il minimo!

5. La logica fuzzy è monotona ?

Il fatto che l'operatore di conseguenza immediata in logica fuzzy sia monotono e conservativo significa che se aumentano le ipotesi aumentano anche le conseguenze da tali ipotesi e che asserzioni dimostrate ad un dato passo del processo deduttivo non possono essere successivamente annullate. Quando una logica verifica tali condizioni si dice che è *monotona*. In questo paragrafo vogliamo mostrare che, se entra in gioco anche il grado di inconsistenza di una teoria fuzzy, la monotonia della logica fuzzy è solo apparente.

Definizione 5.1. Detto D l'operatore di deduzione di una logica fuzzy, chiamiamo *grado di inconsistenza* di un insieme fuzzy di formule v il numero

$$Inc(v) = Inf\{D(v)(\alpha) : \alpha \in F\}. \quad (5.1)$$

Se $Inc(v) = 1$, allora diciamo che v è *totalmente inconsistente*.

Per fare un esempio, riferiamoci alla $\otimes$ -estensione canonica della logica classica del primo ordine. In tale caso abbiamo che il grado di inconsistenza ha una forma particolarmente semplice.

Proposizione 5.2. Consideriamo una $\otimes$ -estensione canonica del calcolo dei predicati classico e sia χ una contraddizione, allora per ogni insieme fuzzy v di formule risulta:

$$Inc(v) = D^{\otimes}(v)(\chi) = Sup\{Incl_{\otimes}(X, v) : X \text{ è inconsistente}\}.$$

Dim. E' sufficiente osservare che, poiché $\chi \vdash \alpha$ per ogni formula α , risulta che $D^{\otimes}(v)(\chi) \leq D^{\otimes}(v)(\alpha)$. $\square$

In altri termini per le $\otimes$ -estensioni il grado di inconsistenza di una teoria fuzzy v è il grado con cui si può provare una contraddizione oppure, equivalentemente, il grado di verità dell'asserzione "v contiene un insieme contraddittorio di formule".

Se v è totalmente inconsistente allora, come nella logica classica, ogni formula può essere provata con grado 1. Può tuttavia capitare che il grado di inconsistenza non sia né nullo né uguale ad 1. In tale caso se siamo in grado di provare una formula α con un grado $D(v)(\alpha)$ diverso da zero ma risulta che $D(v)(\alpha) = Inc(v)$, allora di fatto non esiste nessuna ragione in favore della validità di α . Questo perché il grado di validità di una formula α è rappresentato più dalla differenza $D(v)(\alpha) - Inc(v)$ che dal valore $D(v)(\alpha)$. Una interessante conseguenza di questo fatto è che:

la logica fuzzy, nonostante il fatto che D sia un operatore monotono, non può essere considerata una logica monotona.

Infatti, poiché sia $D(v)(\alpha)$ che $Inc(v)$ crescono al crescere dei valori di v , il valore $D(v)(\alpha) - Inc(v)$ non cresce necessariamente al crescere dei valori di v . Supponiamo ad esempio che l'informazione disponibile ad un istante t sia v e che tale informazione conduca ad affermare α , cioè $D(v)(\alpha) > Inc(v)$. Supponiamo che successivamente, all'istante t' , sia disponibile nuova informazione v' in modo che v sia esteso ad un insieme fuzzy più grande $v \cup v'$ ma che tale nuova informazione sia parzialmente inconsistente con la precedente. Allora, grazie al fatto che il grado di inconsistenza è cresciuto, è possibile che $D(v \cup v')(\alpha) = Inc(v \cup v')$. In tale modo, mentre all'istante t l'informazione v suggeriva la validità dell'asserzione α , all'istante successivo t' la nuova informazione porta ad un cambiamento di opinione per cui non esiste nessuna ragione per credere nella validità di α .

In base a tali considerazioni sarebbe forse più corretto sostituire l'operatore di deduzione con il seguente operatore di deduzione *normalizzato*.

Definizione 5.3. Sia D l'operatore di deduzione di un apparato inferenziale fuzzy, allora definiamo l'*operatore di deduzione normalizzato* D_n ponendo, per ogni $v \in U^F$, $D_n(v)$ uguale alla teoria totalmente inconsistente nel caso in cui $Inc(v) = 1$ e

$$D_n(v)(\alpha) = \frac{D(v) - Inc(v)}{1 - Inc(v)} \quad (5.2)$$

se $Inc(v) \neq 1$.

E' interessante osservare che $D_n(v)(\alpha) = 1$ implica che $D_n(v \cup v')(\alpha) = 1$ per ogni nuova informazione v' . Ciò significa che se siamo assolutamente sicuri della validità della formula α , allora nessuna nuova informazione può cambiare la nostra opinione.

Ovviamente, nonostante il fatto che D_n sia più significativo di D , da un punto di vista tecnico è preferibile riferirsi a D . Infatti all'operatore D è possibile applicare la teoria dei punti fissi per operatori monotoni, teoria che è molto semplice e potente. D'altra parte il considerare D non comporta nessuna perdita di informazione in quanto, dato D , possiamo calcolare D_n .

Un esempio. Esaminiamo un esempio tipico di ragionamento non monotono e vediamo come tale ragionamento si possa inquadrare nella logica fuzzy purché si tenga conto del grado di inconsistenza. Consideriamo un linguaggio contenente i seguenti predicati:

$Sv(x)$: x è un cittadino svedese,

$Gi(x)$: x ha genitori italiani,

$Pr(x)$: x è protestante.

e poniamo

$$T = \{Sv(b) \Rightarrow Pr(b), Sv(b) \wedge Gi(b) \Rightarrow \neg Pr(b), Sv(b)\}.$$

Tale teoria afferma che se b è svedese allora è protestante e che se b è svedese e con genitori italiani allora non è protestante. Inoltre afferma che b è svedese. T è consistente, infatti un suo modello si può ottenere assumendo che $Sv(b)$ e $Pr(b)$ siano asserzioni vere mentre $Gi(b)$ è falso. T inoltre permette di provare $Pr(b)$. Supponiamo ora di apprendere che $Gi(b)$ è vera e quindi di doverci riferire alla teoria $\underline{T} = T \cup \{Gi(b)\}$. Tale teoria è inconsistente perché in essa è possibile provare sia $\neg Pr(b)$ che $Pr(b)$. Tuttavia il senso comune non considera $\underline{T}$ vuota di contenuto ed è portato a ritenere la conclusione $\neg Pr(b)$ più plausibile della conclusione $Pr(b)$. Questo perché, in un certo senso, la regola $Sv(b) \wedge Gi(b) \Rightarrow \neg Pr(b)$ viene considerata prioritaria rispetto alla regola $Sv(b) \Rightarrow Pr(b)$. Questo tipo di argomentazione non è monotono in quanto l'aggiunta di un nuovo fatto, la formula $Gi(b)$, determina un cambiamento di opinione circa la validità di $Pr(b)$. La logica fuzzy suggerisce una via per formalizzare un tale tipo di argomentazione. Infatti

“fuzzifichiamo” la teoria T sostituendola con l’insieme fuzzy di formule v definito ponendo:

$$v(\alpha) = \begin{cases} 1 & \text{se } \alpha = Sv(b) \\ 0.8 & \text{se } \alpha = Sv(b) \wedge Gi(b) \Rightarrow \neg Pr(b), \\ 0.6 & \text{se } \alpha = Sv(b) \Rightarrow Pr(b), \\ 0 & \text{altrimenti.} \end{cases}$$

In questo modo forniamo gradi diversi di “affidabilità” alle varie formule coinvolte nel discorso. Se, per semplificare, scegliamo come norma triangolare il minimo $\wedge$, avremo che:

$$\begin{aligned} D^\wedge(v)(\alpha) &= 1 && \text{se } Sv(b) \vdash \alpha, \\ D^\wedge(v)(\alpha) &= 0.8 && \text{se } T - \{Sv(b)\} \vdash \alpha \text{ e } Sv(b) \not\vdash \alpha, \\ D^\wedge(v)(\alpha) &= 0.6 && \text{se } T \vdash \alpha \text{ e } T - \{Sv(b)\} \not\vdash \alpha \\ D^\wedge(v)(\alpha) &= 0 && \text{altrimenti.} \end{aligned}$$

In tale caso risulta che $Inc(v) = 0$ e quindi $D^\wedge(v) = D_n^\wedge(v)$. Inoltre possiamo dimostrare $Pr(b)$ con grado 0.6. Aggiungiamo ora l’informazione $Gi(b)$ con grado 1, cioè poniamo $\underline{v} = v \cup \{Gi(b)\}$. Allora è subito visto che:

$$\begin{aligned} D^\wedge(\underline{v})(\alpha) &= 1 && \text{se } Sv(b), Gi(b) \vdash \alpha, \\ D^\wedge(\underline{v})(\alpha) &= 0.8 && \text{se } T - \{Sv(b)\}, Gi(b) \vdash \alpha \text{ e } Sv(b), Gi(b) \not\vdash \alpha, \\ D^\wedge(\underline{v})(\alpha) &= 0.6 && \text{altrimenti.} \end{aligned}$$

Ne segue che $Inc(\underline{v}) = 0.6$ e quindi che,

$$\begin{aligned} D_n^\wedge(\underline{v})(\alpha) &= 1 && \text{se } Sv(b), Gi(b) \vdash \alpha, \\ D_n^\wedge(\underline{v})(\alpha) &= 0.5 && \text{se } T - \{Sv(b)\}, Gi(b) \vdash \alpha \text{ e } Sv(b), Gi(b) \not\vdash \alpha, \\ D_n^\wedge(\underline{v})(\alpha) &= 0 && \text{altrimenti.} \end{aligned}$$

In conclusione, se accettiamo che i gradi di dimostrabilità siano forniti dagli operatori di deduzione normalizzati, risulta che pur essendo $\underline{v}$ una estensione propria di v , risulta che

- v permette di dimostrare $Pr(b)$ con grado 0.6 e $\neg Pr(b)$ con grado 0,
- $\underline{v}$ permette di dimostrare $Pr(b)$ con grado 0 e $\neg Pr(b)$ con grado 0.5.

Ipotesi: Tutte le forme di ragionamento nonmonotono e per default possono essere formalizzate nella logica fuzzy fornendo pesi diversi a regole diverse ed introducendo il grado di inconsistenza nella valutazione dei gradi di dimostrabilità.

6. Metalogica fuzzy: ragionamenti con sinonimia

M. Ying in un lavoro del 1994 considera una logica in cui è permessa l'applicazione di una regola di inferenza anche quando gli antecedenti di tale regola non coincidono con asserzioni già provate. Ad esempio, supponiamo che io debba comprare un regalo per mio figlio e che fornisca al commesso di una libreria la seguente regola:

$$x \text{ è un thriller} \Rightarrow x \text{ è buono per mio figlio}$$

Allora nel caso in cui nel negozio non ci sia un libro che sia catalogato come thriller il commesso (che supponiamo si sia laureato in matematica con il massimo dei voti e lode e con una tesi in logica) deve necessariamente rispondere che non ha niente adatto alle mie esigenze. Infatti i dati non gli permettono di trovare un libro d per cui si possa provare " d è buono per mio figlio". Di fatto però se il commesso è un bravo venditore (e non laureato in matematica) tenta comunque di suggerirmi qualche libro che potrebbe andare bene. Ad esempio se un libro, chiamiamolo b , è un giallo, allora il bravo commesso proverà a proporlo anche se è cosciente che l'asserzione b è buono per mio figlio non può essere considerata completamente vera in quanto essere libro giallo ed essere thriller sono cose diverse. Se ora formalizziamo il ragionamento fatto dal commesso, ci accorgiamo che ha la seguente forma:

- | | |
|--|---|
| 1. $x \text{ è un thriller} \Rightarrow x \text{ adatto a mio figlio}$ | + |
| 2. " <u>romanzo giallo</u> " <u>è un sinonimo di</u> "thriller" | + |
| 3. $b \text{ è un romanzo giallo}$ | = |
-

$b \text{ è adatto a mio figlio.}$

In tale inferenza un ruolo particolare è giocato dall' affermazione
 "romanzo giallo" è un sinonimo di "thriller"

che esprime una relazione tra nomi di predicati ed è quindi una affermazione metalinguistica. Ora la sinonimia oltre ad essere un meta-predicato è indiscutibilmente di tipo fuzzy in quanto due predicati possono essere considerati sinonimi con gradi diversi.

Pertanto appare naturale che il grado di verità della conclusione “*b è adatto a mio figlio*” dipenda dal grado di sinonimia tra “*romanzo giallo*” e “*thriller*”.

Naturalmente è possibile che anche i predicati del linguaggio siano vaghi. Ad esempio potremmo considerare una regola del tipo

$$x \text{ è un thriller e } x \text{ economico} \Rightarrow x \text{ è adatto a mio figlio}$$

dove il predicato “economico” è fuzzy. In tale caso potremmo avere un ragionamento del tipo

$$\begin{array}{lcl} 1. & x \text{ è un thriller e } x \text{ economico} & \Rightarrow x \text{ adatto a mio figlio} + \\ 2. & \text{“romanzo giallo” è un sinonimo di “thriller”} & + \\ 3. & b \text{ è economico} & + \\ 2. & b \text{ è un romanzo giallo} & = \end{array}$$

$$b \text{ è adatto a mio figlio.}$$

La conclusione sarà ovviamente accettabile con un grado che dipende

- a) dal grado con cui si ritiene la regola corretta,
- b) dal grado di sinonimia tra “romanzo giallo” e “thriller”,
- c) dal grado con cui è lecito considerare *b* economico
- e) dal grado con cui *b* è catalogabile come romanzo giallo.

Facciamo ora qualche passo per formalizzare le forme di argomentazione di tale tipo.

Definizione 6.1. Chiamiamo $\otimes$ -relazione di sinonimia ogni relazione di $\otimes$ -similarità *Sin* definita sull’insieme dei nomi di relazioni.

Una relazione di sinonimia può essere estesa all’insieme *F* delle formule al modo seguente.

Definizione 6.2. Chiamiamo $\otimes$ -sinonimia tra formule la relazione *Sin* definita ponendo $Sin(\alpha, \beta) = 0$ nel caso in cui α e β non abbiano la stessa struttura⁶⁰ e ponendo altrimenti

⁶⁰ Diciamo che due formule α e β hanno la stessa struttura se β si può ottenere da α sostituendo alcune occorrenze di nomi di predicati (non necessariamente in modo uniforme). Ad esempio

$$\begin{aligned}
\text{Sin}(r(t_1, \dots, t_n), r'(t_1, \dots, t_n)) &= \text{Sin}(r, r') \\
\text{Sin}(\alpha \div \beta, \alpha' \div \beta') &= \text{Sin}(\alpha, \alpha') \otimes \text{Sin}(\beta, \beta') \\
\text{Sin}(\neg \alpha) &= \text{Sin}(\alpha) ; \text{Sin}(\forall x \alpha) = \text{Sin}(\alpha)
\end{aligned}$$

dove $\div$ denota un qualunque connettivo logico binario.

Detto questo modifichiamo le regole di inferenza in modo ovvio. Ad esempio il Modus Ponens assume la forma seguente

$$\frac{\alpha, \alpha' \Rightarrow \beta}{\beta} \quad \frac{\lambda, \mu}{\lambda \otimes \mu \otimes \text{Sin}(\alpha, \alpha')}$$

Non esaminiamo ulteriormente questo modo di procedere ma mettiamo in rilievo che se si vede la deduzione come ricerca di un punto fisso dell'operatore di conseguenze immediate, allora non è difficile estendere ogni logica fuzzy tramite la nozione di sinonimia. Per mostrare come ciò possa essere fatto cominciamo col definire un interessante operatore.

Definizione 6.3. Data una relazione $\text{sin} : F \times F \rightarrow [0,1]$ di $\otimes$ -sinonimia, definiamo l'operatore $\text{SIN} : U^F \rightarrow U^F$ ponendo, per ogni $s \in U^F$ ed $x \in F$,

$$\text{SIN}(s)(x) = \text{Sup}\{\text{sin}(x', x) \otimes s(x') : x' \in F\}. \quad (6.1)$$

Interpretiamo $\text{SIN}(s)$ come l'insieme fuzzy delle formule che sono sinonimo di qualche formula in s .

Proposizione 6.4. $\text{SIN} : U^F \rightarrow U^F$ è un operatore di chiusura continuo.

Tenendo conto che la composizione di due operatori continui è ancora un operatore continuo, appare naturale "comporre" l'operatore di conseguenza immediata C di una qualunque logica fuzzy con l'operatore SIN , ottenendo

$$K = C \circ \text{SIN}. \quad (6.2)$$

Data una valutazione iniziale v , $C(\text{SIN}(v))$ è l'insieme fuzzy delle formule che possono essere dedotte in un passo da formule che sono in v oppure che sono sinonimi di formule che sono in v . Ora è possibile che un tale insieme fuzzy non sia chiuso rispetto alla

la formula $\forall x(r(x,y) \wedge s(z) \Rightarrow s(x))$ ha la stessa struttura sia di $\forall x(r(x,y) \wedge t(z) \Rightarrow g(x))$ che di $\forall x(r(x,y) \wedge s(z) \Rightarrow t(x))$.

relazione di similarità. Allora dobbiamo aggiungere alle asserzioni dedotte anche le formule simili a formule in $C(SIN(v))$, cioè considerare l'insieme fuzzy $SIN(C(SIN(v)))$... Andando avanti in questo modo otteniamo un insieme fuzzy $D_{sin}(v)$ con un processo a limite

$$v \subseteq SIN(v) \subseteq D(SIN(v)) \subseteq SIN(D(SIN(v))) \subseteq \dots \rightarrow D_R(v).$$

Naturalmente, da un punto di vista strettamente matematico, possiamo dire che, essendo K un operatore continuo, viene ad essere definito un operatore D_{sin} che associa ad ogni insieme fuzzy v l'insieme fuzzy

$$D_{sin}(v) = \bigcup_{n \in \mathbb{N}} K^n(v) \quad (6.3)$$

cioè il minimo punto fisso di K maggiore o uguale a v . Chiamiamo *logica basata sulla sinonimia* una logica che sia caratterizzata da tale operatore.

7. La programmazione logica classica

Per capire meglio le potenzialità ed i limiti della logica fuzzy forse la cosa migliore è riferirsi a frammenti di tale logica, come ad esempio quelli legati alla programmazione logica fuzzy. Per fare questo ricordiamo alcune nozioni base della programmazione logica classica. Per semplificare l'esposizione mi riferisco solo a programmi positivi.

Definizione 7.1. Una *regola positiva* è una formula del tipo

$$\alpha_1 \wedge \dots \wedge \alpha_h \Rightarrow \alpha$$

con $\alpha, \alpha_1, \dots, \alpha_h$ formule atomiche. Chiamiamo *clausola positiva di programma* ogni formula che sia atomica o che sia una regola positiva. Una teoria P prende il nome di *programma positivo* se è costituita solo da clausole positive di programma.

Per rappresentare meglio la strategia di deduzione da un programma, useremo la notazione $\alpha \Leftarrow \alpha_1 \wedge \dots \wedge \alpha_h$ invece di $\alpha_1 \wedge \dots \wedge \alpha_h \Rightarrow \alpha$. Il conseguente α viene detto *testa*, l'antecedente $\alpha_1 \wedge \dots \wedge \alpha_h$ *corpo* della clausola. Chiamiamo *fatto* una formula atomica chiusa e chiamiamo *base di Herbrand* l'insieme B_P di tutti i fatti che si possono scrivere nel linguaggio di P . Nel seguito il termine “positivo” sarà sottinteso. Il nome “programma” deriva dal fatto che nel linguaggio di programmazione *Prolog* i programmi degli usuali linguaggi

procedurali sono sostituiti da sistemi di assiomi che sono clausole di programma.

Se ci limitiamo alle teorie che sono programmi positivi e se ci poniamo solo il compito di dimostrare fatti, allora è possibile utilizzare due metodi di dimostrazione particolarmente semplici. Il primo metodo, prende il nome di *metodo di risoluzione*. Il secondo metodo, che esamineremo nel seguito, si riferisce al calcolo dei punti fissi di un particolare operatore, chiamato *operatore di conseguenze immediate*. In entrambi i casi sono utili le seguenti nozioni.

Definizione 7.2. Dato un programma P , indichiamo con

- $Fatt(P)$ l'insieme dei fatti che appartengono a P o che si ottengono dalle formule atomiche di P per particolarizzazione;
- $Reg(P)$ l'insieme delle regole chiuse che appartengono a P o che si ottengono per particolarizzazione dalle regole di P
- $\underline{P} = Fatt(P) \cup Reg(P)$.

Ad esempio se P consiste nelle clausole

$r(a,b)$
 $r(b,c)$
 $r(X,X)$
 $r(X,Y) \Leftarrow r(Y,X)$

allora

$Fatt(P) = \{r(a,b), r(b,c), r(a,a), r(b,b), r(c,c)\}$
 $Reg(P) = \{r(a,a) \Leftarrow r(a,a), r(a,b) \Leftarrow r(b,a), r(a,c) \Leftarrow r(c,a),$
 $r(b,a) \Leftarrow r(a,b), r(b,b) \Leftarrow r(b,b), r(b,c) \Leftarrow r(c,b)$
 $r(c,a) \Leftarrow r(a,c), r(c,b) \Leftarrow r(b,c), r(c,c) \Leftarrow r(c,c)\}.$

E' evidente che anche se si assume che P sia finito, nel caso in cui esistono infiniti termini chiusi $\underline{P}$ risulta in generale infinito.

La seguente proposizione mostra che a tutti gli effetti posso riferirmi al programma $\underline{P}$ invece che al programma P . Questo mostra che la programmazione logica si svolge, in un certo senso, nell'ambito del calcolo proposizionale.

Proposizione 7.3. Ogni programma P è logicamente equivalente a $\underline{P}$.

Il metodo di risoluzione si basa sul processo di *unificazione* di due formule. Esso consiste nel rendere uguali due formule

atomiche tramite una opportuna sostituzione delle variabili. Ad esempio la formula $r(a,x)$ può essere unificata alla formula $r(y,b)$ ponendo al posto di x il valore b ed al posto di y il valore a . In questo caso diciamo che le due formule sono *unificabili*. Chiamiamo unificabili anche due formule che coincidono. Ovviamente può capitare che due formule non siano unificabili come avviene, ad esempio per la coppia di formule $r(a,x)$ e $r(b,y)$. In tale caso si dice che l'unificazione *fallisce*.

Supponiamo ora di volere provare il fatto α o, se si vuole, di effettuare una interrogazione del tipo “è dimostrabile α ?” . Allora possiamo procedere partendo dalla formula α da dimostrare ed “andando all'indietro”. Precisamente possiamo procedere tramite i seguenti passi:

1. verifichiamo se α è unificabile con un formula atomica in P , cioè se $\alpha \in Fatt(P)$. In caso positivo consideriamo α dimostrata, ci fermiamo e rispondiamo “sì”: Altrimenti si passa a 2
2. verifichiamo se α può essere unificata con la testa di una regola, cioè se esiste $\alpha \Leftarrow \alpha_1 \wedge \dots \wedge \alpha_n$ in $Reg(P)$ con testa α . In caso positivo tentiamo di dimostrare $\alpha_1, \dots, \alpha_n$ applicando a ciascuno di tali fatti i passi 1 e 2. Se riusciamo a dimostrare tutti i fatti $\alpha_1, \dots, \alpha_n$ allora consideriamo α dimostrata, rispondiamo “sì” e ci fermiamo. Altrimenti tentiamo con un'altra regola
3. se dopo avere esplorato tutte le possibili regole con testa unificabile ad α ci accorgiamo che tutti i tentativi di dimostrazione falliscono, allora ci fermiamo e rispondiamo “no”.

Per capire l'idea facciamo un esempio considerando il programma

$$P = \{r(a,b), r(c,d), r(x,x), s(x,y) \Leftarrow r(x,y), s(y,x) \Leftarrow r(x,y)\}.$$

e tentando di dimostrare alcuni fatti.

a) Proponiamoci di dimostrare $r(c,d)$.

La dimostrazione è semplice poiché è sufficiente constatare che tale fatto è un assioma (cioè appartiene a P).

b) Proponiamoci ora di dimostrare il fatto $r(c,c)$.

In questo caso è sufficiente accorgersi che tale formula appartiene a $Fatt(P)$ in quanto può essere unificata alla formula atomica $r(x,x)$.

c) Proponiamoci di dimostrare il fatto $s(b,a)$.

Poiché $s(b,a)$ non appartiene a $Fatt(P) = \{r(a,b), r(c,d), r(a,a), r(b,b), r(c,c), r(d,d)\}$ possiamo tentare di unificare $s(b,a)$ con la testa di una regola o, se si vuole, di vedere se $s(b,a)$ coincide con la testa di una regola in $Reg(P)$. In $Reg(P)$ le regole che potrebbero essere utili sono

$$s(b,a) \leftarrow r(b,a) \text{ e } s(b,a) \leftarrow r(a,b).$$

Nel primo caso dobbiamo concentrarci sul compito di dimostrare $r(b,a)$. Tale compito non può riuscire in quanto $r(b,a)$ non compare in $Fatt(P)$ e non compare nella testa di nessuna regola in $Reg(P)$. Allora tentiamo di utilizzare la seconda regola e quindi di provare $r(a,b)$. Poiché ci accorgiamo che $r(a,b)$ appartiene ad $Fatt(P)$ possiamo fermarci e concludere che $r(b,a)$ è dimostrabile. Se non avessimo trovato $r(a,b)$ in $Reg(P)$, poiché non esistono regole con la testa $r(a,b)$, ci saremmo fermati per concludere che $r(b,a)$ non è un teorema. Infatti tutte le possibilità di dimostrazione sono state esplorate.

d) Proponiamoci infine di provare $r(b,d)$.

E' evidente che questa è una impresa impossibile poiché tale fatto non appartiene a $Fatt(P)$ e non è unificabile con nessuna testa di regole del programma.

Il metodo di risoluzione si ottiene in un certo senso meccanizzando questo modo di procedere.⁶¹ E' importante osservare che in effetti l'apparato inferenziale utilizzato non usa nessun assioma logico e si può ridurre al solo uso di tre regole di inferenza: la regola di particolarizzazione, il Modus Ponens e la regola di $\wedge$ -introduzione che possiamo supporre nella forma più generale:

$$\frac{\alpha_1, \dots, \alpha_n}{\alpha_1 \wedge \dots \wedge \alpha_n} \quad (\text{regola della } \wedge\text{-introduzione})$$

La regola di particolarizzazione permette in un certo senso di sostituire P con $\underline{P}$. La regola di $\wedge$ -introduzione permette di

⁶¹In realtà teoricamente si presenta come un metodo di refutazione che parte dalla negazione del fatto α da provare per tentare di pervenire ad una contraddizione.

provare il corpo di una regola, il Modus Ponens di provare la testa.

8. Programmazione logica fuzzy

Le idee della programmazione logica fuzzy sono una ovvia estensione di quelle ora esposte della programmazione logica classica.

Definizione 8.1. Diciamo che una teoria fuzzy $p : F \rightarrow L$ è un *programma* (*programma positivo*) se $Supp(p)$ è un programma (programma positivo).

Dovendoci limitare ai programmi fuzzy è sufficiente considerare un apparato inferenziale che non ha assiomi logici ed utilizza la fuzzyficazione delle regole di inferenza considerate nel paragrafo precedente. Pertanto avremo il solito Modus Ponens fuzzy

$$\frac{\alpha, \alpha \Rightarrow \beta}{\beta} \quad \frac{\lambda, \mu}{\lambda \otimes \mu}$$

la regola di particolarizzazione

$$\frac{\forall x \alpha(x),}{\alpha(x/t)} \quad \frac{\lambda}{\lambda}$$

e la regola di $\wedge$ -introduzione

$$\frac{\alpha_1, \dots, \alpha_n}{\alpha_1 \wedge \dots \wedge \alpha_n} \quad \frac{\lambda_1, \dots, \lambda_n}{\lambda_1 \otimes \dots \otimes \lambda_n}$$

Da notare che tale apparato inferenziale risulta essere un frammento sia di quello della logica di Łukasiewicz⁶² che di quello di altre logiche fuzzy⁶³.

Per la programmazione logica fuzzy il modo di procedere rispetto ad una interrogazione del tipo “con quale grado è dimostrabile α ?” non è molto diverso da quello della programmazione logica classica. Ancora una volta un programma fuzzy risulta essere equivalente all’insieme fuzzy delle sue istanze ground. In questo caso però si deve tenere conto che un fatto o una regola ground possono essere istanze di formule

⁶² Infatti, come abbiamo già osservato, le prime due regole appartengono a tale logica e la terza è una regola derivata.

⁶³ Si veda l’estensione canonica della logica classica tramite una fissata norma triangolare.

diverse ed in questo caso deve essere preso il valore massimo di appartenenza al programma di tali formule.

Definizione 8.2. Dato un programma fuzzy p , definiamo l'insieme fuzzy di fatti $Fatt(p)$ ponendo

- $Fatt(p)(\alpha) = 0$ se α non è un fatto
- $Fatt(p)(\alpha)$ uguale al massimo dei valori che assume p nelle formule che per particolarizzazione permettono di ottenere α altrimenti.

Similmente definiamo l'insieme fuzzy $Reg(p)$ di regole ground ponendo

- $Reg(p)(\alpha) = 0$ se α non è una regola ground
- $Reg(p)(\alpha)$ uguale al massimo dei valori che assume p in formule che per particolarizzazione permettono di ottenere α , altrimenti.

Infine poniamo

$$\underline{p} = Fatt(p) \cup Reg(p).$$

Come nel caso classico p risulta essere logicamente equivalente a $\underline{p}$.

Proposizione 8.3. L'insieme fuzzy di fatti provabili da p coincide con l'insieme dei fatti provabili da $\underline{p}$.

Dato ora un programma p , le possibili dimostrazioni di α coincidono con quelle classiche a partire dal programma $Supp(\underline{p})$. Tuttavia, data una dimostrazione π , è necessario anche calcolare l'informazione $\lambda = I(\pi, p)$ fornita da π . Possiamo anche interpretare λ come il “prezzo da pagare” ogni volta che si utilizza una formula in $Supp(p)$. Dobbiamo allora registrare in un “contatore” tale prezzo ed aggiornarlo ad ogni passo della dimostrazione. Per farlo, nello sviluppare una dimostrazione π possiamo partire ponendo inizialmente $\lambda = 1$ e poi aggiornare tale valore man mano che si va avanti. Precisamente se al passo n di una dimostrazione abbiamo in memoria il valore λ , allora al passo successivo dobbiamo moltiplicare tale valore per:

- il valore di appartenenza a p della regola utilizzata
- il grado con cui si sono dimostrate le formule atomiche presenti nel corpo della regola utilizzata.

Come al solito, dopo avere esplorato tutte le possibili dimostrazioni di α è necessario “fondere” i vari gradi ottenuti.

Per fare un esempio consideriamo il programma fuzzy costituito dai fatti

<i>ama(carlo, elena)</i>	[0.9]
<i>ama(carlo, giovanna)</i>	[0.8]
<i>ama(maria, maria)</i>	[1]
<i>ama(x, x)</i>	[0.9]
<i>donna(maria)</i>	[1]
<i>donna(luisa)</i>	[1]
<i>giovane(maria)</i>	[0.8]
<i>alta(maria)</i>	[0.7]
<i>giovane(elena)</i>	[0.6]

e dalle tre regole

<i>ama(carlo, y) ← donna(y) ∧ bella(y)</i>	[0.9]
<i>bella(x) ← giovane(x)</i>	[0.2]
<i>bella(x) ← alta(x) ∧ giovane(x)</i>	[0.9]

ed assumiamo come norma triangolare l'usuale prodotto. Volendo calcolare p dovremmo scrivere un po' troppe formule e quindi per brevità ci riferiremo ancora a p .

Se si pone l'interrogazione

ama(maria, maria) ?

allora tale formula è ottenibile per particolarizzazione da *ama(x, x)* con grado 0.9. Tuttavia essendo presente anche come fatto con grado 1, essa è provabile con grado $\max\{0.9, 1\} = 1$.

Se si pone l'interrogazione

ama(carlo, maria) ?

ci accorgiamo (dopo avere posto $\lambda = 1$) che una eventuale dimostrazione di tale formula deve usare la prima regola *ama(carlo, y) ← donna(y) ∧ bella(y)* (quindi poniamo $\lambda = 1 \times 0.9$) in quanto tale fatto non appartiene a $Fatt(Supp(p))$. La sostituzione di y con "maria" permette l'unificazione tra *ama(carlo, maria)* e *ama(carlo, y)*. Ciò conduce all'esigenza di dimostrare le formule *donna(maria)* e *bella(maria)*. La formula *donna(maria)* può essere provata con grado 1 (poniamo $\lambda = 1 \times 0.9 \times 1$). La formula *bella(maria)* potrebbe essere provata tramite la seconda regola ($\lambda = 1 \times 0.9 \times 1 \times 0.2$) che a sua volta comporta il provare *giovane(maria)* ($\lambda = 1 \times 0.9 \times 1 \times 0.2 \times 0.8 = 0.144$). Tuttavia *bella(maria)* potrebbe essere provata anche tramite la terza regola

($\lambda = 1 \times 0.9 \times 1 \times 0.9$) che a sua volta comporta il provare *alta(maria)* ($\lambda = 1 \times 0.9 \times 1 \times 0.9 \times 0.7$) e poi *giovane(maria)* ($\lambda = 1 \times 0.9 \times 1 \times 0.9 \times 0.7 \times 0.8 = 0.4536$). Essendo $0.4536 > 0.144$ possiamo concludere che *ama(carlo, maria)* è un teorema con grado 0.4536.

Se si pone invece l'interrogazione

ama(carlo, luisa) ?

allora applicando la prima regola ($\lambda = 1 \times 0.9$) si pone il problema di provare *donna(luisa)* ($\lambda = 1 \times 0.9 \times 1$) e *bella(luisa)*. E' evidente che questo ultimo fatto non è provabile e quindi possiamo concludere che *ama(carlo, luisa)* è provabile con grado 0 (o, se si vuole, non è provabile).

9. Usare la tecnica dei punti fissi

Abbiamo già visto come, in presenza di un apparato inferenziale, sia utile l'utilizzazione della tecnica di calcolo dei punti fissi di un operatore. Nel caso della programmazione logica questa tecnica si dimostra ancora più utile e si basa sull'operatore di conseguenze immediate definito al modo seguente.

Definizione 9.1. Dato un programma positivo P , definiamo l'operatore $T_P : P(B_P) \rightarrow P(B_P)$ ponendo, per ogni sottoinsieme X di B_P

$$T_P(X) = X \cup Fatt(P) \cup \{ \alpha \mid \alpha \Leftarrow \alpha_1 \wedge \dots \wedge \alpha_n \in Reg(P), \alpha_1 \in X, \dots, \alpha_n \in X \}. \quad (9.1)$$

Il significato di tale operatore è che se ad un certo passo del processo inferenziale è stato dimostrato un insieme X di fatti al passo successivo è possibile dimostrare l'insieme $T_P(X)$ di fatti.

Teorema 9.2. L'operatore T_P è compatto e conservativo. Detto M_P l'insieme dei fatti deducibili da un programma P , risulta che M_P è il minimo punto fisso di T_P e quindi che

$$M_P = \bigcup_{n \in \mathbb{N}} T_P^n(\emptyset). \quad (9.2)$$

In altre parole, per ogni $\alpha \in B_P$,

$$P \vdash \alpha \Leftrightarrow \exists n \alpha \in T_P^n(\emptyset).$$

Non esponiamo la dimostrazione di tale teorema ma facciamo un paio di esempi per illustrarne il significato.

Esempio 1. Sia $P = \{r(a,b), r(c,d), s(x,y) \Leftarrow r(x,y), s(x,y) \Leftarrow r(y,x)\}$, allora $U_P = \{a,b,c,d\}$. Inoltre,

$$T(\emptyset) = Fatt(P) = \{r(a,b), r(c,d)\};$$

$$T^2(\emptyset) = \{r(a,b), r(c,d)\} \cup \{s(a,b), s(c,d), s(b,a), s(d,c)\};$$

$$T^3(\emptyset) = T^2(\emptyset).$$

Pertanto l'insieme $T^2(\emptyset) = \{r(a,b), r(c,d), s(a,b), s(c,d), s(b,a), s(d,c)\}$ è il minimo punto fisso di T_P e quindi coincide con l'insieme M_P dei fatti che si possono derivare dal programma P . In tale esempio si raggiunge il punto fisso in soli due passi.

Esempio 2. Consideriamo un linguaggio con una costante c , un simbolo s per una funzione unaria ed un nome di predicato monadico “*dispari*”. Consideriamo il programma positivo

$$dispari(c).$$

$$dispari(s(s(X))) \Leftarrow dispari(X).$$

Allora

$$T(\emptyset) = \{dispari(c)\}$$

$$T^2(\emptyset) = \{dispari(c)\} \cup \{dispari(s(s(c)))\}$$

$$T^3(\emptyset) = \{dispari(c), dispari(s(s(c)))\} \cup \{dispari(s(s(s(s(c))))\}$$

...

In tale esempio tutti gli insiemi $T^n(\emptyset)$ sono diversi tra loro e l'insieme M_P dei fatti dimostrabili coincide con l'unione infinita $\bigcup_{n \in \mathbb{N}} T^n(\emptyset)$ che coincide con l'insieme dei fatti del tipo $dispari(s(...s(a)))$ dove s compare un numero pari di volte.

Possiamo estendere un tale modo di procedere alla programmazione logica fuzzy (si veda [Vojtas 2001] e [Gerla 2005]). A tale scopo dobbiamo definire, con una qualche precauzione, l'operatore di conseguenze immediate.

Definizione 9.3. Sia $p : F \rightarrow L$ un programma fuzzy, allora chiamiamo *operatore di conseguenze immediate* associato a p l'operatore $T_p : U^B \rightarrow U^B$ definito ponendo, per ogni sottoinsieme fuzzy s di B_p ed $\alpha \in B_p$,

$$T_p(s)(\alpha) = s(\alpha) \vee p(\alpha) \vee \underline{T}_p(s)(\alpha)$$

dove

$$\underline{T}_p(s)(\alpha) = Sup\{p(\alpha \Leftarrow \alpha_1 \wedge \dots \wedge \alpha_n) \otimes s(\alpha_1) \otimes \dots \otimes s(\alpha_n) : \alpha_1, \dots, \alpha_n \in B_p\}$$

Ovviamente nel calcolo di $T_p(s)(\alpha)$ possiamo limitarci alle regole che appartengono al supporto di p . Come nel caso classico l'operatore di conseguenze immediate fornisce un elegante e potente modo di calcolare l'insieme fuzzy m_p dei fatti derivabili da p .

Teorema 9.4. L'operatore T_p è compatto e conservativo. Inoltre l'insieme fuzzy m_p dei fatti che sono conseguenze di p coincide con il minimo punto fisso di T_p e quindi

$$m_p = \bigcup_{n \in \mathbb{N}} T_p^n(\emptyset).$$

Se ci riferiamo all'esempio di programma fuzzy proposto del paragrafo precedente otteniamo che $T_p(\emptyset)$ coincide con l'insieme fuzzy di fatti

<i>ama(carlo, elena)</i>	[0.9]
<i>ama(carlo, giovanna)</i>	[0.8]
<i>ama(maria, maria)</i>	[Max{0.9, 1}=1]
<i>ama(carlo, carlo)</i>	[0.9]
<i>ama(elena, elena)</i>	[0.9]
<i>ama(giovanna, giovanna)</i>	[0.9]
<i>ama(luisa, luisa)</i>	[0.9]
<i>donna(maria)</i>	[1]
<i>donna(luisa)</i>	[1]
<i>giovane(maria)</i>	[0.8]
<i>alta(maria)</i>	[0.7]
<i>giovane(elena)</i>	[0.6]

$T_p^2(\emptyset)$ si ottiene unendo $T_p(\emptyset)$ con l'insieme fuzzy di fatti

<i>bella(maria)</i>	[Max{0.8⊗0.2, 0.9⊗0.7⊗0.8}=0.504]
<i>bella(elena)</i>	[0.2⊗0.6 = 0.12]

$T_p^3(\emptyset)$ si ottiene unendo $T_p^2(\emptyset)$ con il fatto fuzzy

<i>ama(carlo, maria)</i>	[0.9⊗1⊗0.504 = 0.4536].
--------------------------	-------------------------

E' evidente che se si va avanti non si ottengono ulteriori ampliamenti di $T_p^3(\emptyset)$ e che quindi $T_p^3(\emptyset)$ è il punto fisso cercato.

10. Sinonimia e corruzione in programmazione logica

Non è molto difficile inventarsi una programmazione logica basata sulla sinonimia. Per quello che riguarda il metodo del punto fisso l'unica variazione è intrecciare l'applicazione dell'operatore di conseguenza logica con l'operatore di similarità

SIN definito nel paragrafo 5. Successivamente si deve procedere al calcolo del minimo punto fisso per l'operatore composto $T_p \circ SIN$.

Più interessante è definire un metodo di risoluzione che si basi su di una versione "rilassata" della unificazione. Supponiamo che nel programma che abbiamo già considerato al posto del fatto

$alta(maria)$ [0.7]

ci fosse il fatto

$altina(maria)$ [0.85]

Supponiamo inoltre di fare la domanda $cotto(carlo, maria)$?. Allora se si accettano le procedure della programmazione logica non esiste nessun modo di procedere alla dimostrazione del fatto $cotto(carlo, maria)$ non essendo esso unificabile con la testa di qualche regola. Tuttavia corrompendo un immaginario "guardiano dell'unificazione" e pagando un opportuno prezzo potremmo comunque fare scattare la prima regola. Il prezzo da pagare è legato al grado di sinonimia tra "ama" e "cotto". Supponiamo che tale grado sia 0.8 e mettiamo in memoria questo numero che deve essere composto con il grado di accettabilità della regola ottenendo quindi $\lambda = 0.9 \cdot 0.8$. Ora possiamo tentare di procedere come abbiamo fatto prima. Purtroppo non essendoci più un fatto del tipo $alta(maria)$ possiamo provare $bella(maria)$ solo con la seconda regola ottenendo $0.8 \cdot 0.2 = 0.16$ e quindi $ama(carlo, maria)$ solo con grado $0.9 \cdot 0.8 \cdot 0.16 = 0.1152$. Tuttavia anche in questo caso possiamo ottenere di meglio. Se infatti accettiamo che "altina" sia un sinonimo di "alta" con grado 0.86, allora possiamo fare scattare lo stesso l'unificazione e quindi utilizzare la terza regola. Il prezzo da pagare al "guardiano dell'unificazione" per provare $bella(maria)$ in questo caso è $0.9 \cdot 0.86 \cdot 0.85 \cdot 0.8$ che composto con il prezzo di $0.9 \cdot 0.8$ per l'uso della prima regola mi restituisce 0.3789504. La cosa si presenta conveniente essendo 0.3789504 maggiore di 0.1152. Pertanto si può concludere che $ama(carlo, maria)$ è provato con grado 0.3789504.

11. Modelli di Herbrand

Nella programmazione logica classica, dato un programma P , l'insieme M_P dei fatti deducibili da P viene detto anche *modello minimo di Herbrand di P* . Il motivo è che a tale insieme è possibile associare un modello di P che risulta in un certo senso il più piccolo possibile. L'idea è che sia possibile costruire modelli

di una teoria a partire dal linguaggio stesso in cui è espressa la teoria. Più precisamente si assume che:

si possa considerare come dominio di interpretazione l'insieme dei "nomi che si possono formulare nel linguaggio".

Per fare questo è utile la seguente definizione.

Definizione 11.1. Dato un linguaggio L che contenga almeno una costante, chiamiamo *universo di Herbrand* l'insieme U_L dei termini chiusi di L .

Se non esistono nomi di operazioni allora U_L coincide con l'insieme delle costanti e quindi è in generale finito. Se invece esistono nomi di funzioni, allora U_L è infinito. Ad esempio se nel linguaggio ho una costante 1 ed un nome di operazione +, allora l'universo di Herbrand sarà costituito dai termini chiusi

1, 1+1, (1+1)+1, 1+(1+1), . . .

Se nel linguaggio ho anche il simbolo 0, allora l'universo di Herbrand sarà costituito dai termini del tipo

0, 1, 1+1, 0+1, 1+0, (1+1)+1, 1+(1+1), . . .

Se il linguaggio contiene solo i nomi "Maria" e "Carlo" allora l'universo di Herbrand si riduce all'insieme {Maria, Carlo}. Se oltre a questi due nomi si introduce anche il nome di funzione *padre_di*, allora nell'universo di Herbrand sarà costituito dagli infiniti termini

Maria, Carlo,

padre_di(Maria), *padre_di*(Carlo),

padre_di(*padre_di*(Maria)), *padre_di*(*padre*(Carlo)),

. . .

Definizione 11.2. Chiamiamo *interpretazione di Herbrand* una interpretazione (U_L, I) con dominio U_L tale che

- per ogni costante c si pone $I_L(c) = c$

- per ogni nome di funzione n -aria h , $I_L(h)$ è definita ponendo, per ogni $t_1, \dots, t_n$ in U_L ,

$$I_S(h)(t_1, \dots, t_n) = h(t_1, \dots, t_n)$$

Non è invece imposta nessuna condizione sul modo di interpretare i nomi di relazione e le interpretazioni di Herbrand differiscono tra loro proprio per il modo come questo viene fatto. La seguente proposizione mostra che è possibile identificare le interpretazioni di Herbrand con i sottoinsiemi di fatti in L .

Proposizione 11.3. Le interpretazioni di Herbrand sono in corrispondenza uno-ad-uno con i sottoinsiemi di B_L .

Dim. Sia S un sottoinsieme di B_L , allora possiamo definire una interpretazione di Herbrand (U_L, I_S) ponendo, per ogni nome di relazione n -aria r ,

$$I_L(r) = \{(t_1, \dots, t_n) : r(t_1, \dots, t_n) \in S\}.$$

E' chiaro che la corrispondenza che associa ad ogni S l'interpretazione (U_L, I_S) è iniettiva. Per mostrare che è suriettiva, consideriamo una qualunque interpretazione di Herbrand (U_L, I) , allora è semplice verificare che posto $S_I = \{\alpha \in B_L : (U_L, I) \models \alpha\}$ risulta che (U_L, I_S) coincide con (U_L, I) . $\square$

Tale proposizione permette di identificare la classe delle interpretazioni di Herbrand con $P(B_L)$ e quindi anche di vedere tale classe come un insieme ordinato (per meglio dire una algebra di Boole). Stante la definizione, possiamo infatti dire che una interpretazione di Herbrand I è *contenuta* in una interpretazione di Herbrand I' se $S_I \subseteq S_{I'}$ cioè se ogni fatto che risulta vero in I risulta vero anche in I' .

Esempio. Sia L il linguaggio le cui costanti sono a, b, c ed avente un predicato binario r ed un predicato monadico s . Risulta allora che $U_L = \{a, b, c\}$. Se si considera ad esempio l'insieme $S = \{r(a,a), r(c,b), r(b,c), s(a), s(c)\}$ di fatti allora viene definita l'interpretazione I_S per cui $I_S(r) = \{(a,a), (c,b), (b,c)\}$ e $I(s) = \{a, c\}$. Poiché B_L contiene $3^2 + 3 = 12$ elementi esistono $2^{12} = 4096$ possibili sottoinsiemi di B_L e quindi 4096 possibili interpretazioni di Herbrand di tale linguaggio.

Se partiamo da un programma P allora scriveremo U_P e B_P per indicare l'universo di Herbrand e la base di Herbrand relativi al linguaggio in cui è scritto P . E' di particolare importanza l'interpretazione di Herbrand definita al modo seguente.

Definizione. 11.4. Dato un programma P chiamiamo *modello minimo di Herbrand* di P l'interpretazione di Herbrand (U_P, I_P) associata all'insieme $M_P = \{\alpha \in B_P : P \vdash \alpha\}$ di fatti che sono conseguenza logica di P .

Allora nel modello minimo di P ogni predicato r viene interpretato ponendo

$$I_P(r) = \{(t_1, \dots, t_n) : P \vdash r(t_1, \dots, t_n)\}.$$

Questo significa che, almeno per i fatti, in tale modello la verità coincide con la dimostrabilità. L'espressione "*modello minimo di Herbrand*" è giustificata dalla seguente proposizione di cui omettiamo la dimostrazione.

Proposizione 11.5. Dato un programma positivo P il minimo modello di Herbrand di P è un modello di Herbrand di P che è contenuto in tutti i possibili modelli di Herbrand di P .

Tale modo di procedere può essere esteso alla programmazione logica fuzzy al modo seguente. Omettiamo le dimostrazioni che possono essere trovate in [Vojtas 2001] e [Gerla 2005]).

Definizione 11.6. Chiamiamo *interpretazione di Herbrand fuzzy* una interpretazione (U_L, I) con dominio U_L tale che

- per ogni costante c si pone $I_L(c) = c$
- per ogni nome di funzione n -aria h , $I_L(h)$ è definita ponendo, per ogni $t_1, \dots, t_n$ in U_L ,
 $I_S(h)(t_1, \dots, t_n) = h(t_1, \dots, t_n).$

Anche in questo caso le varie interpretazioni di Herbrand differiscono solo per il modo di interpretare i predicati. Vale inoltre la seguente proposizione.

Proposizione 11.7. Le interpretazioni di Herbrand fuzzy sono in corrispondenza uno-ad-uno con i sottoinsiemi fuzzy di B_L .

Dim. Sia s un insieme fuzzy di fatti, allora possiamo definire una interpretazione di Herbrand fuzzy (U_L, I_s) ponendo per ogni nome di predicato r , $I_s(r)(t_1, \dots, t_n) = s(r(t_1, \dots, t_n))$. Tale corrispondenza è iniettiva. Per mostrare che è suriettiva consideriamo una qualunque interpretazione di Herbrand fuzzy (U_P, I) e sia s_I l'insieme fuzzy dei fatti veri in tale interpretazione, cioè poniamo $s_I(r(t_1, \dots, t_n)) = I(r)(t_1, \dots, t_n)$. Allora una semplice verifica mostra che l'interpretazione associata a s_I coincide con I . $\square$

Pertanto possiamo identificare le interpretazioni di Herbrand fuzzy con la classe dei sottoinsiemi fuzzy di B_p e quindi definire anche un ordinamento tra tali interpretazioni. Precisamente diciamo che una interpretazione I è *contenuta* in una interpretazione di Herbrand I' se risulta che s_I è contenuto in $s_{I'}$. E' di particolare importanza l'interpretazione di Herbrand definita al modo seguente.

Definizione 11.8. Dato un programma fuzzy p chiamiamo *modello minimo di Herbrand* di p l'interpretazione di Herbrand (U_p, I_p) associata all'insieme fuzzy m_p di fatti che sono conseguenza logica di p .

L'espressione "*modello minimo di Herbrand*" è giustificata dalla seguente proposizione di cui omettiamo la dimostrazione.

Proposizione 11.9. Dato un programma positivo P il minimo modello di Herbrand di P è un modello di Herbrand di P che è contenuto in tutti i possibili modelli di Herbrand di P .

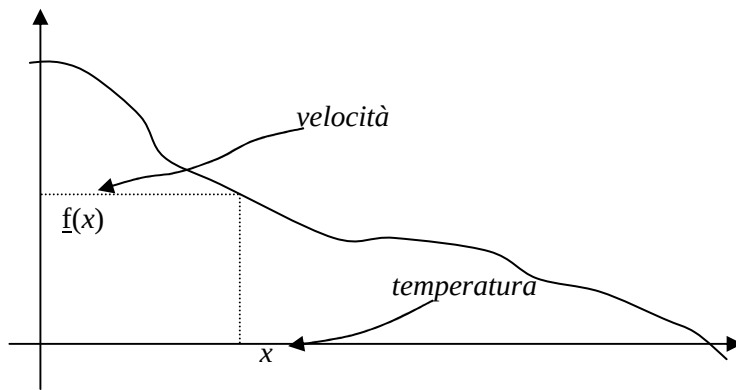
12. Il controllo: usare l'esperienza di un "praticone" ?

I maggiori successi della teoria degli insiemi fuzzy sono legati alla tecnica del controllo fuzzy. Questo successo non è casuale. Infatti, a mio parere, dietro alle tecniche "ingegneristiche" del controllo fuzzy si nasconde la vera essenza e l'importanza filosofica della logica fuzzy.

Ricordiamo che la teoria usuale del controllo ha come scopo quello di "progettare" apparecchi del tipo



capaci, dato un input x , percepito tramite un sensore dall'ambiente, di emettere un output y che determina un'azione sull'ambiente. In altri termini, un tale apparecchio calcola una funzione $\underline{f} : X \rightarrow Y$ tale che $\underline{f}(x)$ sia uguale al controllo da fare, dato x . Ad esempio, X potrebbe essere l'insieme delle possibili temperature, Y l'insieme delle possibili velocità di un ventilatore:



Ora, usualmente, tale funzione è ottenuta a partire da:

1. una teoria generale su come è fatto il mondo (usualmente fornita dalla fisica), teoria che si esprime tramite una serie di equazioni differenziali

+

2. metodi matematici per risolvere le risultanti equazioni ed ottenere pertanto la funzione $\underline{f}$.

+

3. costruzione di un marchingegno capace di realizzare la funzione ottenuta.

In altre parole si applica lo schema

fisica $\rightarrow$ teorie generali $\rightarrow$ equazioni

$\rightarrow$ matematica $\rightarrow$ ingegneria.

Si pone tuttavia il problema se questo sia l'unico modo di procedere che si ricollega ad un problema ben più ampio ed interessante: è sempre necessario ed utile accettare le equazioni

Razionalità = Scienza moderna = fisica + analisi matematica ?

Il punto di vista della teoria degli insiemi fuzzy è che sia possibile ottenere un buon controllo utilizzando l'informazione di tipo qualitativo che un "esperto-praticone" può fornire relativamente al controllo da fare. Per esperto-praticone si deve intendere una persona con nessuna conoscenza teorica dei fenomeni considerati ma con una grande esperienza su tali

fenomeni e grande abilità a trattarli. Tale esperto comunicherebbe la sua competenza ad un apprendista nel linguaggio naturale con espressioni del tipo

SE x è *Basso* ALLORA metti y *Lento*,
SE x è *Alto* ALLORA metti y *Veloce*,
...
SE x è *Medio* ALLORA ...

Tre osservazioni in proposito di tale sistema di regole sono fondamentali:

1. Tale sistema è capace di trasmettere informazione
(un essere umano è capace di capire tali suggerimenti e, in qualche modo, di applicarli)
2. L'informazione trasmessa non è completa (persone diverse interpretano i suggerimenti dell'esperto in modo diverso)
3. La mancanza di completezza dell'informazione può essere colmata solo da una (a volte lunga) esperienza.

Su queste affermazioni non credo ci possano essere dubbi, tuttavia da esse nascono le seguenti domande:

- Quale è la natura di un tale tipo di informazione ?
- E' possibile fornire ad una macchina tale tipo di informazione in modo che la macchina sia capace di eseguire le relative regole come farebbe un essere umano?
- Come comunichiamo tra noi ?

Tali questioni a me sembrano fondamentali sia da un punto di vista filosofico che da un punto di vista tecnico. La speranza è che un vecchio cuoco possa insegnare ad una macchina ad eseguire le sue ricette o che una macchina possa diventare un buon medico dopo avere letto manuali di medicina e dopo avere fatto un po' di esperienza (ammazzato un po' di pazienti). Il controllo fuzzy può essere visto come un tentativo di affrontare tali questioni.

13. La procedura proposta dal controllo fuzzy

Il sistema di regole che abbiamo scritto nel paragrafo precedente non è molto diverso dal procedimento di descrivere una funzione

$\underline{f}$ in modo approssimato tramite una tabella

x	y
x_1	y_1
$\dots$	$\dots$
x_n	y_n

dove y_i è l'immagine di x_i tramite $\underline{f}$ mentre i rimanenti valori di $\underline{f}$ vengono ottenuti con metodi di interpolazione.⁶⁴ Una descrizione linguistica di tale tabella è infatti del tipo

SE x è (approssimativamente uguale a) x_1
ALLORA metti y (approssimativamente uguale a) y_1

SE x è (approssimativamente uguale a) x_2
ALLORA metti y (approssimativamente uguale a) y_2
...

Parlando in termini geometrici, dopo avere disegnato i punti $(x_1, y_1), \dots, (x_n, y_n)$ per approssimare una data funzione $\underline{f}$ il metodo di interpolazione mi permette di tracciare una “buona” curva che passa vicino ai punti $(x_1, y_1), \dots, (x_n, y_n)$.

La procedura del controllo fuzzy non è molto diversa. Infatti, una volta che un esperto abbia fornito un sistema di regole del tipo:

SE x è *Piccolo* ALLORA metti y *Lento*,
SE x è *Medio* ALLORA metti y *Veloce*,
...

possiamo rappresentare tale descrizione tramite una tabella del seguente tipo:

x	y
<i>Piccolo</i>	<i>Lento</i>
<i>Small</i>	<i>Veloce</i>
<i>Medio</i>	<i>Moderato</i>
<i>Grande</i>	<i>Velocissimo</i>
<i>Verybig</i>	<i>Moderato</i>

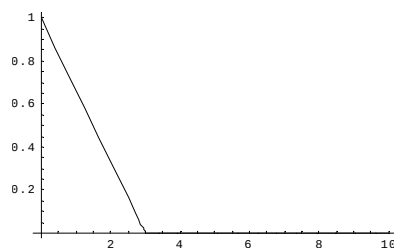
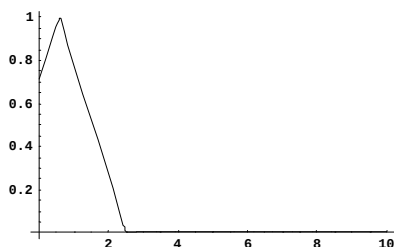
⁶⁴ Tipico è l'esempio delle tavole dei logaritmi o delle tavole trigonometriche che si usavano fino a pochi anni fa.

dove le parole “Piccolo”, “Small”, “Lento” ... sono interpretate come “numeri fuzzy” tramite opportuni sottoinsiemi fuzzy $\text{piccolo} : X \rightarrow U$, $\text{small} : X \rightarrow U$, $\text{lento} : Y \rightarrow U$, A tale tabella dovremo fare corrispondere una unione di corrispondenti *punti fuzzy* $(\text{piccolo}, \text{lento})$, $(\text{small}, \text{veloce})$... Per fare questo interpretiamo un punto fuzzy come $(\text{piccolo}, \text{lento})$ tramite il prodotto cartesiano $\text{piccolo} \times \text{lento} : X \times Y \rightarrow U$ definito ponendo

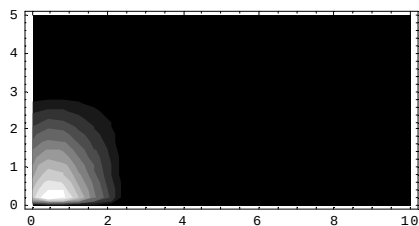
$$(\text{small} \times \text{veloce})(x, y) = \text{small}(x) \otimes \text{veloce}(y).$$

Il coinvolgimento della norma triangolare $\otimes$ nel definire il prodotto cartesiano si giustifica con il fatto che il grado di appartenenza di (x, y) a $\text{small} \times \text{veloce}$ coincide con la valutazione della congiunzione dell’affermazione “ x appartiene a *small*” e l’affermazione “ y appartiene a *veloce*”.

Per fare un esempio, supponiamo che la norma sia il prodotto usuale e che i sottoinsiemi fuzzy *piccolo*, *lento*, ... siano rappresentati da funzioni-triangolo, come in figura:



Allora, se si adotta come norma triangolare il prodotto usuale, il punto fuzzy $\text{piccolo} \times \text{lento}$ sarà rappresentato dalla figura chiara seguente dove per ogni punto (x, y) il numero $\text{piccolo} \times \text{lento}(x, y)$ rappresenta l’intensità di bianco in (x, y) .



Il secondo passo consiste nell'associare al sistema di regole, o se si vuole alla tabella una “funzione fuzzy” che si definisce come estensione della nozione di funzione non

deterministica, cioè come estensione della nozione di relazione.

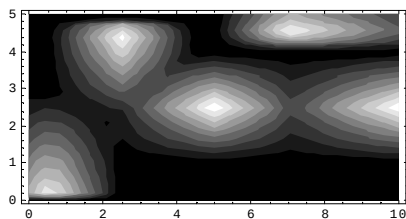
Definizione 13.1. Dati due insiemi non vuoti X ed Y chiamiamo *funzione fuzzy* ogni relazione fuzzy di X in Y cioè ogni sottoinsieme fuzzy di $X \times Y$.

Se $\tilde{f} : X \times Y \rightarrow U$ è una funzione fuzzy ed $\underline{x} \in X$ allora l'insieme fuzzy $s : Y \rightarrow U$ definito ponendo $s(y) = f(\underline{x}, y)$ può essere considerato come l'insieme fuzzy dei possibili output. In altre parole $\tilde{f}(\underline{x}, y)$ è il grado con cui si può considerare y un output corrispondente all'input $\underline{x}$.

Detto questo, al sistema di cinque regole associamo la funzione fuzzy non determinista $f : X \times Y \rightarrow U$ ottenuta come unione di cinque punti fuzzy, cioè la funzione fuzzy

$$\tilde{f} = (\text{piccolo} \times \text{lento}) \cup (\text{small} \times \text{veloce}) \cup (\text{medio} \times \text{moderato}) \cup (\text{grande} \times \text{velocissimo}) \cup (\text{verybig} \times \text{moderato}).$$

La rappresentazione di tale funzione sarà allora la seguente:

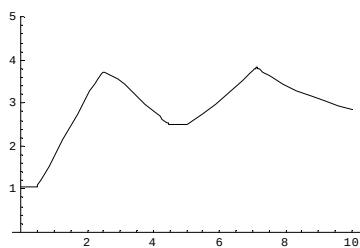


Dobbiamo ora associare ad $\tilde{f}$ (funzione fuzzy non determinista) una funzione f (classica e determinista). I procedimenti adottati a tale scopo prendono il nome di *processi di*

defuzzificazione. Ad esempio può essere adottato il *metodo del baricentro* che consiste nel definire f tramite l'equazione

$$f(r) = \frac{\int_Y f(r, y) \cdot y \, dy}{\int_Y f(r, y) \, dy}$$

Il grafico corrispondente a tale funzione o, se si vuole, la curva interpolante, sarà del tipo:



Infine il passo più importante consiste nel calibrare tutto tramite un processo di “addestramento”.⁶⁵ A tale scopo si confronta la funzione ottenuta f con la funzione che si voleva ottenere. Se il risultato non è soddisfacente

si cambia l'interpretazione delle parole vaghe *Small*, *Veloce*, ... cioè si cambiano i sottoinsiemi fuzzy con qualche tipo di strategia. Questo ultimo passo è necessario poiché l'informazione trasmessa dall'esperto non è completa e può essere completata solo tramite un processo di apprendimento.

Riassumiamo la procedura proposta dal controllo fuzzy.

STEP 1. Le parole

“Piccolo”, “Small”, “Lento” ...

sono interpretate tramite “quantità fuzzy” in maniera ragionevole.

STEP 2. Si associa ad ogni regola un punto fuzzy ed al sistema di regole la funzione fuzzy $\tilde{f} : X \times Y \rightarrow U$ ottenuta come unione dei punti fuzzy corrispondenti.

STEP 3. Tramite un processo di *defuzzificazione* si associa ad f (funzione fuzzy non determinista) una funzione determinista $f : X \rightarrow Y$.

STEP 4. Se f è una approssimazione adeguata di f ci si ferma; altrimenti si cambia l'interpretazione delle parole vaghe *Small*, *Veloce* ... e si va a STEP 2.

14. Logica e modelli di Herbrand per il controllo fuzzy

A dispetto del grande successo del controllo fuzzy, non esiste una convincente giustificazione logica delle tecniche applicate. Inoltre si sente l'esigenza di potere considerare descrizioni linguistiche più complesse di quelle esprimibili tramite una

⁶⁵ Questo ruolo dell'addestramento-apprendimento è da tempo anche la caratteristica fondamentale delle reti neurali ed ha un significato filosofico fondamentale.

semplice tabella. Vediamo allora se si possa inquadrare quanto proposto dal controllo fuzzy nell'ambito della logica matematica che sembra essere lo strumento più adeguato per la rappresentazione e la trattazione delle conoscenze di un esperto. Naturalmente la prima tentazione è quella di interpretare la struttura del tipo SE-ALLORA di una regola come una implicazione logica. Ad esempio potremmo interpretare una regola come "SE x è *Piccolo* ALLORA y è *Lento*" tramite una formula della logica del primo ordine del tipo $Piccolo(x) \Rightarrow Lento(y)$. Sfortunatamente ciò non può essere fatto in un modo così diretto. Infatti, supponiamo che r sia un input tale che $Piccolo(r)$ sia falso, allora la formula $Piccolo(r) \Rightarrow Lento(t)$ sarebbe vera per ogni t in Y e quindi ogni $t \in Y$ fornirebbe un controllo corretto. Ciò non è sicuramente nelle intenzioni dell'esperto che ha fornito le regole. La cosa sarebbe meno grave se i predicati che si applicano agli input

Piccolo, Small, Medio, Grande, Verybig

costituissero una sorta di partizione dell'insieme di tali input ed infatti nel controllo fuzzy si tende a fare in modo che questo avvenga. Tuttavia nel comportamento di un esperto esiste sicuramente la possibilità che per particolari input non si sappia quale sia la regola da applicare.

Partendo da tali considerazioni in [Gerla 2001]*a* e [Gerla 2005] ho proposto un approccio al controllo fuzzy che permette di vederlo come un capitolo della programmazione logica fuzzy. L'idea è che dato un input r ed un possibile output t , il numero $\tilde{f}(r,t)$ rappresenti il grado di verità dell'asserzione " t è una buona risposta dato r ". Ciò significa che la $\tilde{f}$ è ottenuta come interpretazione della relazione vaga *Buono* in una logica a più valori. Per illustrare una tale idea, supponiamo di avere un linguaggio del primo ordine Λ che contenga un simbolo di relazione binaria *Buono*. Associamo poi al sistema di regole proposte nel paragrafo precedente il programma fuzzy $p : F \rightarrow U$ definito ponendo:

$Buono(x,y) \Leftarrow Piccolo(x) \wedge Lento(y)$	$[\lambda_1]$
$Buono(x,y) \Leftarrow Small(x) \wedge Veloce(y)$	$[\lambda_2]$
$Buono(x,y) \Leftarrow Medio(x) \wedge Moderato(y)$	$[\lambda_3]$
$Buono(x,y) \Leftarrow Grande(x) \wedge Velocissimo(y)$	$[\lambda_4]$
$Buono(x,y) \Leftarrow Verybig(x) \wedge Moderato(y)$	$[\lambda_5]$

$Piccolo(\underline{r})$	$[piccolo(r)]$
$Small(\underline{t})$	$[small(t)]$
...	

dove indichiamo con $\underline{r}$ e $\underline{t}$ delle costanti che denotano r e t , al variare di $r \in X$ e $t \in Y$, rispettivamente. Osserviamo che in questo caso le parole *Piccolo*, *Lento*, *Small*,... denotano predicati e non quantità fuzzy come avviene nell'approccio usuale al controllo fuzzy.

Fissiamo una norma triangolare, calcoliamo l'insieme dei fatti deducibili da tale programma, cioè il suo minimo modello di Herbrand ed in particolare consideriamo l'estensione *buono* : $X \times Y \rightarrow U$ del predicato *Buono* in tale modello. E' immediato verificare che se $\tilde{f}$ è la funzione fuzzy ottenuta tramite la procedura esposta nel paragrafo precedente e tutti i λ_i sono uguali ad 1, allora $\tilde{f}(r,t)$ coincide con *buono*(r,t).

Teorema 14.1. Dato un sistema di regole per un controllo fuzzy, e detta $\tilde{f}$ la funzione fuzzy definita come nel paragrafo precedente, risulta che, per ogni $r \in X$ e $t \in Y$,

$$\tilde{f}(r,t) = D^{\otimes}(p)(Buono(\underline{r},\underline{t})) = m_p(Buono(\underline{r},\underline{t})).$$

Dim. Non esponiamo una dimostrazione generale, peraltro ovvia, ma illustriamo la questione riferendoci all'esempio che abbiamo considerato fino ad ora. Calcoliamo allora *buono*(r,t) cercando tutte le possibili dimostrazioni di *Buono*($\underline{r},\underline{t}$). Considerando la prima regola, otterremo la seguente dimostrazione, che indichiamo con π_1 :

$Buono(x,y) \Leftarrow Piccolo(x) \wedge Lento(y)$	$[\lambda_1]$
$Buono(\underline{r},\underline{t}) \Leftarrow Piccolo(\underline{r}) \wedge Lento(\underline{t})$	$[\lambda_1]$
$Piccolo(\underline{r}) \wedge Lento(\underline{t})$	$[piccolo(r) \otimes lento(t)]$
$Buono(\underline{r},\underline{t})$	$[\lambda_1 \otimes piccolo(r) \otimes lento(t)]$

Pertanto l'informazione fornita da tale dimostrazione su *Buono*($\underline{r},\underline{t}$) è $I(\pi_1,p) = \lambda_1 \otimes piccolo(r) \otimes lento(t)$. Analogamente, tramite la seconda regola otteniamo una ulteriore dimostrazione π_2 di *Buono*($\underline{r},\underline{t}$) che fornisce l'informazione $I(\pi_2,p) = \lambda_2 \otimes small(r) \otimes veloce(t)$. Procedendo per ogni regola e fondendo le informazioni ottenibili in questo modo abbiamo che

$$D(p)(Buono(\underline{r}, \underline{t})) = \\
Max\{\lambda_1 \otimes piccolo(r) \otimes lento(t), \lambda_2 \otimes small(r) \otimes veloce(t), \\
\lambda_3 \otimes medio(r) \otimes moderato(t), \lambda_4 \otimes grande(r) \otimes velocissimo(t), \\
\lambda_5 \otimes verybig(r) \otimes moderato(t)\}.$$

Se si suppone che $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 1$, si ottiene l'uguaglianza $\tilde{f}(r, t) = D(p)(Buono(\underline{r}, \underline{t}))$. $\square$

In questo modo abbiamo visto come l'approccio logico al controllo fuzzy permetta di riottenere e quindi anche di giustificare le procedure utilizzate usualmente per tale tipo di controllo. Naturalmente il potere espressivo della logica va molto oltre quello che si ottiene con programmi positivi. Allora niente vieta che si possano considerare teorie più complesse per ottenere controlli più efficienti. Per prima cosa appare naturale un controllo in cui i gradi di accettabilità $\lambda_1, \lambda_2, \dots$ delle regole non siano necessariamente uguali ad 1. In tale caso nel processo di apprendimento potrebbero essere variate non solo le interpretazioni dei predicati vaghi ma anche i pesi $\lambda_1, \dots, \lambda_n$ assegnati a ciascuna regola. Naturalmente potrebbe persino accadere che nel processo di apprendimento alcune delle regole assumano peso 0 e che quindi tali regole vengano di fatto cancellate.

Ancora, si potrebbero considerare programmi più complessi. Ad esempio potremmo accettare nel corpo di una regola la negazione. Oppure potremmo supporre che oltre ad esistere regole che dicano ciò che è buono ci siano anche regole che dicano ciò che andrebbe evitato in una azione di controllo. Ad esempio supponiamo che nel nostro linguaggio esista un predicato *Cattivo* e che al programma fuzzy p si siano aggiunte le regole

$$\begin{aligned} Rischioso(x, y) &\Leftarrow (\neg Grande(x)) \wedge Veloce(y) & [0.8] \\ Rischioso(x, y) &\Leftarrow (\neg Grande(x)) \wedge Velocissimo(y) & [1] \end{aligned}$$

Per calcolare la relazione fuzzy *rischioso* : $X \times Y \rightarrow U$ che interpreta il predicato *Rischioso* nel modello minimo di Herbrand, avremo una formula del tipo

$$rischioso(x, y) = Max\{(1 - grande(x, y)) \otimes veloce(y) \otimes 0.8, \\ (1 - grande(x, y)) \otimes velocissimo(y)\} ??$$

In tale caso potremmo fondere i due tipi di informazione considerando un ulteriore predicato *Ottimo* ed aggiungendo al fuzzy programma p la regola

$$Ottimo(x,y) \Leftarrow Buono(x,y) \wedge \neg Rischioso(x,y) \quad [1].$$

Successivamente dovremmo calcolare il valore $ottimo(r,t) = D(p)(Ottimo(r,t))$. Ad esempio, se si considera la norma di Łukasiewicz, si avrà che

$$ottimo(x,y) = |buono(x,y) - cattivo(x,y)|_n$$

Naturalmente il processo di defuzzificazione deve poi essere applicato alla relazione fuzzy $ottimo : X \times Y \rightarrow U$.

15. Quantità contro qualità nella scienza moderna

Come già accennato nel paragrafo 11 del primo capitolo, il controllo fuzzy può essere inquadrato nell'ambito di quella scienza ingenua che viene elaborando l'uomo nella sua quotidiana "lotta per la sopravvivenza". Tale scienza ingenua elabora teorie di carattere qualitativo che vengono espresse nel linguaggio comune e si riferiscono agli oggetti dell'esperienza quotidiana. In particolare, il controllo fuzzy può essere visto come tentativo di definire modelli numerici del modo come agiscono tali teorie nell'interazione individuo-mondo allo scopo di rendere possibile adeguate implementazioni. In tale senso esso si pone come un capitolo dell'intelligenza artificiale (riproduzione delle attività intelligenti dell'uomo) e non come alternativa alle usuali "scienze forti". E' l'oggetto di investigazione che è diverso. Nella logica fuzzy si tratta di trovare modelli matematici di una modalità, di carattere essenzialmente qualitativo, che assume l'intelligenza umana. Nelle scienze forti si tratta invece di trovare modelli matematici della struttura intima del mondo con un approccio, di carattere principalmente quantitativo, che vuole esplicitare la dipendenza funzionale tra le grandezze fisiche in gioco.

D'altra parte la scienza moderna è nata proprio da una precisa accusa contro la scienza aristotelica preesistente: quella di procedere ad un tipo qualitativo e "verbale" di indagine scientifica in contrapposizione all'esigenza di un approccio

quantitativo e “numerico”.⁶⁶ Ad esempio per la scienza antica lo studio del comportamento dei corpi liberi, si traduce nell’osservare che la natura degli oggetti li spinge verso il proprio luogo naturale e che per alcuni tale luogo è la terra, per altri il cielo. Non ci si pone il problema di quanto aumenta la velocità di un corpo che si vede precipitare verso terra.

Per esaminare come sia avvenuto questo passaggio dal qualitativo al quantitativo si deve risalire alle discussioni medioevali sulle “qualità” in cui veniva osservato che le “qualità” aristoteliche potevano anche realizzarsi con gradi diversi. Ad esempio “essere caldo” è una qualità che può variare in modo continuo dal freddo all’incandescente. Lo stesso si può dire per la qualità “essere in moto”. Un momento cruciale sarà quando Nicola Oresme in testi come il “*Trattato sulle configurazioni delle qualità e dei moti*” (1353) o *De uniformitate et difformitate intensionum* considera il caso di una qualità che possieda una data *intensità* per ogni punto di un corpo oppure per ogni istante di un dato intervallo di tempo. Ad esempio l’intensità nel caso della qualità “essere caldo” corrisponde alla temperatura che può variare da punto a punto di una barretta. Nel caso della qualità “essere in moto” l’intensità corrisponde alla velocità che può variare da istante ad istante. Oresme rappresenta geometricamente questo fenomeno considerando una linea orizzontale in cui sono rappresentati i punti di un corpo lineare o gli istanti di tempo e rappresentando l’intensità della qualità tramite segmenti verticali che si poggiano verticalmente a tale retta⁶⁷. Da notare che Oresme non si limita alle qualità fisiche, e ciò è espressione del fatto che il suo discorso vuole essere generale. Ad esempio leggiamo il seguente passo circa la misura e la rappresentazione grafica del dolore la cui intensità $d(t)$ è concepita come una funzione del tempo.

⁶⁶ L’accusa ad Aristotele non poteva essere certo di sottovalutare l’osservazione della natura e del mondo che ci circonda. Infatti per Aristotele l’osservazione diretta aveva un ruolo forse maggiore che negli scienziati che lo contestavano, come ad esempio Galileo, Cartesio o Newton.

⁶⁷ Ricordiamo che dovranno passare molti anni prima che la moderna geometria analitica venga inventata.

*Io suppongo quindi che il dolore o l'afflizione sia una certa qualità dell'anima che si sviluppa nel tempo e può intensificarsi con gradi diversi. Quindi è possibile per due di tali qualità essere semplicemente uguali oppure anche che una sia più spiacevole di un'altra.*⁶⁸

Ora l'accettazione di livelli diversi di intensità per una qualità e la sua rappresentazione funzionale hanno qualcosa di simile alla rappresentazione delle proprietà vaghe tramite gli insiemi fuzzy. Tuttavia mentre nella modellizzazione proposta da quest'ultimi i gradi di intensità sono gradi di verità rappresentati da numeri nell'intervallo U , in Oresme l'intensità è una grandezza fisica rappresentata da segmenti che, in termini attuali, corrispondono a numeri reali. Niente connettivi logici quindi, ma derivazione ed integrazione ! Da questo momento quindi:

le qualità escono dal mondo della logica ed entrano in quello dell'analisi matematica e dall'aristotelico Oresme nasce l'approccio quantitativo alla scienza che spodesta la polverosa logica di Aristotele.

Diventa centrale lo studio di funzioni di variabile reale $f: R^n \rightarrow R$ capaci di rappresentare le dipendenze funzionali causa-effetto che regolano il sistema-natura e diventa indiscussa l'equazione:

Scienza moderna = scienza delle quantità

Tale equazione ha avuto tanto successo che la ricerca della rispettabilità scientifica ha spinto e spinge, da Galileo ai giorni nostri, la stragrande parte dei ricercatori ad applicarla nei contesti più improbabili.⁶⁹

⁶⁸ Con le espressioni “semplicemente uguale” e “più spiacevole di” intendeva “avere lo stesso integrale” o “avere integrale maggiore” per le funzioni-intensità del dolore corrispondenti.

⁶⁹ Nel bel film “L'attimo fuggente” si può trovare un esempio di questo desiderio di ridursi a tale paradigma. Nella sua prima lezione il Professor Keating (il Capitano) cita il seguente passo tratto dal libro “Comprendere la poesia” del Professore Emerito Johnathan Evans Prichard.

“Per comprendere appieno la poesia, dobbiamo, innanzitutto, conoscere la metrica, la rima e le figure retoriche e, poi porci due

16. Ritorno alla qualità: il successo del controllo fuzzy

Ora, anche se è vero che tale equazione ha avuto ed ha un valore enorme, non è affatto scontato che tutte le forme di razionalità si possano ridurre ad essa. Infatti se si passa dallo studio della natura ad altre scienze, come le scienze sociali, la psicologia, l'economia, la medicina, il diritto, allora subito appare evidente il limite di un tale paradigma di razionalità. Proviamo ad esempio ad esaminare un'asserzione nel campo delle scelte economiche di questo tipo:

SE il governo non è molto stabile

E

le persone sono sfiduciate

E

non pensi di dovere affrontare spese imminenti

ALLORA

reinvesti buona parte dei tuoi titoli in beni immobili.

E' una asserzione ragionevole ed asserzioni simili si trovano nelle pagine di economia di tutti i giornali. Ma che significano parole come "molto stabile", "sfiduciate", e tutte le altre ? Certamente sarebbe bello avere un libro di economia con una teoria tanto ben elaborata da permettere di ricavare una funzione $y = f(x)$ che ci dica esattamente quale è la somma y da investire in beni immobili in una data situazione x . In fisica esistono molti libri che contengono teorie e formule che fanno una cosa del genere perché non dovrebbe essere possibile fare la stessa cosa in economia ?

domande: uno con quanta efficacia sia stato il fine poetico e due, quanto sia importante tale fine. La prima domanda valuta la forma di una poesia, la seconda ne valuta l'importanza. Una volta risposto a queste domande, determinare la grandezza di una poesia diventa una questione relativamente semplice. Se segniamo la perfezione di una poesia sull'asse orizzontale di un grafico e la sua importanza su quello verticale, sarà sufficiente calcolare l'area totale della poesia per misurarne la grandezza...

La risposta del Professor Keating è netta:

"Cacca! Ecco cosa penso delle teorie di J. Evans Prichard. Non stiamo parlando di tubi, ... stiamo parlando di poesia!"

Similmente, apriamo un testo di legge e leggiamo un articolo del tipo:

“Se una persona ha molestato sessualmente un'altra allora dovrà essere condannata ad un numero y di anni di detenzione variabili da 1 anno a 5 in misura corrispondente alla gravità della molestia.”

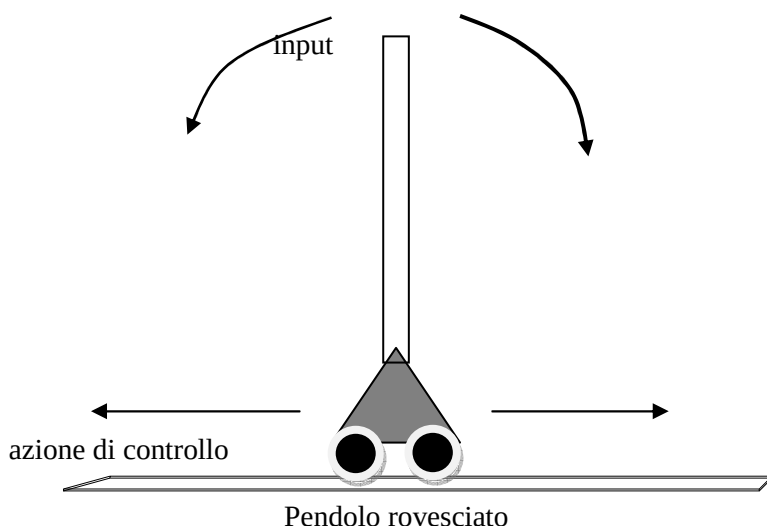
E' chiaro che una giuria nel dovere applicare tale norma si sentirà fortemente imbarazzata a stabilire quale dovrebbe essere il numero corretto di anni di pena da assegnare. Insomma si dovrebbe elaborare una teoria giuridica che definisca una funzione $y = f(x)$ che stabilisca, a seconda del grado x di molestia, il numero y di anni di pena. Si potrebbe in proposito elaborare una casistica complessa ed esaustiva e descrivere in una tabella il numero di anni attribuibili nei vari casi.⁷⁰

In definitiva è un fatto che esistono numerosissimi campi in cui il punto di vista quantitativo non sembra praticabile. Questo ad esempio accade tutte le volte che:

- non si è riusciti a trovare una teoria adeguata
- esiste una teoria ma non si sanno risolvere le equazioni da essa fornite
- esiste una teoria, si sanno risolvere le equazioni da essa fornite ma tali funzioni sono troppo complicate da calcolare (tempo, memoria, costo).

Da notare che tale non praticabilità non si restringe solo alle scienze umane. Essa può manifestarsi anche in campi tradizionalmente appartenenti alla fisica ed all'ingegneria. Ad esempio si cita spesso un importante successo nell'ambito di un problema che appariva irrisolvibile dal punto di vista tradizionale: quello detto del “*pendolo rovesciato*”.

⁷⁰ Mi sembra di ricordare che il rito della confessione nella religione cattolica contempli o abbia contemplato tabelle di tale tipo per una corretta corrispondenza *peccato* $\rightarrow$ *numero di preghiere*. L'uso del controllo fuzzy potrebbe essere un sistema molto più moderno, scientifico ed efficiente e permetterebbe di automatizzare buona parte del lavoro degli uomini di religione permettendo di diminuirne il numero e quindi consentendo allo stato di eliminare o ridurre il 7 per mille.



Il problema è simile a quello di un giocoliere che deve tenere in equilibrio un'asticella poggiata verticalmente su di un dito e può essere illustrato dalla figura indicata sopra dove compare un carrello, che può scorrere lungo un binario, con un asta incernierata nel punto C. Si vuole trovare una forma di controllo che permetta all'asta di mantenersi verticale anche se una piccola oscillazione ad un dato istante la porterebbe a cadere. Il controllo dovrebbe consistere in uno spostamento in avanti o indietro del carrello. Se l'asta verticale tende a cadere in avanti (baricentro in avanti rispetto al punto di appoggio A), allora si deve fare un rapido spostamento in avanti del carrello fino a quando l'asta non raggiunga la posizione (quasi) verticale. L'inerzia però porta a fare sì che il baricentro si trovi ora spostato indietro e quindi che l'asta tenda a cadere indietro. Se ciò accade allora un rapido spostamento indietro del carrello permette di farle assumere una posizione verticale ... e così via. Si manifesta allora il tipico ondeggiare del giocoliere che in generale non ha grandi difficoltà a mantenere in equilibrio l'asticella.

Ebbene, è avvenuto che mentre la realizzazione di un meccanismo di controllo nell'ambito dell'approccio usuale (teoria + equazioni differenziali) non è riuscita, tale realizzazione è risultata invece estremamente semplice tramite il controllo fuzzy. Si è proceduto solo a tradurre in regole del tipo SE ... ALLORA il comportamento naturale di un giocoliere.

CAPITOLO 5

ESISTENZA POTENZIALE, EFFETTIVITA'

1. Esistenza potenziale ed esistenza attuale

Supponiamo che P sia una proprietà vaga definita in un dominio D e che, pur non esistendo un oggetto $d \in D$ verificante pienamente P esistano oggetti in D i quali la verificano con un grado sempre maggiore e sempre più prossimo ad uno. Se indichiamo con $r : D \rightarrow U$ l'estensione di P , ciò significa che esiste una successione $d_1, d_2, \dots$ di elementi di D tale che

- a) $(r(d_n))_{n \in \mathbb{N}}$ è crescente
- b) $r(d_n) \neq 1$ per ogni $d_n \in D$.
- c) $\lim_{n \rightarrow \infty} r(d_n) = \text{Sup}\{r(d_n) \mid n \in \mathbb{N}\} = 1$.

In tale caso la formula $\exists x P(x)$ risulta essere vera nonostante non esista un “testimone semantico” per cui P risulti vera. Ora noi potremmo considerare tale situazione come una “patologia” della logica fuzzy e della sua interpretazione dei quantificatori. Tuttavia potremmo anche guardare a questo fenomeno in positivo, come possibilità di rappresentare due tipi di esistenza: quella potenziale e quella attuale. Questa possibilità è esaminata in [Gerla 2006]a.

Definizione 1.1. Sia α una formula che abbia come sola variabile libera x_i e (D, I) un modello fuzzy, allora diciamo che la formula $\exists x_i \alpha$ è *attualmente vera* se esiste $d \in D$ tale che $I(\alpha, x_i/d) = 1$. Diciamo che $\exists x_i \alpha$ è *vera solo potenzialmente* se è vera ma non attualmente.

Nel primo caso parleremo di *esistenza attuale* nel secondo caso di *esistenza potenziale*.

Definizione 1.2. Diciamo che l'interpretazione fuzzy (D, I) è *ground* se in essa non esistono formule esistenziali che siano vere solo potenzialmente.

Inutile dire che tutte le interpretazioni fuzzy di dominio finito sono ground. Sono ancora ground tutte le interpretazioni che coinvolgono un numero finito di elementi di U , almeno nel caso

in cui la struttura di valutazione considerata è tale che ogni sua sottoalgebra finitamente generata è finita.⁷¹

Nei prossimi paragrafi considereremo due esempi di modelli fuzzy che non sono ground ed in cui è presente un interessante fenomeno di esistenza solo potenziale. Quando non viene specificato altrimenti nel seguito mi riferirò alla logica di Łukasiewicz del primo ordine.

2. Un esempio: la geometria senza punti

Sotto il nome di *geometria senza punti* vanno una serie di tentativi di fondare la geometria senza assumere come primitiva la nozione di punto. Al posto di tale nozione, si preferisce quella di *regione* o *corpo solido*. Per una panoramica sulle varie proposte fatte suggerisco di vedere [Gerla 1995], per un articolo divulgativo consiglio di vedere [Gerla 2006]d.

Il primo a porsi tale problema probabilmente è stato il famoso filosofo inglese A. N. Whitehead nei suoi due libri: *An Inquiring Concerning the Principles of Natural Knowledge* e *The Concept of Nature*. In tali libri è assunta come primitiva, oltre la nozione di *regione*, quella di *inclusione* tra regioni. Pertanto egli considera strutture ordinate del tipo $(R, \subseteq)$ per le quali ipotizza alcune proprietà⁷². Inutile dire che una regione non deve essere intesa necessariamente come insieme di punti e l'inclusione non deve essere intesa necessariamente in senso insiemistico. I punti vengono invece definiti tramite un processo di astrazione al modo seguente.

Definizione 2.1. Una *classe astrattiva* è una successione decrescente di regioni $(r_n)_{n \in \mathbb{N}}$ tale che non esiste una regione che sia contenuta in tutte le sue regioni.

⁷¹ La nozione di interpretazione ground è collegata a quella di interpretazione *witnessed* proposta da Hájek in una serie di fondamentali lavori. Hájek chiama in questo modo le interpretazioni tali che tutte le valutazioni di formule esistenziali si ottengono tramite un massimo e tutte le valutazioni di formule universali tramite un minimo (vedere [Hájek 2007]). Infatti è evidente che ogni interpretazione *witnessed* è anche ground.

⁷² Whitehead non vuole scrivere un testo matematico e quindi non propone un sistema assiomatico ben preciso. Piuttosto vuole mettere in evidenza alcuni aspetti fondamentali della nozione di regione.

L'idea è che una classe astrattiva rappresenta un processo di generazione di regioni sempre più piccole, processo che “al limite” definisce un ente geometrico astratto che potrebbe essere un punto, una linea o una superficie.⁷³ Naturalmente potrebbe capitare che due classi astrattive siano due modi diversi di rappresentare lo stesso ente geometrico. Ad esempio si guardi la

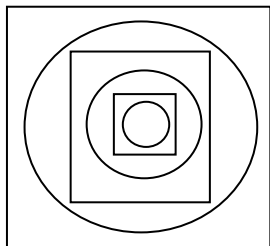


figura affianco in cui è presente una successione di cerchi concentrici ed una di quadrati concentrici. Allora si deve in qualche modo definire una equivalenza tra classe astrattive.⁷⁴ Per affrontare tale questione Whitehead propone la seguente definizione.

Definizione 2.2. Una classe astrattiva $(r_n)_{n \in \mathbb{N}}$ *copre* una classe astrattiva $(s_n)_{n \in \mathbb{N}}$ se per ogni regione r_n esiste una regione s_m tale che $r_n \supseteq s_m$.

La copertura è una relazione di preordine che, al solito modo, induce nell'insieme delle classi astrattive una relazione di equivalenza ed un corrispondente quoziente.

Definizione 2.3. Due classi astrattive sono dette *equivalenti* se si coprono a vicenda. Un *ente geometrico astratto* è una classe completa di equivalenza rispetto tale relazione. Nella classe degli enti geometrici astratti è definita una relazione d'ordine $\leq$ ponendo :

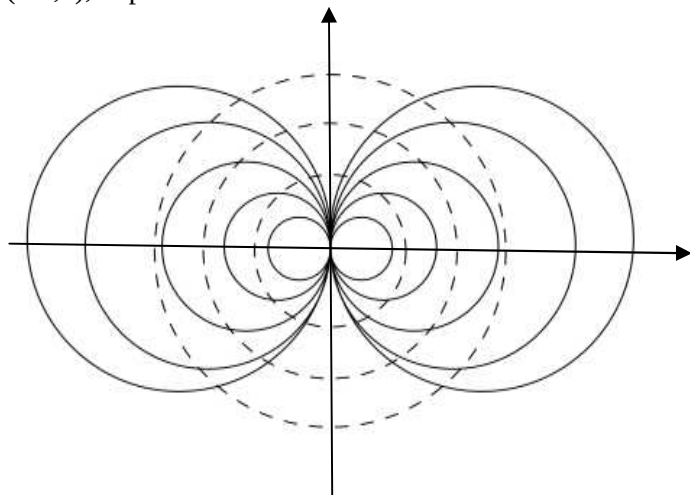
$$[(r_n)_{n \in \mathbb{N}}] \leq [(s_n)_{n \in \mathbb{N}}] \Leftrightarrow (r_n)_{n \in \mathbb{N}} \text{ è coperta da } (s_n)_{n \in \mathbb{N}}.$$

Definizione 2.4. Chiamo *punto* un ente geometrico astratto $P = [(r_n)_{n \in \mathbb{N}}]$ “che non ha parti” cioè che sia minimale rispetto a $\leq$.

⁷³Whitehead non pone la condizione di numerabilità per le classi astrattive. Ho preferito una tale formulazione perché mi sembra meglio rappresentare la dimensione “potenziale” del processo di astrazione che permette di passare dalle regioni all'idea di punto.

⁷⁴Una esigenza simile a quella di introdurre la nozione di equiconvergenza quando si definiscono i numeri reali tramite le successioni di Cauchy.

Ora la proposta di Whitehead anche se molto affascinante ed interessante appare inadeguata. Riferiamoci ad esempio al modello, alquanto naturale, definito dagli aperti del piano cartesiano e dalla relazione di inclusione. In tale modello consideriamo tre successioni di regioni. La prima è la successione $G = (B_n)_{n \in \mathbb{N}}$ delle sfere aperte con centro nell'origine $P = (0,0)$ e raggio $r_n = 1/n$. La seconda e la terza sono invece le successioni G_1 e G_2 definite tramite le sfere di raggio $1/n$ e centri $(-1/n, 0)$ e, $(1/n, 0)$, rispettivamente.



Risulta allora che G copre entrambe le classi G_1 and G_2 , ma che G non è equivalente né alla prima né alla seconda classe. Allora $[G]$ non essendo minimale non rappresenta un punto: pertanto, poiché G , G_1 e G_2 non sono equivalenti, l'origine viene a spezzarsi in (almeno) tre elementi geometrici $P^{\leftarrow} = [G_1]$, $P = [G]$, $P^{\rightarrow} = [G_2]$. Ovviamente questo fenomeno è molto lontano dalle intenzioni di Whitehead e dalla intuizione comune.⁷⁵

Anche se non mi risulta che Whitehead si accorga di tale fenomeno deve comunque avere percepito l'inadeguatezza della sua proposta se pochi anni dopo l'uscita dei primi due libri egli propone un terzo libro intitolato *Process and Reality* in cui viene

⁷⁵ Il fatto che la sola nozione di inclusione si dimostri inadeguata per una approccio alla geometria senza punti non è sorprendente poiché la nozione di inclusione non avendo carattere topologico o metrico non può essere abbastanza potente da rappresentare in modo adeguato lo spazio geometrico.

assunta come primitiva la nozione topologica di *connessione* e quella metrica di *regione ovale*.

Ora in [Gerla e Miranda 2004] e [Coppola, Gerla, Miranda 2010] si esplora la possibilità di utilizzare la logica fuzzy per recuperare il primitivo approccio basato sull'inclusione. Basta accettare che l'inclusione sia una relazione graduata e vedere la definizione di punto come un passaggio dal potenziale all'attuale. Infatti, consideriamo un linguaggio del primo ordine con due costanti proposizionali **0** ed **1**, in cui sia presente il nome di relazione binaria *Incl*. L'interpretazione intuitiva di *Incl* è quella di rappresentare una inclusione graduata tra regioni. Riferendoci a tale linguaggio utilizzeremo:

- $x \subseteq y$ come abbreviazione della formula $Incl(x,y) \leftrightarrow \mathbf{1}$
- $Ug(x,y)$ come abbreviazione della formula $Incl(x,y) \wedge Incl(y,x)$
- $Punto(x)$ come abbreviazione della formula $\forall y(y \leq x \Rightarrow Ug(x,y))$.

L'interpretazione di $x \subseteq y$ è che il grado di inclusione di x in y è 1, l'interpretazione di $Ug(x,y)$ è che x è approssimativamente uguale ad y . Infine l'interpretazione di $Punto(x)$ è, in un certo senso, che " x non ha parti" che coincide con la famosa definizione di punto di memoria euclidea. Infatti tale formula dice che ogni parte di x in realtà è approssimativamente uguale ad x . Detto questo, possiamo considerare una teoria che contenga gli assiomi

A1 $Incl(x,x)$ (proprietà riflessiva)

A2 $Incl(x,y) \wedge Incl(y,z) \Rightarrow Incl(x,z)$ (proprietà transitiva)

A3 $(x \subseteq y) \wedge (y \subseteq x) \Rightarrow (x=y)$ (anti-riflessiva)

A4 $\exists x Punto(x)$.

Chiameremo una tale teoria *teoria degli spazi ad inclusione graduata*.

Teorema 2.5. Esiste un modello fuzzy della teoria degli spazi ad inclusione graduata il cui dominio è la classe $\mathcal{A}$ degli aperti di uno spazio metrico euclideo ed in cui A4 è solo potenzialmente verificata.

Non entriamo nei particolari della dimostrazione che si possono trovare nei lavori citati. Mi limito a dire che in tale modello

l'inclusione graduata è definita tramite la funzione eccesso⁷⁶ $e : \mathcal{A} \times \mathcal{A} \rightarrow [0, \infty]$ ponendo, per ogni coppia di aperti x ed y ,

$$incl(x, y) = |1 - e(x, y)|_n.$$

In tale caso il predicato Ug è interpretato dalla relazione fuzzy u tale che

$$u(x, y) = |1 - d(x, y)|_n$$

avendo denotato con d la distanza di Hausdorff. Il predicato $Punto$ è interpretato dall'insieme fuzzy p tale che

$$p(x) = |1 - D(x)|_n$$

dove $D(x)$ denota il diametro di x . In altre parole, se si prescinde dal problema della normalizzazione, l'inclusione graduata coincide con il duale dell'eccesso, l'uguaglianza graduata con il duale della distanza di Hausdorff ed il predicato "essere puntiforme" con il duale del diametro. Che la formula $\exists x Punto(x)$ sia potenzialmente ma non attualmente verificata segue dal fatto che in uno spazio euclideo si possono trovare insiemi aperti di diametro arbitrariamente piccolo ma non si possono trovare insiemi aperti di diametro nullo.

3. Un altro esempio: la teoria degli insiemi enormi

Consideriamo un linguaggio con identità adeguato per la teoria degli insiemi, cioè contenga:

- i simboli di relazione $\subseteq, \in, \equiv$, per denotare l'inclusione, l'appartenenza e l'equipotenza,
- la costante $\emptyset$, le operazioni insiemistiche $\cup, \cap, -$ e la funzione singleton, denotata con $\{ _ \}$.

- il nome di relazione Eq , che chiamiamo *equipotenza graduata*. L'interpretazione di riferimento è che gli oggetti di cui parliamo sono gli insiemi finiti e che Eq rappresenta una sorta di equipotenza graduata che estende l'usuale equipotenza. Se denotiamo con $Inf(x)$ la formula $Eq(x, x \cup \{x\})$, il significato di $Inf(x)$ è che il numero di elementi di x rimane lo stesso se si aggiunge ad x un elemento, proprietà questa che caratterizza gli insiemi infiniti. Pertanto si potrebbe interpretare tale predicato

⁷⁶ Ricordiamo che in uno spazio metrico (S, d) la *funzione eccesso* si definisce ponendo $e(x, y) = Sup\{d(P, y) : P \in x\}$ dove $d(P, y) = Inf\{d(P, Q) : Q \in y\}$ è la distanza di P da y . La funzione eccesso è utilizzata per definire la famosissima distanza di Hausdorff.

come “essere enorme” e chiamare infinito un insieme che è enorme con grado 1.

Definizione 3.1. La *teoria degli insiemi enormi* è una teoria che esprima le proprietà più importanti della classe degli insiemi finiti e che contenga inoltre i seguenti assiomi:

- A1. $Eq(x,x)$ (riflessività)
- A2. $Eq(x,y) \Rightarrow Eq(y,x)$ (simmetria)
- A3. $Eq(x,y) \wedge Eq(y,z) \Rightarrow Eq(x,z)$ (transitività)
- A4. $x \equiv x' \Rightarrow (Eq(x,y) \Leftrightarrow Eq(x',y))$ (compatibilità).
- A5. $Inf(x) \wedge (x \subseteq y) \Rightarrow Inf(y)$
- A6. $\neg Inf(\emptyset)$
- A7. $\exists x Inf(x)$ (assioma dell'infinito).

Per la teoria degli insiemi enormi vale il seguente teorema.

Teorema 3.2. Esiste un modello fuzzy della teoria degli insiemi enormi il cui dominio è la classe degli insiemi finiti ed in cui A7 è solo potenzialmente verificato.

Dim. Consideriamo l'interpretazione fuzzy (D,I) in cui D coincide con la classe Fin degli insiemi finiti di un qualunque modello della teoria degli insiemi ed I è definita su tutti i simboli della teoria degli insiemi in modo usuale. In più I interpreta Eq ponendo

$$I(Eq)(x,y) = 1 \quad \text{se } x = y = \emptyset$$

$$I(Eq)(x,y) = \frac{\min\{n(x), n(y)\}}{\max\{n(x), n(y)\}} \quad \text{altrimenti}$$

dove $n(x)$ ed $n(y)$ indicano il numero di elementi di x e di y , rispettivamente. In tale interpretazione il predicato Inf è interpretato dall'insieme fuzzy inf definito ponendo

$$inf(x) = n(x)/(n(y)+1).$$

Una semplice verifica mostra che tale interpretazione è un modello della teoria degli insiemi enormi in cui l'esistenza di un insieme infinito è potenziale ma non attuale. $\square$

4. Ultrapotenze per passare dal potenziale all'attuale

Si pone il problema se sia sempre possibile trasformare un modello fuzzy (D, I) in cui sono presenti fenomeni di esistenza potenziale in un modello (D^*, I^*) che sia equivalente a (D, I) ma che sia invece ground. Possibilmente ciò dovrebbe essere fatto individuando un metodo per passare da un processo di generazione di oggetti che *potenzialmente* verificano una proprietà ad oggetti astratti che *attualmente* verificano tale proprietà. Tale questione è esaminata in [Gerla 2006]*a* ed è collegata ad una questione proposta in [Hájek 2007].

Il fenomeno non è nuovo in matematica. Ad esempio l'introduzione dei numeri reali tramite successioni di razionali (successioni di Cauchy) può essere vista proprio in tale chiave. Supponiamo di avere un processo che mi permette di trovare numeri razionali che verificano la proprietà "avere il quadrato prossimo a 2" con un grado sempre più vicino ad uno. In termini di analisi matematica ciò significa che ho individuato una successione $(q_n)_{n \in \mathbb{N}}$ di razionali tali che $\lim_{n \rightarrow \infty} |q_n^2 - 2| = 0$. Allora posso trasformare tale processo in un nuovo "oggetto matematico", che è la classe completa di equivalenza $[(q_n)_{n \in \mathbb{N}}]$ modulo la equiconvergenza. Tale nuovo numero verifica la proprietà "avere il quadrato prossimo a 2" con grado 1.

Naturalmente bisognerebbe vedere quanto questo modo di procedere possa essere considerato generale. Un tentativo interessante può essere fatto tramite la nozione di ultrapotenza definita al modo seguente.

Definizione 4.1. Dato un insieme non vuoto S , un *filtro* su S è un insieme di parti $\mathcal{U}$ di S tale che

$$X \in \mathcal{U} \text{ e } Y \in \mathcal{U} \Rightarrow X \cap Y \in \mathcal{U} ;$$

$$X \in \mathcal{U} \text{ e } Y \supseteq X \Rightarrow Y \in \mathcal{U}.$$

Diciamo che $\mathcal{U}$ è un *ultrafiltro* se oltre a queste due proprietà verifica anche la seguente :

$$X \in \mathcal{U} \Leftrightarrow -X \notin \mathcal{U}.$$

Un esempio tipico di filtro è quello costituito dalla classe dei cofiniti, cioè degli insiemi il cui complemento è finito. Un esempio di ultrafiltro si ottiene fissando x in S e ponendo $\mathcal{U} = \{X \in P(S) : x \in X\}$. Si dimostra (tramite l'assioma della scelta) che

esiste un ultrafiltro contenente il filtro dei cofiniti. Sia P una proprietà definita in S , allora nel caso accada che $\{x \in S : x \text{ verifica } P\} \in \mathcal{U}$ allora si dice che P è *quasi ovunque vera*. L'idea è che un filtro sia una classe $\mathcal{U}$ di insiemi X che sono tanto grandi o i cui elementi siano tanto importanti che è sufficiente che una proprietà P sia vera in un insieme in $\mathcal{U}$ perché si possa considerare vera.

Nel seguito considereremo sempre il caso in cui S coincide con l'insieme N dei naturali. E' possibile definire una nozione di convergenza di una successione rispetto ad un filtro.

Definizione 4.2. Sia $\mathcal{U}$ un filtro in N ed $s \in U^N$ una successione di elementi di U . Diciamo che l è il *limite di s rispetto ad $\mathcal{U}$* e scriviamo $l = \lim_{\mathcal{U}} s$ oppure $l = \lim_{\mathcal{U}} s(i)$ se per ogni intorno I di l risulta che $\{i \in N : s(i) \in I\} \in \mathcal{U}$, cioè $s(i) \in I$ quasi ovunque.

Nel caso in cui $\mathcal{U}$ sia il filtro dei cofiniti tale nozione coincide con quella usuale di convergenza di una successione. Inoltre essa verifica tutte le usuali proprietà della nozione classica. Ad esempio vale il teorema di unicità del limite e se $g : U^n \rightarrow U$ è una funzione continua ed $s_1, \dots, s_p$ sono successioni in U^N , risulta

$$g(\lim_{\mathcal{U}} s_1(i), \dots, \lim_{\mathcal{U}} s_p(i)) = \lim_{\mathcal{U}} g(s_1(i), \dots, s_p(i)).$$

E' possibile inoltre provare le seguenti proprietà.

Proposizione 4.3. Sia $\mathcal{U}$ un ultrafiltro, allora per ogni successione $s \in U^N$, esiste il limite di s rispetto ad $\mathcal{U}$. Inoltre, se $\mathcal{U}$ contiene il filtro dei cofiniti, la convergenza usuale implica la convergenza rispetto a $\mathcal{U}$, cioè

$$\lim_{n \rightarrow \infty} s(n) = l \Rightarrow \lim_{\mathcal{U}} s(n) = l.$$

Dato un insieme non vuoto D ed un ultrafiltro $\mathcal{U}$ in N definiamo una relazione di equivalenza $\equiv$ nella classe D^N di tutte le successioni di elementi di D ponendo

$$f \equiv g \Leftrightarrow \{n \in N : f(n) = g(n)\} \in \mathcal{U}.$$

In altre parole diciamo che due successioni sono equivalenti se sono *quasi ovunque uguali*. Indichiamo con D^* l'insieme $D^N / \mathcal{U} = \{[f] : f \in D^N\}$ delle classi di equivalenza modulo $\equiv$ dove $[f] =$

$\{f \in D^N : f \equiv g\}$. Se $d \in D$, allora indicheremo con $[d]$ la classe di equivalenza della successione costantemente uguale a d . Questo significa che D^* può essere visto come una estensione di D .

Definizione 4.4. Sia (D, I) una interpretazione fuzzy ed $\mathcal{U}$ un ultrafiltro. Allora l'ultrapotenza di (D, I) modulo $\mathcal{U}$ è il modello fuzzy (D^*, I^*) definito ponendo:

$$D^* = D^N / \mathcal{U}$$

$$I^*(c) = [I(c)] \text{ per ogni costante } c.$$

$$I^*(h)([s_1], \dots, [s_p]) = [\langle I(h)(s_1(n), \dots, s_p(n)) \rangle_{n \in N}] \text{ per ogni nome di funzione } h$$

$$I^*(r)([s_1], \dots, [s_p]) = \lim_{\mathcal{U}} I(r)(s_1(n), \dots, s_p(n)) \text{ per ogni nome di relazione } r,$$

Tale definizione coincide con quella classica tranne che, ovviamente, per l'interpretazione dei predicati. Per essa è possibile dimostrare il seguente fondamentale teorema.

Teorema 4.5. Consideriamo una logica fuzzy i cui connettivi logici siano interpretati da funzioni continue e sia $\mathcal{U}$ un ultrafiltro non principale. Allora data una interpretazione fuzzy (D, I) , per ogni formula α ,

$$I^*(\alpha, [s_1], \dots, [s_n]) = \lim_{\mathcal{U}} I(\alpha, s_1(i), \dots, s_n(i)). \quad (4.1)$$

Quindi, in particolare, poiché per ogni $d_1, \dots, d_n \in D$,

$$I^*(\alpha, [d_1], \dots, [d_n]) = I(\alpha, d_1, \dots, d_n), \quad (4.2)$$

(D^*, I^*) è una estensione elementare di (D, I) . Inoltre risulta che (D^*, I^*) è ground.

5. Ma manca un approccio totalmente soddisfacente

Tale approccio al problema del passaggio dal potenziale all'attuale è estremamente generale e molto elegante da un punto di vista matematico. Tuttavia ha vari difetti. Ad esempio proviamo ad applicarlo alla teoria degli spazi di inclusione graduata partendo dai modelli definiti nel teorema 2.5. Supponendo che la dimensione sia uno consideriamo la serie armonica $(r_n)_{n \in N}$ di numeri reali che si ottiene ponendo $r_1 = 1$ e $r_{n+1} = r_n + 1/n$. Tale serie determina una successione di intervalli aperti $I_n = (r_n, r_{n+1})$ uno successivo all'altro

$$(1.0, 1.5), (1.5, 1.8), (1.8, 2.1), (2.1, 2.3), (2.3, 2.5), \dots$$

Poiché

$$p(I_n) = (1 - D(I_n)) \vee 0 = n/(n+1),$$

tale successione permette di affermare che la formula $\exists x \text{Point}(x)$ è potenzialmente soddisfatta. Passando all'ultrapotenza modulo un ultrafiltro non principale $\mathcal{U}$, consideriamo l'elemento $P = [(I_n)_{n \in N}]$ ed osserviamo che

$$I^*(\text{Point})(P) = \lim_{\mathcal{U}} I(\text{Point})(I_n) = \lim_{n \rightarrow \infty} n/(n+1) = 1,$$

Pertanto, come previsto, P soddisfa il predicato Point con grado 1. Purtroppo il “punto” P è alquanto strano poiché essendo per ogni aperto limitato X l'insieme $\{n \in N : X \supseteq I_n\}$ vuoto o finito, nessuna regione limitata contiene tale punto. Questo fatto è molto lontano dalla nostra intuizione e dalla definizione di Whitehead di processo di astrazione anche se appare naturale a chiunque si sia occupato di analisi non standard.

Un discorso simile può essere fatto in rapporto alla teoria degli insiemi enormi. Infatti consideriamo il modello (D, I) che abbiamo proposto nel teorema 3.1. In tale modello consideriamo la classe di intervalli in N definita ponendo

$$I_n = [n(n-1)/2, n(n+1)/2]$$

che costituisce un ricoprimento del tipo

$$[0,1], [1,3], [3,6], [6,10], [10,15], \dots$$

Poiché $I(\text{Inf}(x), I_n) = n/(n+1)$ una tale classe consente di asserire che la formula $\exists x \text{Inf}(x)$ è potenzialmente vera. Consideriamo ora una ultrapotenza (D^*, I^*) di (D, I) ed in tale ultrapotenza l'elemento $[i] \in D^*$ definito dalla successione $i = (I_n)_{n \in N}$. Allora poiché dal fatto che $\lim_{n \rightarrow \infty} n/(n+1) = 1$ segue che $\lim_{\mathcal{U}} n/(n+1) = 1$, risulta che $I^*(\text{Inf}(x), [i]) = 1$. Siamo passati da un infinito potenziale ad un infinito attuale in accordo peraltro con il teorema 4.5. Tuttavia a meno di non essere un seguace dell'analisi non-standard bisogna ammettere che l'oggetto $[i]$ è alquanto strano. Ad esempio, comunque fissiamo $\underline{n} \in N$, poiché l'insieme $\{n \in N : \{\underline{n}\} \subseteq I_n\}$ contiene al più due elementi il suo complemento è cofinito e quindi appartiene ad $\mathcal{U}$. Da ciò segue che $\{\underline{n}\}$ non è “contenuto” in i o, se si vuole, che $\underline{n}$ non è un elemento di $[i]$. Detto in termini di analisi non standard l'oggetto $[i]$ contiene solo elementi non standard.

Probabilmente il metodo delle ultrapotenze non è adeguato poiché si riferisce ad una nozione di equivalenza che non tiene conto della nozione di approssimazione che dovrebbe essere

definita all'interno del dominio. Si dovrebbe potere esprimere il fatto che un elemento del dominio sia una maggiore o minore approssimazione dell'ente astratto che si vuole definire. D'altra parte in Whitehead il processo che porta a trasformare una successione $(d_n)_{n \in \mathbb{N}}$ di regioni in un ente geometrico astratto d è basato sull'idea che d_{n+1} e d_n sono approssimazioni differenti di d e che d_{n+1} sia più prossimo a d di quanto lo sia d_n . Questa idea è trasportabile nel caso della teoria degli insiemi enormi. Si potrebbe allora utilizzare la seguente definizione.

Definizione 5.1. Supponiamo che in un modello fuzzy (D, I) in D sia definita una relazione d'ordine $\leq$, allora chiamiamo *processo di astrazione* ogni successione decrescente di elementi di D . Denotiamo con $D^{\leq}$ la classe dei processi di astrazione. Dato un ultrafiltro $\mathcal{U}$ possiamo allora definire una interpretazione fuzzy $(D_{\leq}^*, I^*)$ tale che

$$D_{\leq}^* = \{[(x_n)_{n \in \mathbb{N}}] \in D^* : (x_n)_{n \in \mathbb{N}} \text{ è un processo di astrazione}\}.$$

Da notare che in tale definizione è necessaria una ipotesi sulla monotonia delle operazioni e dei predicati rispetto a $\leq$. Anche su tale definizione esistono dei dubbi e non conosco una formalizzazione completamente valida per il passaggio dal potenziale all'attuale. Qualche ulteriore (piccola) informazione su tale argomento può essere trovata in [Gerla 2006]a.

6. Effettività e vaghezza⁷⁷

Come è noto, le nozioni di “decidibilità” e di “semidecidibilità” (o “enumerabilità effettiva”) giocano un ruolo centrale nell'ambito della logica classica. Ad esempio consideriamo quello che è il teorema fondamentale della logica del primo ordine: il teorema di completezza. Se svincoliamo il suo contenuto dal particolare sistema inferenziale proposto, ci accorgiamo che di fatto esso consiste nell'asserire che l'insieme

⁷⁷ Per capire quello che si dirà nel seguito sarebbe utile la conoscenza di qualche elemento di teoria della computabilità. Tuttavia di fatto faremo riferimento solo alla nozione di “funzione computabile” da intendere come funzione che può essere computata in un linguaggio di programmazione fissato (supponendo che non esistano limiti di memoria nell'interpretazione di un programma).

delle conseguenze logiche di un insieme decidibile di assiomi è almeno semidecidibile o, equivalentemente, che esiste un processo effettivo per potere produrre le conseguenze logiche di una teoria decidibile. D'altra parte tali nozioni sono strumenti fondamentali per la dimostrazione dei famosi teoremi limitativi, ad esempio la incompletezza e la non decidibilità dell'aritmetica (si veda ad esempio [Shoenfield 1967]).⁷⁸ In definitiva senza la dimensione "decidibilità" una buona parte della logica verrebbe annullata o banalizzata ed essa si limiterebbe alla semantica.

Ora questo discorso si applica altrettanto bene alla logica fuzzy: tale logica non può avere un senso se non si chiarisce prima che cosa significa attribuire al suo apparato deduttivo i caratteri dell'effettività. Purtroppo i risultati che si sono ottenuti in tale ambito non sono molto incoraggianti. Infatti se si guarda all'insieme delle formule logicamente valide allora ci si accorge che manca il requisito essenziale della effettiva enumerabilità (si veda [Scarpellini 1962], [Hájek 1998] e [Montagna 2001]). Addirittura anche la logica monadica non risulta essere decidibile in contrasto con la ben nota decidibilità del calcolo dei predicati monadico classico (si veda [Baaz et al. 2001]).⁷⁹ Pelletier sottolinea in modo tragico (per meglio dire ironico) l'importanza di questo fatto per la logica fuzzy:

"Many observers might think this should be the death knell for fuzzy logic".

Sembra quasi dire che solo i fuzzisti fanno finta di niente nei confronti di questo neo, non certo piccolo, della loro logica. Per sfuggire al suono macabro di tali campane tuttavia esistono delle possibili risposte. Quella più diffusa si riferisce all'approccio "non graded" alla logica fuzzy ed alle nozioni classiche di decidibilità e semidecidibilità. La semidecidibilità dell'insieme delle formule valide è ottenuta al prezzo di cambiare la definizione di validità che viene data non riferendosi ad una

⁷⁸Se non ci preoccupassimo della effettività allora potremmo considerare direttamente come teoria dei numeri interi l'insieme delle formule vere dell'aritmetica. Per definizione questa teoria è consistente, completa ed in essa tutte le asserzioni vere dell'aritmetica sono dimostrabili: nessun problema!

⁷⁹ Si ricordi che logica di Aristotele è monadica e quindi tale indecidibilità potrebbe comportare una difficoltà nell'estendere la teoria dei sillogismi ai predicati vaghi.

singola struttura di valutazione ma ad una sottoclasse di una intera varietà di strutture di valutazione. Tale modo di procedere a mio parere non è molto soddisfacente poiché mi sembra un tantino “ad hoc”. Per chi fosse interessato, tale questione viene esaminata in un mio articolo che può essere trovato in rete [Gerla 2009].

Un altro modo di rispondere, che si riferisce all’approccio “graded” alla logica fuzzy, consiste nel proporre definizioni di semidecidibilità e decidibilità per un insieme fuzzy che tengano conto che ci si muove nell’ambito del continuo. Nei prossimi paragrafi esamineremo come ciò possa essere fatto.

Nota. Il fenomeno della non decidibilità di molti predicati classici è stato utilizzato da M. A. Changizi per dare una interpretazione di tipo informatico della vaghezza [Changizi 1999]. Changizi osserva che esistono proprietà P che, pur non essendo decidibili sono parzialmente decidibili nel seguente senso. Esistono algoritmi i quali rispetto ad una domanda del tipo “ x verifica P ?” possono convergere o meno. Se convergono forniscono risposte corrette. Ad esempio nella logica del primo ordine, data una teoria decidibile T , un predicato rappresentabile dalla formula $P(x)$ ed un numero naturale x , posso generare uno dopo l’altro tutti i teoremi di T . Se compare la formula $P(\underline{x})$ posso fornire in output il valore 1, se compare la formula $\neg P(\underline{x})$ posso fornire il valore 0. Se P non è decidibile in T allora tale algoritmo potrebbe non terminare. Un elemento $\underline{x}$ per cui non esiste convergenza è considerato *borderline*. La vaghezza nascerebbe dalla esistenza di tali casi borderline e quindi nascerebbe dal fenomeno della non decidibilità.

Tale teoria non mi sembra molto convincente poiché la vaghezza si presenta quasi sempre in ambiti finiti (che sono quelli che incontriamo nella vita quotidiana). D’altra parte in un dominio finito tutti i sottoinsiemi sono decidibili e quindi non ha senso fare entrare in gioco questioni di non decidibilità. Un’altra difficoltà è di carattere tecnico. Se $\underline{x}$ è borderline rispetto ad un dato algoritmo ma di fatto verifica P , allora basta aggiungere all’algoritmo una istruzione del tipo IF $x = \underline{x}$ THEN YES per avere un nuovo algoritmo in cui $\underline{x}$ non è più borderline. Pertanto la definizione di borderline non mi sembra ben data.

7. Le macchine di Turing fuzzy

Nel primo capitolo abbiamo già esposto un esempio di algoritmo che coinvolge nozioni vaghe. Si tratta di una semplice ricetta per fare una crema:

1. metti una pentola *quasi piena* di latte sopra una fiamma *debole*
2. aggiungi un uovo
3. aggiungi *un po'* di zucchero
4. se il contenuto è *liquido* allora effettua un giro di cucchiaino altrimenti vai all'istruzione 5
5. spegni il fuoco.

Che significa eseguire tale algoritmo ? Una cosa è certa, non si tratta di un algoritmo deterministico. Persone diverse in generale fanno creme diverse e perfino la stessa persona in giorni diversi potrebbe fare creme diverse. Nel leggere la ricetta, ad ogni passo in cui è coinvolta una nozione vaga siamo in presenza di una decisione da prendere: come interpretare quel “quasi piena di latte” oppure quel “fiamma debole”? Sembra inevitabile rinunciare immediatamente ad ogni pretesa di determinismo. Un'altra cosa da osservare è che ogni esecuzione dell'algoritmo può essere considerata più o meno corretta, cioè più o meno in accordo con quanto scritto. Il grado globale di correttezza può essere calcolato considerando i gradi di correttezza dei singoli passi. Ovviamente, in accordo con la nozione di algoritmo, non stiamo valutando la bontà della crema (cioè se l'output è quello che volevamo). Un cuoco cuoce anche cose che non gli piacciono. Stiamo valutando fino a che punto l'esecuzione della ricetta è stata corretta.

Le prime proposte per formalizzare la nozione di algoritmo in ambito fuzzy furono fatte da E. S. Santos in una interessante serie di lavori. In particolare Santos, partendo da una idea di L. Zadeh [Zadeh 1968], propose la nozione di macchina di Turing fuzzy e la conseguente nozione di funzione fuzzy computabile. In accordo con quanto osservato prima, tale nozione estende quella classica di macchina di Turing non deterministica di cui ricordiamo la definizione.

Definizione 7.1. Una *macchina di Turing non deterministica* è una struttura del tipo $(S, A, I, M, b, q_0, q_f)$ dove:

- (1) S è un insieme finito i cui elementi chiamiamo *stati*
- (2) A è un alfabeto finito
- (3) I è un insieme finito di *simboli input* $I \subseteq A$
- (4) $M \subseteq S \times A \times S \times A \times \{-1, 0, 1\}$ è l'insieme delle *mosse possibili*
- (5) $b \in A - I$ è il simbolo di *spazio vuoto*
- (6) q_0 e q_f sono due stati chiamati *stato iniziale* e *stato di accettazione*, rispettivamente.

Si assume che in ogni istante la macchina stia guardando la casella di un nastro infinito e che stia in uno determinato stato. I valori $-1, 0, +1$ corrispondono ad un passo indietro, fermo, un passo avanti, rispettivamente. Una *descrizione istantanea* è una qualche espressione capace di indicare come sono distribuite le lettere dell'alfabeto sul nastro, quale è lo stato assunto dalla macchina e quale è la posizione della testa di lettura. Ad esempio possiamo chiamare descrizione istantanea una parola Q del tipo $a_1 \dots a_n s a_{n+1} \dots a_m$ il cui significato è che sul nastro è scritta la parola $a_1 \dots a_m$, che la macchina è nello stato s ed osserva la lettera a_{n+1} . Denotiamo con DI l'insieme delle descrizioni istantanee. Data una descrizione istantanea $Q = a_1 \dots a_n s a_{n+1} \dots a_m$, una possibile mossa $(s_1, t_1, s_2, t_2, d) \in M$ permette di passare ad un'altra descrizione istantanea Q' se $s_1 = s$, $t_1 = a_{n+1}$ e Q' si ottiene da Q

- a) sostituendo al posto di t_1 il simbolo t_2
- b) effettuando lo spostamento d della testina
- c) sostituendo al posto di s_1 lo stato s_2 .

In questo caso scriviamo $Q \rightarrow Q'$. Se $x \in I^+$ è una parola dell'alfabeto-input I , chiamiamo *computazione con input x* una successione $Q_1 \dots Q_n$ di descrizioni istantanee tale che Q_1 coincide con $s_0 x$ e $Q_i \rightarrow Q_{i+1}$. Diciamo che una tale computazione è di *accettazione per x* se Q_n contiene lo stato di accettazione.

Definizione 7.2. Diciamo che una parola $x \in I^+$ è *accettata dalla macchina di Turing* $(S, A, I, M, b, q_0, q_f)$ se esiste una computazione di accettazione per x . Un sottoinsieme X di I^+ viene detto *Turing-semidecidibile* se è l'insieme delle parole accettate da una macchina di Turing non deterministica. Diciamo che X è *Turing-decidibile* se sia X che il suo complemento sono Turing-semidecidibili.

Per ottenere la nozione di macchina di Turing fuzzy basta sostituire all'insieme M delle mosse possibili un sottoinsieme fuzzy di M . Ricordiamo che con $\ddot{U}$ denotiamo l'insieme dei numeri razionali in U .

Definizione 7.3. Una *macchina di Turing fuzzy* è una struttura $F = (S, A, I, M, b, q_0, q_f, \mu, \otimes)$ dove $(S, A, I, M, b, q_0, q_f)$ è una macchina di Turing non determinista, $\mu : M \rightarrow \ddot{U}$ è un sottoinsieme fuzzy di M e $\otimes$ è una norma triangolare.

Il numero $\mu(s_1, t_1, s_2, t_2, d)$ viene interpretato come il grado di accettabilità della mossa (s_1, t_1, s_2, t_2, d) .

La nozione di computazione coincide con quella usuale per le macchine di Turing non deterministe ma ad essa si affianca anche una valutazione del grado di accettabilità della computazione.

Definizione 7.4. Il *grado di accettabilità* di una computazione $Q_1 \dots Q_n$ si definisce per ricorsione ponendo:

$$V(Q_1) = 1$$

$$V(Q_1 \dots Q_i) = V(Q_1 \dots Q_{i-1}) \otimes \mu(r) \text{ se } r \in M \text{ è la mossa che ha permesso il passaggio da } Q_{i-1} \text{ a } Q_i.$$

Ovviamente, dato un input x , possono esserci più computazioni di accettazione per x e tali computazioni possono avere gradi diversi di accettabilità. Se si tiene presente che nel caso delle logiche fuzzy il quantificatore “esiste” viene interpretato tramite l'estremo superiore, appare naturale la seguente definizione dove $Acc(x)$ denota l'insieme delle computazioni di accettazione per x .

Definizione 7.5. Data una macchina di Turing fuzzy F l'insieme fuzzy $s : I^+ \rightarrow U$ delle parole accettate da F si definisce ponendo, per ogni $w \in I^+$,

$$s(w) = \text{Sup}\{V(Q) : Q \in Acc(w)\}. \quad (7.1)$$

Diciamo che un insieme fuzzy $s : I^+ \rightarrow U$ è *Turing-semidecidibile* se è l'insieme fuzzy delle parole accettate da una macchina di Turing fuzzy.

Si noti che, coerentemente con il fatto che l'estremo superiore in U dell'insieme vuoto è uguale a 0, se non esistono computazioni di accettazione per x allora $s(x) = 0$.

Come nella teoria della computabilità classica, a partire dalla nozione di semidecidibilità possiamo definire sia quella di insieme fuzzy decidibile che quella di funzione fuzzy computabile.

Definizione 7.6. Un insieme fuzzy s è detto *Turing-decidibile* se sia s che il suo complemento $\neg s$ sono Turing-semidecidibili (si veda [Biacino e Gerla 1989]). Una funzione fuzzy $f: I^+ \times I^+ \rightarrow U$ è *Turing-computabile* se, vista come fuzzy sottoinsieme di $I^+ \times I^+$ è Turing-semidecidibile.⁸⁰

8. Ma tali macchine non si dimostrano adeguate

E' possibile dimostrare che in (7.1) l'estremo superiore è un massimo e che quindi se esiste una computazione di accettazione esiste anche la migliore possibile (si veda [Gerla 2007]). Per provarlo è necessario il seguente importante lemma la cui dimostrazione è in [Wiedermann 2002].

Lemma 8.1. Sia $(A, \otimes, 1)$ un sottomonoido finitamente generato di $(U, \otimes, 1)$, allora ogni sottoinsieme non vuoto di A ammette un massimo, cioè il duale di $(A, \leq)$ è un buon ordine.

Nel seguito indicheremo con $\text{cod}(s)$ il codominio di un insieme fuzzy s , cioè l'insieme dei valori assunti da s .

Proposizione 8.2. Se $\text{Acc}(w) \neq \emptyset$ allora in (7.1) l'estremo superiore risulta essere un massimo.

Proposizione 8.3. Sia s Turing-semidecidibile, allora $\text{cod}(s)$ è tale che ogni suo sottoinsieme non vuoto ammette massimo. Se s è Turing-decidibile, allora $\text{cod}(s)$ è finito.

Possiamo ora provare che la nozione di Turing-semidecidibilità è troppo restrittiva.

⁸⁰ Non è difficile codificare $I^+ \times I^+$ come un insieme di parole di un ampliamento dell'alfabeto I . Da notare che con questa definizione l'output viene fornito fin dall'inizio: alla macchina compete il compito di calcolare il suo grado di accettabilità. Questo perché la macchina è stata considerata come meccanismo di accettazione.

Proposizione 8.4. Consideriamo l'insieme fuzzy $big : I^+ \rightarrow U$ dei “numeri grandi” definito dal porre

$$big(w) = 1 - 1/lunghezza(w)$$

per ogni $w \in I^+$. Allora big non è Turing-semidecidibile a dispetto del fatto che intuitivamente sia decidibile.

Dim. Basta osservare che il codominio $cod(big) = \{1-1/n : n \in N\}$ non ammette massimo. $\square$

Per questo motivo nel prossimo paragrafo proponiamo una nozione più generale di semidecidibilità e quindi di decidibilità.

9. Effettività come limite di approssimazioni calcolabili

Quando si considerano macchine di Turing le nozioni di computabilità e di semidecidibilità vengono date con riferimento all'insieme delle parole I^+ in un dato alfabeto I . Poiché I^+ è in corrispondenza biettiva con N tramite processi di codifica effettiva, le stesse definizioni valgono con riferimento all'insieme dei naturali N . Valgono anche per qualunque insieme S che sia in corrispondenza biettiva con N purché tale corrispondenza abbia il carattere dell'effettività. Questo significa che possiamo parlare di computabilità per funzioni di più variabili, per funzioni definite sull'insieme delle formule di una logica del primo ordine, per funzioni a valori razionali e così via (si veda [Rogers 1976]).

Nella teoria classica della computabilità tra le varie definizioni di sottoinsieme effettivamente enumerabile di un insieme S scegliamo la seguente.

Definizione 9.1. Il sottoinsieme X di S si dice *effettivamente enumerabile* o *semidecidibile* se $X = \emptyset$ oppure se X è il codominio di una funzione computabile $f : N \rightarrow S$. X si dice *decidibile* se è effettivamente enumerabile insieme al suo complemento.

Pertanto dire che un insieme non vuoto X è effettivamente significa che esiste una funzione computabile capace di generare tutti gli elementi di X , cioè tale che

$$X = \{f(1), f(2), \dots\}.$$

Si dimostra che la nozione di semi-decidibile secondo Turing coincide con quella ora data di effettivamente enumerabile. La

seguinte proposizione mostra che può essere anche definita in termini di limite.

Proposizione 9.2. Un sottoinsieme X di S è effettivamente enumerabile se e solo se esiste una funzione computabile $h : S \times \mathbb{N} \rightarrow \{0,1\}$ crescente rispetto alla seconda variabile e tale che, per ogni $x \in S$,

$$c_X(x) = \lim_{n \rightarrow \infty} h(x,n). \quad (9.1)$$

Tale proposizione suggerisce di estendere ai sottoinsiemi fuzzy la definizione di enumerabilità effettiva al modo seguente (si veda [Biacino e Gerla 1987, 1988]).

Definizione 9.3. Un sottoinsieme fuzzy $s : S \rightarrow U$ di S si dice *semidecidibile* se esiste una funzione totale computabile $h : S \times \mathbb{N} \rightarrow \bar{U}$ che sia crescente rispetto alla seconda variabile e tale che, per ogni $x \in S$,

$$s(x) = \lim_{n \rightarrow \infty} h(x,n). \quad (9.2)$$

Una funzione fuzzy di S in S è *computabile* se è un sottoinsieme fuzzy semidecidibile di $S \times S$.

Si osservi che, per l'ipotesi di crescita su h , (9.2) equivale a

$$s(x) = \sup\{h(x,n) : n \in \mathbb{N}\}.$$

La proposizione 9.2 mostra che la definizione proposta estende quella usuale per gli insiemi classici. E' possibile provare che inoltre essa è in accordo con la teoria dei domini effettivi (si veda [Smyth 1977] e [Gerla 2007b]).

Definizione 9.4. Supponiamo che la negazione sia interpretata da una buona negazione computabile sui razionali, allora chiamiamo *decidibile* un insieme fuzzy s che sia semidecidibile insieme al suo complemento $\neg s$.

Osserviamo che dire che $\neg s$ è semidecidibile equivale a dire che

$$\neg s(x) = \lim_{n \rightarrow \infty} h(x,n)$$

con h computabile e crescente rispetto ad n . Ciò equivale a dire che

$$s(x) = \neg \lim_{n \rightarrow \infty} h(x,n) = \lim_{n \rightarrow \infty} \neg h(x,n)$$

e quindi che $s(x) = \lim_{n \rightarrow \infty} k(x, n)$ con k funzione computabile e decrescente rispetto ad n . Possiamo allora enunciare la seguente proposizione.

Proposizione 9.5. Un insieme fuzzy s è decidibile se e solo se esistono due funzioni computabili $h : S \times \mathbb{N} \rightarrow \bar{U}$ e $k : S \times \mathbb{N} \rightarrow \bar{U}$ tali che, per ogni $x \in S$,

- $h(x, n)$ è crescente e $k(x, n)$ è decrescente rispetto ad n ,
- per ogni $n \in \mathbb{N}$, $h(x, n) \leq s(x) \leq k(x, n)$,
- ed inoltre

$$\lim_{n \rightarrow \infty} h(x, n) = s(x) = \lim_{n \rightarrow \infty} k(x, n).$$

Equivalentemente, possiamo esprimere al modo seguente tale caratterizzazione.

Proposizione 9.6. Un insieme fuzzy s è decidibile se e solo se, per ogni $x \in S$, $s(x)$ è il limite di una successione decrescente ed effettivamente computabile di intervalli ad estremi razionali (uniformemente rispetto ad x).

In particolare, da tali proposizioni segue che i valori che può assumere un insieme fuzzy decidibile sono numeri reali ricorsivi.

10. Una tesi di Church⁸¹ per la computabilità fuzzy

Si pone naturalmente il problema se la definizione 9.3 sia la migliore possibile: in altri termini ci si chiede:

E' possibile proporre una sorta di tesi di Church che affermi che tale definizione sia l'adeguata formalizzazione della nozione intuitiva di computabilità in ambito fuzzy ?

Chiameremo “*tesi di Church estesa*” una tesi di questo genere. Come avviene per la tesi di Church classica è evidente che non è possibile darne una dimostrazione poiché il concetto di “nozione

⁸¹ Come è noto la famosa “Tesi di Church” asserisce che la nozione di computabilità tramite macchina di Turing (equivalentemente la nozione di funzione parziale ricorsiva) è completamente adeguata a rappresentare la nozione intuitiva di computabilità. A supporto di tale tesi c'è il fatto che tutte le definizioni proposte di computabilità si sono dimostrate equivalenti a quella proposta da Turing.

intuitiva di computabilità in ambito fuzzy” non è precisato. Tuttavia esistono vari fatti in supporto a tale tesi. Il primo è dato dalle seguente proposizione che mostra come la definizione da noi proposta estenda strettamente quella fornita dalle macchine di Turing.

Proposizione 10.1. Ogni insieme fuzzy che sia Turing-semidecidibile (Turing-decidibile) è anche semidecidibile (decidibile) secondo la definizione 9.3. Tuttavia il viceversa non vale.

Dim. La prima parte della proposizione è ovvia conseguenza della definizione di Turing-semidecidibilità. Per provare che esistono insiemi fuzzy che sono semidecidibili ma non Turing-semidecidibili è sufficiente osservare che se s è computabile come funzione di S in $\tilde{U}$ allora s è anche decidibile ed quindi semidecidibile. Ne segue che l’insieme fuzzy dei numeri grandi che abbiamo considerato nel paragrafo precedente è decidibile anche se, come abbiamo visto, non è Turing-decidibile. $\square$

Naturalmente sarebbe di grande interesse riformulare la definizione di macchina di Turing fuzzy in modo da ottenere una equivalenza delle due nozioni.

A favore della tesi di Church estesa esistono molti altri risultati simili che sono relativi alla nozione di grammatica fuzzy, di algoritmo normale di Markov fuzzy, di programma fuzzy e così via. Inoltre sino ad ora non mi è capitato mai di incontrare qualche forma di algoritmo in ambito fuzzy che non fosse catturato dalla definizione di semidecidibilità.

In particolare è da mettere in rilievo che le definizioni proposte si accordano bene con la nozione di apparato deduttivo nella logica fuzzy. Infatti vale il seguente teorema dove si suppongono alcune ovvie ipotesi di effettività per il sistema deduttivo fuzzy.

Teorema 10.2. Se l’insieme fuzzy v di ipotesi è decidibile allora l’insieme fuzzy $D(s)$ delle formule da esso deducibili è semidecidibile. Se v è anche completo, cioè tale che per ogni formula α , $D(v)(\neg\alpha) = 1 - D(v)(\alpha)$, allora $D(v)$ è decidibile.

Da tale teorema segue il seguente corollario che sembrerebbe in contrasto con il fatto che, come messo in rilievo nel paragrafo 6, sia stato provato che “l’insieme” delle formule logicamente valide non è semidecidibile per le principali logiche fuzzy.

Corollario 10.3. Sotto opportune ed ovvie ipotesi di effettività per l’apparato deduttivo, l’insieme fuzzy $D(\emptyset)$ delle formule logicamente valide risulta essere semidecidibile.

Il contrasto è tuttavia solo apparente. Infatti l’insieme delle formule logicamente valide considerato in letteratura è il taglio di altezza uno di $D(\emptyset)$ e dal fatto che un insieme fuzzy sia semidecidibile non si può inferire che un suo taglio sia semidecidibile. D’altra parte una formula α è logicamente valida con grado 1 se e solo se, detta h la funzione computabile di cui $D(\emptyset)$ è il limite, $\forall k \exists m (|h(x, m) - 1| \leq 1/k)$. Ciò assicura solo che l’insieme delle formule logicamente valide con grado 1 appartiene al livello Π_2 della gerarchia di Kleene.

11. Sfumare o precisare per ottenere decidibilità

In [Gerla 2001]b affronto la questione del rapporto tra la decidibilità e la possibilità di modificare un insieme fuzzy (in particolare un insieme crisp) aumentandone o diminuendone la vaghezza. Mi rifaccio ai formalismi proposti in alcuni articoli fondamentali di Aldo de Luca e Settimo Termini in cui si propongono le nozioni di entropia e di energia nell’ambito della teoria degli insiemi fuzzy (si veda ad esempio [De Luca e Termini 1972]).

Definizione 11.1. Dati due elementi λ_1 e λ_2 in U , poniamo $\lambda_2 \preccurlyeq \lambda_1$ se risulta che

$$\lambda_1 > 1/2 \Rightarrow \lambda_2 \geq \lambda_1 \text{ e } \lambda_1 < 1/2 \Rightarrow \lambda_2 \leq \lambda_1.$$

Se s ed s' sono due sottoinsiemi fuzzy di un insieme S , allora poniamo $s \preccurlyeq s'$ se $s(x) \preccurlyeq s'(x)$ per ogni $x \in S$.

Se $s \preccurlyeq s'$ diciamo che s' è una *versione sfumata* di s o che s è una *versione più precisa* di s' . In altri termini risulta che $s \preccurlyeq s'$ se per ogni $x \in S$ il valore $s(x)$ si ottiene accorciando la distanza di $s'(x)$ da un valore booleano 0, 1. Chiamiamo *borderline* un elemento x tale che $s(x) = 0.5$, ed indichiamo con $s^{0.5}$ l’insieme fuzzy in cui

tutti gli elementi sono borderline, cioè l'insieme fuzzy costantemente uguale a 0.5. Allora $s^{0.5}$ è il massimo rispetto a $\leq$. Gli insiemi crisp sono elementi minimali.

Ci poniamo ora due questioni.

Questione 1.

E' sempre possibile sfumare un insieme fuzzy indecidibile in modo da ottenerne uno decidibile ?

Tale questione si pone in particolare per un insieme crisp poiché si può sperare che la sua indecidibilità possa essere eliminata accettando di perdere in precisione. Ora, ogni insieme fuzzy può essere sfumato nell'insieme $s^{0.5}$ o in un insieme in cui tutti gli elementi sono borderline tranne al più un numero finito. Poiché tali insiemi fuzzy sono decidibili, la prima risposta a tale questione è banalmente positiva. E' opportuno pertanto riformulare la questione perché sia significativa. Ad esempio al modo seguente:

E' sempre possibile sfumare un insieme fuzzy indecidibile in un insieme fuzzy decidibile che non abbia infiniti elementi borderline ?

In questo caso la risposta è negativa.

Teorema 11.2. Sia W un qualunque insieme ricorsivamente enumerabile che non sia decidibile. Allora ogni versione sfumata di W che sia decidibile ha necessariamente infiniti elementi borderline.

Questione 2.

Dato un insieme fuzzy indecidibile s è sempre possibile modificare s in modo da ottenere una versione più precisa di s che sia decidibile ?

Ora se s è crisp, poiché non esiste nessuna versione più precisa di s che non sia s stesso, la risposta non può che essere negativa. Pertanto il problema assume un certo interesse solo se s non è crisp. Il caso più favorevole è che esistano infiniti elementi borderline sui quali si possa agire liberamente. Riformuliamo allora la questione al modo seguente:

Dato un insieme fuzzy s con infiniti elementi borderline è sempre possibile modificarlo in modo da ottenere una versione più precisa di s che sia decidibile ?

Il seguente teorema, la cui dimostrazione richiederebbe un po' troppo tempo (si rimanda a [Gerla 2001]b), fornisce una risposta negativa.

Teorema 11.3. Esiste un insieme fuzzy s semidecidibile e con infiniti elementi borderline tale che nessuna versione più precisa di s può essere decidibile.

Sarebbe interessante ritrovare risultati simili ma legati alla complessità piuttosto che alla decidibilità. Infatti tali risultati mostrerebbero che non esiste un rapporto tra complessità ed vaghezza.

12. Conclusioni (?)

A questo punto dovrei trarre le conclusioni da tutto quanto scritto nel libro. Sarò breve poiché ho trovato sempre insopportabili le conclusioni in quanto sembrano legate all'idea che una questione sia conclusa, cosa che quasi mai avviene. In particolare, il problema della vaghezza è enormemente complesso, forse non dominabile e certamente lontano da conclusioni anche parziali. Infatti, nonostante il fatto che la logica fuzzy sia ormai passata da una fase iniziale "*ingenua*" ad una fase di rigorosa matematizzazione (intesa come produzione di molti articoli che usano strumenti matematici "difficili") esiste ancora qualcosa di insoddisfacente proprio nell'idea di base di utilizzare gli insiemi fuzzy per rappresentare un concetto vago.

Infatti l'interpretazione tramite un insieme fuzzy non può certamente essere un processo preciso. Come abbiamo già osservato nel primo capitolo, la nozione di "grande" può essere rappresentata da diversi insiemi fuzzy, alcuni dei quali sembrano buoni modelli ed altri meno buoni. In un certo senso si dovrebbe interpretare una nozione come quella di "grande" con una classe fuzzy di insiemi fuzzy. 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. In effetti è come se nella nostra mente esistesse una nozione di grande che cambia continuamente. La stessa interpretazione sembra potersi evolvere nel tempo e

principalmente nell'interazione tra soggetti in un gioco linguistico difficilmente dominabile dal formalismo matematico.

Più in generale tutte le nozioni proposte dalla logica fuzzy appaiono troppo rigide. Consideriamo ad esempio la nozione di modello di un insieme fuzzy di ipotesi proposta nella definizione 8.3 del secondo capitolo. Si accetta che m sia un modello di un insieme fuzzy v di ipotesi se risulta che $m(\alpha) \geq v(\alpha)$ per ogni formula α . Ora è evidente che sia m che v non possono essere valutazioni definitive in quanto m dipende da decisioni soggettive di modellizzazione di un predicato vago e v dipende da una valutazione soggettiva circa i possibili valori di verità delle formule. Ora supponiamo che m oppure v siano soggetti a leggere variazioni in seguito ad un processo di "tuning" cioè di aggiustamento (una componente essenziale sia dell'idea di vaghezza sia in tutte le applicazioni della teoria degli insiemi fuzzy). Allora è possibile che m cessi completamente di essere modello di v mentre dovrebbe essere naturale aspettarsi che m continui ad essere, con qualche grado, un modello di v . E' probabile che si possa riformulare la logica fuzzy in modo da evitare questo "paradosso". Ad esempio si potrebbe sostituire la nozione di inclusione crisp tra insiemi fuzzy con una nozione più flessibile di inclusione graduata.

Indice

alfabeto; 55
algoritmi vaghi, esempi; 19
assiomi della logica di Łukasiewicz; 124
assiomi fuzzy della teoria degli insiemi; 128
base di Herbrand; 140
bordeline, casi; 6
bordeline, elemento; 193
buon ordine, perchè quello in N non lo è; 114
catena continua
 di interpretazioni; 73
 di relazioni; 72
Cervino, esempio; 74
clausola di programma; 140
compattezza
 per operatori su insiemi classici; 116
 per operatori su insiemi fuzzy; 118
 per operatori su insiemi fuzzy secondo Pavelka; 116
co-norma triangolare; 48
continuità di un operatore in un reticolo; 117
controllo fuzzy
 approccio logico; 160
 approccio usuale; 157
defuzzificazione; 159
dimostrazione di una formula; 93
dimostrazione fuzzy; 96
eccesso; 176
equipotenza graduata; 176
esistenza, attuale; 171
esistenza, potenziale; 171
fatto; 140
finitismo stretto; 22
formula
 atomica; 57
 universale; 57
formula di una logica del primo ordine; 57
Frege contro la vaghezza; 13
funzione caratteristica; 38
funzione fuzzy
 computabile; 190
 definita nel controllo fuzzy; 159
 definizione; 159
 Turing-computabile; 188
geometria senza punti; 172

giustificazioni in una dimostrazione; 96
grado di appartenenza; 39
grado di inconsistenza; 133
gruppo fuzzy; 66
 $I(\pi, v)$ informazione fornita da una dimostrazione π dato v ; 97
implicazione associata ad una norma; 54
inclusione graduata; 175
 associata alla norma $\otimes$; 61
 come probabilità; 33
insieme
 decidibile; 189
 degli assiomi logici; 93
 diretto; 117
 effettivamente enumerabile; 189
 Turing-decidibile; 186
 Turing-semidecidibile; 186
insieme fuzzy
 decidibile; 190
 degli assiomi logici; 96
 delle conseguenze logiche; 90
 delle tautologie; 91
 semidecidibile; 190
 Turing-decidibile; 188
 Turing-semidecidibile; 187
insiemi
 enormi; 177
insiemi fuzzy
 inclusione, unione, intersezione, complemento; 42
interpretazione fuzzy; 58
 ground; 171
 witnessed; 172
inviluppo inferiore; 120
limite rispetto ad un ultrafiltro; 179
logica monotona; 133
Łukasiewicz logica; 123
macchina di Turing fuzzy; 187
macchina di Turing non deterministica; 185
modello di una valutazione; 88
modificatori linguistici; 4
negazione; 43
 associata ad una norma triangolare; 54
 buona; 49
norma triangolare; 46
oggetto
 crisp; 77

- vago; 77
- omonimia; 79
- operatore di conseguenza
 - in logica classica; 90
 - in logica fuzzy; 90
- operatore di conseguenze immediate
 - in logica fuzzy; 119
 - in programmazione logica classica; 147
 - in programmazione logica fuzzy; 148
- operatore di deduzione; 94
- operatore di deduzione fuzzy; 106
- operatore di deduzione normalizzato; 134
- Oresme Nicola
 - teoria delle qualità; 165
- paradosso
 - dei test uguali; 131
 - del mentitore; 27
 - del mucchio di grano; 7
 - della meccanica relativistica; 10
 - della nave di Teseo; 85
 - della teoria dell'evoluzione; 9
 - dell'uomo calvo; 7
 - di Evans; 84
 - di Poincaré o della indiscernibilità; 97
- parola in un alfabeto; 55
- principio degli indiscernibili di Leibnitz; 63
- principio di comprensione per le proprietà vaghe; 37
- principio di induzione
 - e paradosso del mucchio di grano; 23
 - enunciato; 23
 - perché è falso; 113
- programma; 140
- qualità
 - teorie medioevali; 165
- ragionamenti vaghi, esempi; 18
- regione ovale; 175
- regola
 - del Modus Ponens; 8
 - del Modus Ponens fuzzy; 20
 - di introduzione della congiunzione; 96
 - di particolarizzazione; 95
- regola di inferenza
 - classica; 93
 - estensione canonica di una; 95
 - fuzzy; 94

- regole di inferenza
 - per la logica di Łukasiewicz; 124
- relazione fuzzy; 39
- Russell contro la vaghezza; 13
- semantica della logica probabilistica; 120
- semantica fuzzy
 - astratta; 89
 - vero-funzionale; 58
- similarità; 62
- sinonimia
 - logica basata sulla; 140
 - relazione fuzzy; 138
- sistema deduttivo fuzzy; 96
- sistema inferenziale; 93
- soluzione della logica fuzzy
 - del paradosso dei test uguali; 131
 - del paradosso del mentitore; 28
 - del paradosso del mucchio di grano; 19
 - del paradosso dell'uomo calvo; 130
 - del paradosso della nave di Teseo; 85
 - del paradosso di Evans; 84
 - del paradosso di Poincaré della indiscernibilità; 99
- struttura di valutazione; 44
- supervalutazioni teoria; 24
- taglio
 - di un insieme fuzzy; 72
 - di una interpretazione fuzzy; 73
- teorema di completezza in logica fuzzy; 106
- teoria fuzzy universale; 144
- termine; 56
- Tesi di Church
 - per la logica classica; 191
 - per la logica fuzzy; 191
- ultrafiltro; 178
- ultrapotenza di una interpretazione fuzzy; 180
- vaghezza
 - come si riconosce; 4
 - definizione euristica; 3
- valore di verità di una formula; 59
- valutazione; 88
- versione
 - più precisa di un insieme fuzzy; 193
 - sfumata di un insieme fuzzy; 193

Bibliografia

- Aguzzoli S., Gerla B., Marra V., [2008], De Finetti's no-Dutch-book criterion for Gödel logic, *Studia Logica*, 90, 25-41.
- Ajello M., Spagnolo F., [2002], Alcune osservazioni sperimentali fra senso comune e logica fuzzy, *Quaderni di Ricerca Didattica*, 98-109.
- Arcelli F., Formato F., Gerla G., [1996], Similarity-based Unification as a Foundation of Fuzzy Logic Programming, *Proceedings of the First Workshop on Logic Programming and Soft Computing*, Bonn, Germany.
- Baaz M., Ciabattoni A. Fermüller C. G., [2001], Herbrand's Theorem for Prenex Gödel Logic and its Consequences for Theorem Proving, in *Proceedings of the Artificial Intelligence on Logic for Programming*, Springer-Verlag, London, 201-215.
- Belohlavek R., [2007], Do exact shapes of fuzzy sets matter ?, *Int. J. of General Systems*, 36, 527-555.
- Belohlavek R., [2002], *Fuzzy Relational Systems: Foundations and Principles*, Kluwer Academic/Plenum Press, New York.
- Bennett B., [2002], Physical Objects, Identity and Vagueness, *Principles of Knowledge Representation and Reasoning: Proceedings of the Eighth International Conference*, edited by D. Fensel, D. McGuinness and M. A. Williams, pp 395-406, Morgan Kaufmann, San Francisco, CA.
- Biacino L., Gerla G., [1987], Recursively enumerable L -sets, *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 33, 107-113.
- Biacino L., Gerla G., [1992]a, Generated necessities and possibilities, *International Journal of Intelligent Systems*, 7, 445-454.
- Biacino L., Gerla G., [1992]b, Fuzzy subsets: A constructive approach, *Fuzzy Sets and Systems* 45, 161-168.
- Biacino L., Gerla G., [1996], Connection structures: comparing Grzegorzczuk's and Whitehead's definitions of point, *Notre Dame Journal of Formal Logic*, 37, 431-439.
- Biacino L., Gerla G., [1998], Logics with approximate premises, *International Journal of Intelligent Systems* 13, 1-10.
- Biacino L., Gerla G., [2002], Fuzzy logic, continuity and effectiveness, *Archive for Mathematical Logic*, 41, 643-667.
- Biacino L., Gerla G., Ying M., [2000], Approximate reasoning based on similarity, *Math. Log. Quart.*, 46, 77-86.

- 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.
- Borgo S., Guarino N., Masolo C. [1996]a, A Naive Theory of Space and Matter, in P. Amsili, M. Borillo, and L. Vieu (eds.), *Time, Space and Movement: Meaning and Knowledge in the Sensible World* (Proceedings of the 5th International Workshop), Toulouse: COREP, 29–32.
- Borgo S., Guarino N., Masolo C. [1996]b, *Stratified Ontologies: The Case of Physical Objects*, ECAI-96 Workshop on Ontological Engineering.
- Borgo S., Guarino N., Masolo C. [1996]c, Towards an Ontological Theory of Physical Objects, in *Proceedings of IMACS-IEEE/SMC Conference "Computational Engineering in Systems Applications"* (CESA 96).
- Bottani, A. [2001], Oggetti vaghi e identità vaghe, in *Prospettive della logica e della filosofia della scienza*, a cura di V. Fano, G. Tarozzi e M. Stanzione, Rubbettino.
- Bozzi P., *Fisica ingenua: studi di psicologia della percezione*, Garzanti, 1998.
- Casati R., Varzi A., 2002, *Buchi ed altre superficialità*, Garzanti.
- Chang C. C., Keisler H. J. [1966], *"Continuous Model Theory"*, Princeton University Press.
- Changizi K., [1999], Vagueness, Rationality and Undecidability: A Theory of Why there is Vagueness, *Synthese* 120, 345-374.
- Chakraborty M. K., [1994], Graded consequence: further studies, in c.s. su *Journal of Applied non-classical Logic*.
- Cignoli R., D'Ottaviano I. M. , Mundici D., [1999], *Algebraic Foundations of Many-Valued Reasoning*, Springer London.
- Coppola C., Gerla G., Miranda A., [2010], Point-free foundation of geometry and multi-valued logic, *Notre Dame Journal of Formal Logic*, 51, 383-405.
- De Cock M., Kerre E., [2003], On (un)suitable fuzzy relations to model approximate equality, *Fuzzy Sets and Systems*, 133, 137-153.
- De Luca A., Termini S., [1972], A definition of nonprobabilistic entropy in the setting of fuzzy set theory, *Information and Control*, 20, 301-312.
- De Witt R., [1993], Vagueness, Semantics and the Language of Thought, *PSYCHE*, 1, 039-723

- Di Concilio A., Gerla G., [2006], Quasi-metric spaces and point-free geometry, *Math. Struct. in Comp. Science* 16, 115-137.
- Di Nola A., Gerla G., [1986], Fuzzy models of first order languages, *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 32, 331-340.
- Dubois D., Lang J., Prade H., [1991], Fuzzy sets in approximate reasoning, part 2: logical approaches, *Fuzzy Sets and Systems* 40, 203-244.
- 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.
- Dubois D., Esteva F., Godo L., Prade H., [2007], *Fuzzy set based logics: An history-oriented presentation of their main developments*, in Handbook of the History of Logic. Volume 8, Dov M. Gabbay and John Woods (Editors). Elsevier BV.
- Dummett M., [1975], Wang's paradox, *Synthese*, 30, 301-324.
- Evans G., [1978], Can There Be Vague Objects ?, 38, *Analysis*.
- Fadini A., [1962], Il calcolo delle classi in una logica a tre valori di verità, *Giornale di matematiche di Battaglini*, vol. XC.
- Fadini A., [1979], *Introduzione alla teoria degli insiemi sfocati*, Liguori Editore.
- Ferrante F., Gerla G., [1998], Grasping Infinity by Finite Sets, *Math. Log. Quart.*, 44, 383-393.
- Fine, K., [1975], Vagueness, truth and logic, *Synthese*, 30, 265-300.
- Frege G., [1986], *Scritti postumi* (a cura di Eva Picardi), Bibliopolis, Napoli.
- Genito D., Gerla G., Vignes A., [2011], Handling the valuations of the predicates in a fuzzy model, manoscritto.
- Gerla G., [1986], The category of the fuzzy models and Lowenheim-Skolem theorem, in *Mathematics of Fuzzy Systems*, Di Nola and Ventre (eds), Verlag TUV Rheinland, Germany, 121-141.
- Gerla G., [1990], Pointless metric spaces, *Journal of Symbolic Logic*, 55, 207-219.
- Gerla G., [1994]a, An extension principle for fuzzy logics, *Mathematical Logic Quarterly*, 40, 357-80.
- Gerla, G., [1994]b, Inferences in probability logic, *Artificial Intelligence*, 70, 33-52.

- Gerla G. [1994]c, Fuzzy refutations for probability and multivalued logics, *International J. of Approximate Reasoning*, 4, 257-279.
- Gerla G., [1997], Probability-Like Functionals and Fuzzy Logic, *J. Math. Anal. Appl.*, 216, 438-465.
- Gerla, G. [1995], Pointless Geometries, In: Buekenhout, F., Kantor, W. (eds), *Handbook of incidence geometry: buildings and foundations*. Amsterdam: North-Holland. 1015-1031.
- Gerla G., [1997], Graded consequence relations and closure operators, *Journal of Applied non-classical Logic*, 6, 369-379.
- Gerla G., [1998], Closure operators, fuzzy logic and constraints, in *Fuzzy Sets, Logics and Knowledge-based reasoning*, D. Dubois Editor, Kluwer.
- Gerla G., [2001]a, Fuzzy control as a fuzzy deduction system, *Fuzzy Sets and Systems*, 121, 409-425.
- Gerla G., [2001]b, *Fuzzy logic: Mathematical Tools for Approximate Reasoning*, Kluwer Academic Press.
- Gerla G., [2005], Fuzzy Logic Programming and fuzzy control, *Studia Logica*, 79, 231-254.
- Gerla G. [2006]a, Multivalued logic to transform potential into actual existence, *Studia Logica*, 82, 1-18.
- Gerla G., [2006]b, Effectiveness and Multivalued Logics, *Journal of Symbolic Logic*, 71, 137-162.
- Gerla G., [2006]c, Fuzzy submonoids, fuzzy preorders and quasi-metrics, *Fuzzy Sets and Systems*, 157, 2356-2370.
- Gerla G., [2006]d Un punto dal volto di gatto: fare a meno dei punti, (I) e (II), *Periodico di Matematiche*, 3, 9-20, 4, 15-25.
- Gerla G., [2007]a, Point-free Geometry and Verisimilitude of Theories, *Journal of Philosophical Logic* 36, 707-733.
- Gerla G., [2007]b, Multi-valued Logics, Effectiveness and Domains, in *Computation and Logic in the Real World, Lecture Notes in Computer Science 4497*, Springer, 336-347.
- Gerla G., [2008], Approximates similarities and Poincaré paradox, *Notre Dame Journal of Formal Logic*, 49, 203-226
- Gerla G. [2009], Compactness and Effectivity for fuzzy logics: discussing on some criticisms. See my Homepage.
- Gerla G., [2010], Why I have an extra-terrestrial ancestor, In: I. Capuzzo Dolcetta, M. Transirico, A. Vitolo. *Percorsi Incrociati: in ricordo di Vittorio Cafagna*. Rubbettino ed., Salerno: Collana Scientifica di Ateneo (2010).

- Gerla G., Miranda A. [2004], Graded inclusion and point-free geometry, *International Journal of Pure and Applied Mathematics*, 11, 63-81.
- Gerla G., Sessa M., [1999], Similarity in Logic Programming, in *Fuzzy Logic and Soft Computing*, G. Chen, M. Ying, Kai-Yuan Cai (editors) by Kluwer Academic Publishers (Boston, Dordrecht, London), 19-31.
- Goguen, J.A., [1968/69], The logic of inexact concepts, *Synthese*, 19, 325-373.
- Gottwald S., [2001], *A Treatise on Many-Valued Logics*. Studies in Logic and Computation, vol. 9, Research Studies Press: Baldock, Hertfordshire, England.
- Gottwald S., [2005], Mathematical fuzzy logic as a tool for the treatment of vague information, *Information Sciences*, 72, 41-71
- Gottwald S., [2008], Mathematical fuzzy logic, *Bulletin of Symbolic Logic*, 14, 210-239.
- Hahnle R., [1993], *Automated Deduction in Multiple-valued Logics*, Clarendon Press, Oxford.
- Höhle U., [1991], Foundations of fuzzy sets, *Fuzzy Sets and Systems*, 40, 257-296.
- Hájek P., [1995], Fuzzy logic and arithmetical hierarchy, *Fuzzy Sets and Systems*, 3, 359-363.
- Hájek P., [1998], *Metamathematics of Fuzzy Logic*, Kluwer Academic Publishers, Dordrecht.
- Hájek P., [2002], Fuzzy logic and Lindström's theorem, *Technical report n. 874 Institute of Computer Science, Academy of Sciences of the Czech Republic*.
- Hájek P., Novák V., [2003], The Sorites paradox and fuzzy logic, *Internat. J. General Systems*, 32, 373-383.
- Hájek P., [2007], On witnessed models in fuzzy logic, *Math. Log. Quart.*, 53, 66-77.
- Hájek P., [2009], *Fuzzy Logic*, in Stanford Encyclopedia of Philosophy.
- Jucker A. H., Smith S. W., Ludge T. [2003], Interactive aspects of vagueness in conversation, *Journal of Pragmatics*, 35, 1737-1769.
- Luzzati C., [1999], Le metafore della vaghezza, *Analisi e diritto*, a cura di P. Comanducci e R. Guastini.
- Kosko B. [1995], *Il fuzzy-pensiero*, Teoria e Applicazioni della logica fuzzy, Baldini&Castoldi.

- Klauer D. [1965], *Über einen Ansatz zur mehrwertigen Mengenlehre*, Monatsberichte der Deutschen Akademie der Wissenschaften Berlin, 7, 859-867.
- Klir G., Yuan B. [1995], *Fuzzy sets and fuzzy logic; theory and applications*, Prentice Hall.
- Kuehr J., Mundici D., [2007], De Finetti theorem and Borel states in U -valued algebraic logic, *International Journal of Approximate Reasoning*, 46, 605-616.
- Mamdani, E., [1977], Application of fuzzy logic to approximate reasoning using linguistic systems, *Fuzzy Sets and Systems* 26, 1182-1191.
- Montagna F., [2001], Three complexity problems in quantified fuzzy logic, *Studia Logica*, 68, 143-152.
- Montagna F., [2005], On the predicate logic of continuous t -norm BL-algebras, *Archive for Math. Logic*, 44, 97-114.
- Moruzzi S. [2002], La nuova teoria epistemica della vaghezza, *Atti del VIII Congresso Filosofia & Linguaggio in Italia*, sez. 6, 132-138.
- Noll W., [1993], The geometry of contact, separation, and reformation of continuous bodies, *Arch. Ration. Mech. Anal.*, 122, 197-212.
- Novák V., [1987], First-Order fuzzy logic, *Studia Logica*, 46, 87-109.
- Novák V., [1990], On the syntactico-semantical completeness of first-order fuzzy logic I e II, *Kybernetika*, 26, 47-66 e 134-152.
- Novák V., [2007], Fuzzy logic with countable evaluated syntax revisited, *Fuzzy Sets and Systems* 158 929 – 936.
- Novák V., Perfilieva I., Močkoř [1999], *Mathematical principles of fuzzy logic*, Kluwer Ac. Publishers, London.
- Novák V. and Perfilieva I., [2000] *Discovering the World With Fuzzy Logic*. Springer-Verlag, Heidelberg.
- Oresme N., [1353]. *Trattato sulle configurazioni delle qualità e dei moti*.
- Oresme N., [?]. *De uniformitate et difformitate intensionum*.
- Papamarcou A., Fine T.L. [1986], A note on undominated lower probabilities, *The Annals of Probability*, 14, 710-723.
- Pavelka J., [1979]a, On fuzzy logic I: Many-valued rules of inference, *Zeitschr. F. math. Logik und Grundlagen d. Math.*, 25, 45-52.

- Pavelka J., [1979]b, On fuzzy logic II: Enriched residuated lattices and semantics of propositional calculi, *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 25, 119-134.
- Pavelka J., [1979]c, On fuzzy logic III: Semantical completeness of some many-valued propositional calculi, *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 25, 447-464.
- Pelletier F. J., [1989], Another Argument Against Vague Objects, *The Journal of Philosophy*, 86, 481-492.
- Pelletier F. J., [2000], Review of "Metamathematics of Fuzzy Logic" by P. Hájek, *The Bulletin of Symbolic Logic*, 6, 342-346.
- Pelletier F. J., [2004], On Some Alleged Misconceptions about fuzzy logic, *Artificial Intelligence Review*, 22, 71-82.
- Poincaré H., [1992], *La science et l'hypothèse*, Flammarion Paris.
- Poincaré H., [1904], *La valeur de la science*, Flammarion Paris.
- Reiger A., [2001], The Liar, the strengthened Liar and bivalence, *Erkenntnis*, 54, 195-203.
- Rogers H. [1967], *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, New York.
- Russell, B., Vagueness, [1923], *The Australian Journal of Philosophy and Psychology*, 1, 84-92.
- Sanford, D. H., [1975], Infinity and vagueness, *Philosophical Review*, 84, 521-522.
- Sangalli A. [2000], L'importanza di essere fuzzy: Matematica e computer, Bollati Boringhieri.
- Santos E. S., Wee W. G. [1968]a, General formulation of sequential machines, *Information and Control*, 12, 5-10.
- Santos E. S. [1968]b, Maximin, Minimax, and Composite Sequential Machines, *J. of Math. Anal. Appl.*, 24, 246-259.
- Santos E. S. [1969], Maximin Sequential Chains, *J. of Math. Anal. Appl.*, 26, 28-38.
- Scarpellini B. [1962], Die Nichtaxiomatisierbarkeit des unendlichwertigen Prädikatenkalküls von Łukasiewicz, *J. of Symbolic Logic*, 27, 159-170.
- Smith B. e Casati, R., 1994, 'Naive Physics', *Philosophical Psychology*, 7, 225-44.
- Smyth M., [1977], Effectively given domains, *Theoretical Computer Science*, 5, 257-274.
- Sorensen, R. [2001], *Vagueness and contradiction*, Clarendon Press, Oxford.

- Termini S., [1998], Su alcuni sviluppi delle scienze dell'informazione, *Atti Sem. Mat. Fis. Univ. Modena*, Supplemento al Vol. XLVI, 971-989.
- Toffalori C., [2011], *L'aritmetica di cupido: Matematica e letteratura*, Guanda.
- Turunen E., [1999], *Mathematics Behind Fuzzy Logic*, Physica-Verlag.
- Vallauri L. [1981], *Corso di Filosofia del diritto*, Cedam.
- Valverde L., [1985], On the Structure of *F*-Indistinguishability Operators, *Fuzzy Sets and Systems*, 17, 313-328.
- Varzi A. C., [1999], *Le strutture dell'ordinario*, in Luigi Lombardi Vallauri (ed.), *Logos dell'essere, logos della norma*, Bari: Editrice Adriatica, 489-530.
- Varzi A. C. [2001], I confini del Cervino, in *Prospettive della logica e della filosofia della scienza*, a cura di V. Fano, G. Tarozzi e M. Stanzione, Rubbettino, 431-445.
- Varzi A. C. [2002], Ontologia e metafisica, in *Storia della Filosofia Analitica*, a cura di Franca D'Agostini e Nicla Vassallo, Torino: Einaudi, 81-117, 521-526.
- Varzi A. C., [2005], *Ontologia*, Laterza (Biblioteca Essenziale).
- Varzi A. C., [2010], Vaghezza e ontologia, di prossima pubblicazione in "*Storia dell'ontologia*", a cura di Pietro Kabau e M. Ferraris, Bompiani, Milano.
- Vojtáš P., [2001], Fuzzy logic programming, *Fuzzy Sets and Systems*, 124, 361-370.
- Weichselberger K. , Pohlmann S. [1990], *A methodology for Uncertainty in Knowledge-Based Systems*, Springer-Verlag, Berlin.
- Whitehead, A. N. [1919], *An Inquiry Concerning the Principles of Natural Knowledge*, Univ. Press. Cambridge.
- Whitehead, A. N. [1920], *The Concept of Nature*, Univ. Press. Cambridge.
- Whitehead, A. N. [1929], *Process and Reality*, Macmillan, N.Y.
- Wiedermann, J. [2002], Fuzzy Turing Machines revised, *Inform. and Computing*, 21, 1-13.
- Williamson, T., [1994], *Vagueness*, Routledge, London.
- Williamson, T., [2000], *Knowledge and its Limits*, Oxford University Press, Oxford.
- Wright, C., [1995], 'The epistemic conception of vagueness', *The Southern Journal of Philosophy*, 33 (Supplement).

- Ying M.S., [1994], A logic for approximate reasoning, *The Journal of Symbolic Logic*, 59, 830-837.
- Zadeh, L. A., [1965], Fuzzy Sets, *Information and Control*, 12, 338-353.
- Zadeh L. A. [1973], Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Systems, Man, and Cybernetics* 1, 28-44.
- Zadeh L. A. [1975], The concept of a linguistic variable and its application to approximate reasoning I, II, III, *Inf. Sci.* 8, 9 199-275, 301-357, 343-80.
- Zadeh, L. A., [1975], Fuzzy logic and approximate reasoning, *Synthese*, 30, 407-428.
- Zadeh L. A. [1979]. "Liar's paradox and truth-qualification principle". *ERL Memo M79/34*.

In italiano e divulgativi

Cammarata S., [1994], *Sistemi fuzzy. Un'applicazione di successo dell'intelligenza artificiale*, EtasLibri.

Cammarata S., [1997], *Sistemi a logica fuzzy. Come rendere intelligenti le macchine*, EtasLibri.

Gerla G., [1999], Logica fuzzy e paradossi. *Lettera Matematica Pristem*, 32, 31-39.

Gerla G., [2006]a, Un punto dal volto di gatto: fare a meno dei punti, (I), *Periodico di Matematiche*, 3, 9-20.

Gerla G., [2006]b, Un punto dal volto di gatto: fare a meno dei punti, (II), *Periodico di Matematiche*, 4, 15-25.

Kosko B., [1995], *Il fuzzy pensiero. Teoria ed applicazioni della logica fuzzy*, Baldini e Castoldi.

Paganini E., [2008], *La vaghezza*, Carrocci.

Pizzaleo A. G., [2003], *Fuzzy logica. Come insegneremo alle macchine a pensare da uomini*, Castelvechi.

Sangalli A., [2000], *L'importanza di essere fuzzy*, Bollati Boringhieri, Torino.

Veronesi M., Visioli A., [2000], *Logica fuzzy: fondamenti teorici ed applicazioni pratiche*, Franco Angeli.