DOI: 10.13137/2039-8646/33405

Sillogismi sfocati: il modus ponens*

ANDREA SGARRO
Dipartimento di Matematica e Geoscienze
Università di Trieste
sgarro@units.it

LAURA FRANZOI
Center for Computational Linguistics, Department of Computer Science
University of Bucharest (Romania)
laura.franzoi@gmail.com

ABSTRACT

Classical logic is usually associated with the name of Aristotle and is strictly binary (a statement is either false or true). Fuzzy logic, initiated by L. A. Zadeh (1921-2017), makes use of all the gradations of gray, intermediate between white = false and black = true, and has proved extremely successful in building intelligent machines and industrial plants, ranging from simple household appliances to complex subway networks. The inferential machinery of fuzzy control, used to make suitable decisions, is a bold generalization of the classical syllogism called modus ponens. This article will deal with this topic.

PAROLE CHIAVE

LOGICA SFOCATA / FUZZY LOGIC; COMPUTAZIONE FLESSIBILE / SOFT COMPUTING; SILLOGISMO / SYLLOGISM; MODUS PONENS / MODUS PONENS.

Allez en avant, et la foi vous viendra
Jean-Baptiste d'Alembert (1717-1783)

1. IL *MODUS PONENS*

Il più noto dei sillogismi aristotelici è il *modus ponens* che di solito viene esemplificato coinvolgendo Socrate:

premessa: Socrate è un uomo

regola d'inferenza: gli uomini sono mortali

conclusione o *deduzione*: Socrate è mortale

* *Title: Fuzzy syllogisms:* modus ponens.

(*Attenzione*: in molti manuali di logica l'ordine seguito è quello di far precedere la regola d'inferenza alla premessa, e di chiamarle rispettivamente *premessa maggiore* e *premessa minore*[1]). Riformuliamo la regola di inferenza in modo da avvicinarci ai simboli della matematica:

regola d'inferenza: se u è un uomo, allora u è mortale

Per avvicinarci ancor di più alla matematica abbandoniamo Socrate e passiamo ai numeri:

premessa: 6 è un numero pari
regola d'inferenza: se n è pari, allora n è divisibile per 2
conclusione: 6 è divisibile per 2

Implicito è un *universo di riferimento* cui la variabile *n* è vincolata ad appartenere, nel nostro caso l'insieme dei numeri naturali $1, 2, 3, 4, 5, 6, 7$ e così via. Fin qui nulla di insolito, ma come reagiremmo di fronte a una regola del tipo:

regola d'inferenza: se x è grande, allora il suo inverso $\frac{1}{x}$ è piccolo
sull'universo dei numeri reali positivi $x > 0$?

Cominceremo con l'osservare che il problema è serio, perché di regole d'inferenza di questo tipo il linguaggio *naturale* - quello che parliamo e su cui si fondano i nostri ragionamenti - abbonda e di regole siffatte si servono persino i matematici più raffinati, sia pure in fasi preliminari dell'elaborazione di un loro teorema, quando la dimostrazione formale deve ancora venir elaborata. Sono il linguaggio naturale e le deduzioni su di esso basate, pur nella loro "debolezza" logica, che consentono alla nostra specie non solo di sopravvivere ma anche di prosperare e magari di mettere a rischio la sopravvivenza di specie diverse: non meraviglia che l'intelligenza artificiale, *AI* per usare l'acronimo americano di *Artificial Intelligence*, miri a imitare l'intelligenza

[1] Per la terminologia in materia, e per tutta la teoria del *sillogismo classico*, consiglio il *Manuale di Logica* di Willard Van Orman Quine (cfr. Van Orman Quine 1961, pp. 100-118).

naturale, umana, cercando di impadronirsi di regole *vaghe* o *sfocate* come quella appena enunciata e di controllarle (gestirle) senza però snaturarle.
Certo, è impossibile definire con precisione accettabile, o per meglio dire con *im*precisione accettabile, che cosa vogliano dire *grande* o *piccolo*, ma non appena chiariamo il contesto in cui stiamo operando il discorso si fa meno allarmante. Se per esempio parliamo di automobili e di velocità in chilometri all'ora, 180, lo converranno tutti a parte i più sofistici, è un numero grande, mentre 15 è un numero piccolo. Se parliamo di condizionatori d'aria e di temperature misurate in gradi centigradi, 12 è piccolo, mentre 28 è grande. L'ultimo esempio non è affatto casuale: condizionatori d'aria "intelligenti" basati sulla logica *vaga*, *sfumata*, o, per usare il termine tecnico inglese *fuzzy = sfocata*[2], esistono, sono in commercio e funzionano sorprendentemente bene.
Neppure la logica sfocata, tuttavia, può fare a meno dei numeri "normali", quelli *nitidi* o *crisp* senza ombra di vaghezza: le parole del discorso naturale, *verbale*, per poterle gestire vanno prima "numerizzate".

2. NUMERIZZAZIONE E VERBALIZZAZIONE

Nelle sezioni che seguono percorreremo prima le due fasi della *numerizzazione* e della *verbalizzazione*, per poi passare alla *fuzzification* e alla *defuzzification*, o per usare due riprovevoli neologismi di pronuncia incerta ma di uso standard nell'ambiente dei "fuzzisti": alla *fuzzificazione* e alla *defuzzificazione* - ma forse sarebbe meglio dire *focalizzazione* o *messa a fuoco* nell'ultimo caso, rimanendo aperto il problema di trovare un termine che traduca *fuzzification*: forse *sfocalizzazione*?
Per esemplificare partiamo da un numero "vago" X che potrebbe essere, in un certo istante, la velocità X di una macchina in chilometri all'ora. Se della velocità X ciò che sappiamo si limita a un *attributo linguistico*, vale a dire che la velocità è grande, converrà procedere a numerizzare questo attributo in una maniera forse arbitraria, ma non per questo irragionevole. Se avete l'impressione che a partire da questo punto stiamo

[2] Cfr. KOSKO 1995; SGARRO, FRANZOI (cfr. Siti web).

"dando i numeri", a ben pensarci è proprio quello che fa un coscienzioso insegnante quando numerizza tramite un voto in decimi o in trentesimi un attributo verbale come "discreto" relativo a un esame appena concluso: in effetti pensare ai *valori logici* che via via assegneremo come a dei *voti* sulla scala da 0 a 1 può essere illuminante. Per procedere alla numerizzazione degli attributi verbali ci aiuterà il contesto in cui stiamo operando: con una comune automobile forse un po' *vintage* come quella del primo autore, velocità superiori a 140 chilometri all'ora sono decisamente grandi, quelle inferiori a 30 altrettanto decisamente non lo sono. In un caso simile la numerizzazione abituale, quella più semplice com'è richiesto da quel ramo dell'intelligenza artificiale che chiamiamo *soft computing* o *computazione flessibile*, tenendo anche conto che velocità superiori ai 250 chilometri all'ora escono via via dalla portata del motore e quelle superiori ai 300 chilometri all'ora sono tecnicamente irraggiungibili (il motore fonderebbe), è quella di ricorrere a un trapezio *fuzzy* $(0, 30, 250, 300)$ come in Figura 1:

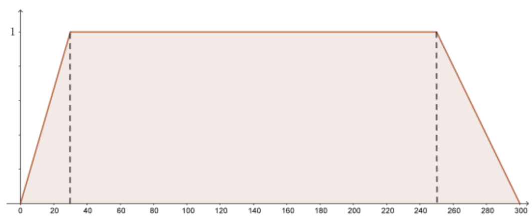

Figura 1. Il trapezio *fuzzy* (0, 30, 250, 300).

Perché proprio un trapezio che lascia i 140 chilometri all'ora al centro del suo lato superiore? Perché è semplice, perfino primitivo, e nel *soft computing*, a costo di esser rozzi, la semplicità è una virtù irrinunciabile. Nella figura possiamo vedere la funzione a valori che variano fra 0 e 1 come una specie di *funzione di densità* $f(x)$ simile a quelle che usano i fisici o i probabilisti. I valori della funzione sono nel nostro caso *valori logici* ossia *gradi di verità*: possono essere uguali a 0 o 1 come nella logica classica

0 = falso = del tutto impossibile = no = bianco,

1 = vero e dunque pienamente possibile = sì = nero

ma possono anche assumere valori intermedi, consentendo tutte le sfumature di grigio, proprio come avviene nel linguaggio naturale, ricco di proposizioni quasi vere, abbastanza vere, possibili ma non sicure e così via. Nel nostro caso la funzione di densità $f(x)$ *codifica* l'attributo linguistico "grande" e ha un andamento trapezoidale sull'intervallo $[0, 300]$ ed è nulla altrove; il valore logico ossia il grado di verità della proposizione «La velocità è uguale a 20 km all'ora» è di $\frac{2}{3}$. Attenzione: le funzioni di densità che compariranno nella premessa e nella regola d'inferenza saranno tutte *normali*, nel senso che il valore 1 viene sempre raggiunto: non sarà così, come vedremo, nel caso della funzione che comparirà nella conclusione, che sarà di solito (e lo sarà sempre nei nostri esempi) *subnormale* o *incompleta*, il che aprirebbe un interessante disquisizione filosofica in cui tuttavia non ci addentreremo, né apriremo un dibattito se i termini "valore logico" e "grado di verità" possano esser considerati sinonimi, una nostra scelta forse opinabile ma tutto sommato irrilevante ai fini di quanto segue.

In modo simile, se il numero vago o *fuzzy* X fosse uguale a circa 80 km all'ora potremmo servirci di un triangolo $(79, 80, 81)$ che escluda velocità inferiori ai 79 o superiori agli 81 km all'ora (cfr. Figura 2).

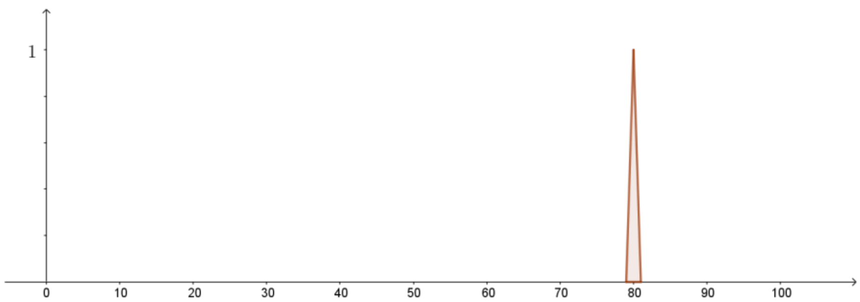

Figura 2. Il numero *fuzzy* 80 codificato con il triangolo (79, 80, 81).

Come nella logica classica, anche in quella *fuzzy* i connettivi logici fondamentali sono la *disgiunzione* (= oppure indifferentemente, OR), la *congiunzione* (= e anche, AND) e la *negazione* (= NOT). La disgiunzione OR e la congiunzione AND ricorrono rispettivamente al massimo e al minimo dei due gradi di verità coinvolti, ciò che non sorprende: se ad esempio abbiamo assegnato i due gradi di verità (i due "voti") 0,4 e 0,9 alle proposizioni

«Sofia ha una bella voce di contralto» e «Sofia è una virtuosa del violino» il grado di verità della disgiunzione OR è $0,9$ e potrebbe interessare chi abbia deciso di usare Sofia in un complesso musicale a cui finora mancavano contralto e violino (rimarrebbe aperto il problema di trovare un buon contralto), mentre il grado di verità della congiunzione AND è solo $0,4$ e potrebbe interessare chi abbia bisogno di una musicista che sia canti da contralto sia suoni il violino. La negazione "complementa" a 1 il grado di verità, per cui il valore logico della proposizione «Sofia *non* è una virtuosa del violino» è $1 - 0,9 = 0,1$.
Decisamente sorprendente è invece l'*implicazione fuzzy*, la cui struttura è *antecedente implica conseguente*, nella sua variante più usata nelle applicazioni (nel *controllo fuzzy*): l'*implicazione di Mamdani*, così chiamata dal nome dell'ingegnere anglo-egiziano che l'ha proposta, si ispira al cosiddetto *periodo ipotetico della realtà* del linguaggio naturale e ne abbiamo parlato in esteso in un lavoro comparso in un precedente numero della rivista[3].
L'implicazione di Mamdani si discosta fortemente dall'implicazione della logica classica, aristotelica: è *del tutto equivalente alla congiunzione logica* AND per cui ha un alto grado di verità solo quando *sia* l'antecedente *sia* il conseguente dell'implicazione sono veri o quasi veri. Il linguaggio naturale tiene ben separati il periodo ipotetico della realtà («Se vieni ti offro un caffè») contrapposto a quello dell'irrealtà («Se il ferro fosse oro sarei ricco sfondato»), ricorrendo a tempi verbali diversi, indicativo (realtà) contro congiuntivo-condizionale (irrealtà). L'implicazione «Se il ferro fosse oro sarei ricco sfondato» ha dunque un valore di verità *basso* o magari nullo secondo Mamdani, alto secondo la scuola megarica (V-IV secolo prima dell'era comune) e per Aristotele (nella logica classica dal falso si può dedurre quel che si vuole e l'implicazione resta "sana").
Supponiamo di avere due *variabili linguistiche* A e B, che potrebbero essere velocità e decelerazione, variabili che possono godere di attributi come grande o piccolo, attributi che dobbiamo numerizzare mediante funzioni di densità di tipo logico.
Qual è la struttura di un sillogismo *fuzzy*, sillogismo che è alla base del ragionamento logico *fuzzy* ed è la "macchina" che ne consente il moto e l'avanzamento? Abbandoniamo

[3] Cfr. SGARRO, FRANZOI 2016.

l'automobile e rimaniamo per un po' sull'astratto: ecco finalmente il sillogismo del *modus ponens*, scritto in una maniera molto sintetica anche se poco amichevole:

premessa: A è $f(x)$
regola d'inferenza: se A è $\varphi(x)$ allora B è $\psi(y)$
conclusione: dunque B è $g(y)$

Meno sinteticamente la premessa andava scritta:

- A gode dall'attributo linguistico codificato dalla funzione $f(x)$;

la *regola d'inferenza*, che ha carattere ipotetico, si può scrivere più in esteso:

- nel caso in cui A goda dall'attributo linguistico codificato dalla funzione $\varphi(x)$ allora B gode dall'attributo linguistico codificato dalla funzione $\psi(y)$.

Le tre funzioni $f(x), \varphi(x)$ *e* $\psi(y)$ sono ben note (è già avvenuta la numerizzazione dei rispettivi attributi linguistici) e vanno usate nel calcolo di $g(y)$.
Il *modus ponens* sfocato, pur ispirandosi all'omonimo sillogismo aristotelico sulla mortalità di Socrate, se ne allontana parecchio per cui converrà argomentarlo su un esempio specifico di tipo linguistico con riferimento proprio ai linguaggi naturali come l'italiano, e non a quelli simbolici, formali, mirando fin da subito a imitare la logica del buon senso, o *common sense logic*, di cui il linguaggio ordinario sa servirsi con incredibili e quotidiani successi pratici:

premessa: X è circa 3
regola d'inferenza: se X è grande, allora Y è piccolo
conclusione: Y gode di un attributo linguistico che va calcolato sulla base della premessa e della regola d'inferenza

Il problema è quello, fatti i calcoli, di riuscire a definire in maniera adeguata l'attributo incognito che figura nella conclusione (beninteso si tratta di un attributo linguistico di natura numerica, proprio come, a codifica avvenuta, "circa 3" nella premessa e "grande" e "piccolo" nella regola d'inferenza), in maniera da poter esplicitare le affermazioni su Y contenute nella premessa e nella regola d'inferenza, anche se in maniera ancora solo implicita e nascosta.

Pensiamo ai valori *nitidi* x che compaiono nella premessa e che competono alla X con un certo valore logico più o meno alto come se fossero gli anelli di una catena. Fissiamo per il momento uno solo di questi anelli, ossia blocchiamo la x. Il valore logico della conclusione è alto quando alto è *sia* il valore logico della premessa e *inoltre* è alto il valore logico dell'implicazione che figura nella regola di inferenza. Dunque, per x temporaneamente bloccato e usando i simboli di congiunzione e di implicazione logica, rispettivamente ∧ e →:

$$g_x(y) \;=\; f(x) \,\wedge\, [\varphi(x) \,\rightarrow\, \psi(y)] \;=\; \psi(y) \,\wedge\, f(x) \,\wedge\, \varphi(x)$$

Ci siamo serviti dell'implicazione di Mamdani equivalente alla congiunzione che restituisce il minimo dei due valori logici coinvolti, e abbiamo usato l'associatività e la commutatività dell'operazione aritmetica di minimo il cui simbolo è di nuovo il "cuneo" ∧ ad esempio $3 \,\wedge\, 5 \,=\, 5 \,\wedge\, 3 \,=\, 3$ (che lo stesso simbolo indichi al contempo un'operazione logica fra gradi di verità e un'operazione aritmetica fra numeri non crea problemi visto come abbiamo definito la congiunzione sfocata AND). Come scegliere l'anello della catena? Si usa l'anello più forte, quello che rende più vero il sillogismo *in toto*:

$$g(y) = \max_{\mathrm{x}}[\psi(y) \,\wedge\, f(x) \,\wedge\, \varphi(x)] \;=\; \psi(y) \wedge \max_{\mathrm{x}}[f(x) \,\wedge\, \varphi(x)]$$

dove, com'è legittimo, abbiamo messo in evidenza il termine $\psi(y)$ che non dipende dalla x (i matematici più attenti avrebbero usato l'estremo superiore invece del massimo, ma le funzioni coinvolte sono di norma così regolari che tutto torna). Il massimo in x è un certo valore numerico che chiameremo h e dunque:

$$g(y) = \psi(y) \wedge h\,, \qquad h \,=\max_{\mathrm{x}}[f(x) \,\wedge\, \varphi(x)]$$

per cui in pratica la funzione di densità $g(y)$ è identica alla funzione-attributo $\psi(y)$, solo "tagliata" ad altezza h. A ritroso potremmo chiederci quale sia l'aggettivo della lingua italiana che viene numerizzato mediante la funzione-attributo $g(y)$ che abbiamo appena costruito. Ovviamente tale aggettivo non è verosimile pescarlo nello Zingarelli, ma potremmo inventarlo, e decidere che il neologismo *grungo* sia un aggettivo ben numerizzato dalla $g(y)$ nello stesso senso in cui l'aggettivo *piccolo* è ben numerizzato dalla $\psi(y)$. Allora il nostro ragionamento diventa, in buon toscano audacemente ampliato:

se X è circa 3 e se X grande implica Y piccolo, allora Y è *grungo*

ciò che ne evidenzia la natura linguistica, da linguaggio ordinario e da logica del buon senso. Attenzione: ci siamo resi colpevoli di quattro scelte arbitrarie, che sono tre numerizzazioni, quella di *circa* 3 tramite il triangolo $(2, 3, 4)$, quella di *grande* tramite il trapezio $(2, 5, 10, 10)$ (stiamo ipotizzando che valori superiori a 10 non abbiano più senso nel nostro contesto) e quella di *piccolo* tramite il trapezio $(0, 0, 3, 5)$, oltre a una verbalizzazione che ci ha fatto inventare l'aggettivo *grungo*.

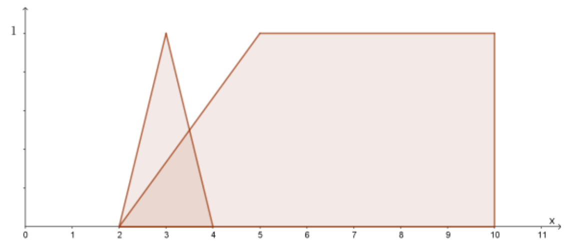

Figura 3. Codifica di *circa 3* e di *grande* mediante un triangolo e un trapezio.

La quarta scelta, quella che corrisponde alla verbalizzazione finale, viene di regola evitata, non servendo a nulla per costruire robot e condizionatori intelligenti, e rimane soltanto implicita. Nella Figura 3 abbiamo evidenziato $f(x) \wedge \varphi(x)$, mentre nella Figura 4 compaiono sia $\psi(y)$ che $g(y)$.

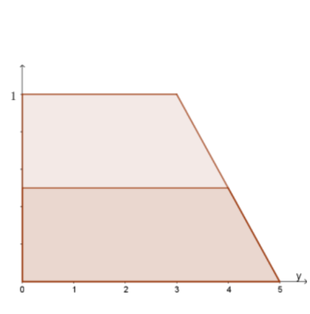

Figura 4. Codifica di *piccolo* e di *grungo*.

3. FUZZIFICATION E DEFUZZIFICATION

Il *modus ponens fuzzy* è usato nel controllo *fuzzy* e nel *soft computing* (nella *computazione flessibile* che preferisce accettare un po' di imprecisione piuttosto che perdersi in calcoli esatti ma interminabili), e ha dunque un notevolissimo impatto applicativo. Nel seguito considereremo l'esempio di due variabili linguistiche, velocità V e decelerazione D, quest'ultima misurata in chilometri al secondo per secondo, nell'istante in cui un sensore ha appena segnalato un ostacolo davanti alla macchina guidata da un robot: va immediatamente presa una decisione.

Le due variabili V e D possono essere descritte da attributi linguistici non necessariamente *fuzzy*, vaghi, ma anche *crisp*, nitidi, e nel nostro caso l'attrezzatura della macchina è in grado di misurare con perfetta precisione la velocità nell'istante corrente.

Premessa: la velocità è esattamente x
Regola: se la velocità è alta la decelerazione è brusca
Conclusione: dunque la decelerazione è?

Per mettere in moto l'armamentario della sezione precedente, basterà vedere un numero nitido x come se fosse un numero fuzzy, "fuzzificarlo" come ormai si dice. Abbiamo già visto numeri triangolari del tipo (a, x, b): ebbene, un numero nitido è appunto un numero *fuzzy* triangolare "limite" (x, x, x) la cui funzione di densità $f(x)$ vale 1 in x e vale 0 altrove.

Per passare a numeri concreti, sia la velocità x pari a 100 km all'ora; ipotizzando che per ragioni tecniche la macchina non possa superare i 220 km all'ora, codifichiamo l'attributo "grande" riferito alla velocità con il trapezio $(80, 140, 220, 220)$, che descrive la funzione $\varphi(x)$. Attenzione: il minimo tra $f(x)$ e $\varphi(x)$ è sempre $f(x) = 0$ tranne per $x = 100$ dove risulta essere $[f(100) \wedge \varphi(100)] = [1 \wedge \varphi(100)] = \varphi(100) = \frac{1}{3}$. La costante di taglio h vale allora $\frac{1}{3}$.

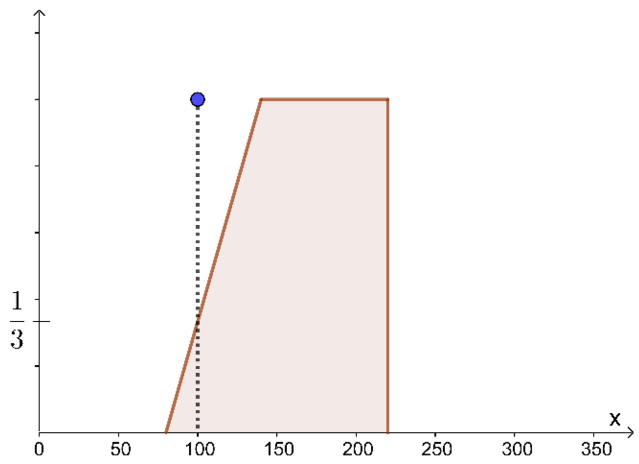

Figura 5. Codifica di *esattamente 3* e di velocità *grande.*

Se decidiamo di codificare decelerazioni brusche misurate in metri al secondo per secondo con il trapezio $(4, 20, 50, 50)$, che descrive la funzione $\psi(y)$ dando per scontati limiti tecnici che non sono quelli di una Ferrari in gara, dopo il taglio ad altezza h la decelerazione $g(y)$, senza sforzarci a inventare neologismi tipo di "grungo", è come in Figura 6.

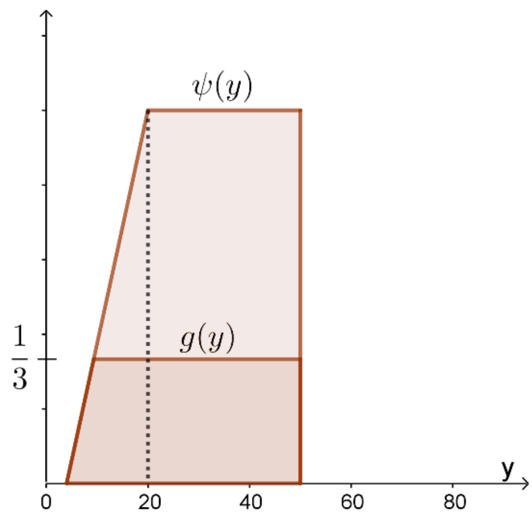

Figura 6. La conclusione *g(y)* in attesa della messa a fuoco.

Se la fuzzificazione operata nella premessa è ovvia o perfino deludente, più problematica e operativamente decisiva è l'operazione inversa, la defuzzificazione, ossia la focalizzazione, la messa a fuoco. D'accordo, la decelerazione è descritta dalla funzione $g(y)$, ma come deve decelerare il povero robot cui abbiamo affidato la guida della macchina, visto che c'è un intervallo di valori possibili di decelerazione che sono tutti veri con diverso grado di verità, quale più vero e quale meno vero: come passare a un unico e singolo valore, quello della decelerazione che il robot concretamente e nitidamente effettuerà?

L'idea cui si ricorre di solito è mutuata dalle cosiddette *mode* del calcolo della probabilità, cioè dai valori y cui compete probabilità massima ossia valore logico massimo $g(y)$ (grado di verità massimo) nel nostro caso.

Nel controllo *fuzzy*, quello alla base della costruzione di macchine e robot intelligenti, nella totalità dei casi reali (siamo nel mondo del *soft computing* dove le complicazioni sono aborrite) si giunge nella conclusione del sillogismo a una funzione $g(y)$ che o è "unimodale" (c'è un *unico* valore y^* che massimizza la funzione, la moda appunto), oppure i valori modali costituiscono un intervallo di cui si sceglie il centro y^* (*centro dei massimi*) che sta esattamente a metà strada dell'intervallo. Negli esempi delle Figure 4 e 6 i valori nitidi conseguenti alla defuzzificazione sono 2 che rende nitido (defuzzifica) l'attributo di "grungo", mentre la decelerazione che il robot dovrà selezionare sarà di 35 metri al secondo per secondo.

4. Conclusione

Grazie alla logica *fuzzy* e al *modus ponens,* che ne è il motore inferenziale fondamentale, è possibile costruire prodotti industriali che vanno dagli elettrodomestici intelligenti al controllo di complesse reti metropolitane e ferroviarie. Chi legge troverà troppo arbitrario e persino opinabile quanto precede, ma il successo applicativo del controllo *fuzzy* è un inoppugnabile dato di fatto!

Ripensando all'esergo di d'Alembert, che risale a un'epoca in cui l'analisi matematica già mieteva straordinari successi ma appariva ancora fondata su basi incerte e talora poco convincenti: *procedete e la fede verrà.*

BIBLIOGRAFIA

VAN ORMAN QUINE W.
1961, *Manuale di logica*, Feltrinelli.

KOSKO B.
1995, *Il fuzzy pensiero*, Baldini & Castaldi.

SGARRO A., FRANZOI L.
2016, «Logica e illogica delle lingue naturali: il caso della *consecutio temporum*», *QuaderniCIRD,* 13, pp. 26-38, scaricabile dall'indirizzo web: <http://hdl.handle.net/10077/13814>.

SITI WEB

LELAND STANFORD JUNIOR UNIVERSITY
Logic: firs-order modus ponens,
<https://stanford-cs221.github.io/autumn2020-extra/modules/logic/first-order-modus-ponens.pdf>, sito consultato il 12.1.2021.

SGARRO A., FRANZOI L.
SOFT COMPUTING e LOGICA FUZZY. Elaborazione degli stati di conoscenza incompleti: incertezza e vaghezza. Probabilità eterodosse, logiche multivalenti. Calcolo flessibile e aritmetica sfocata,
<https://www.dmi.units.it/~sgarro/livre_flou.pdf>, sito consultato il 12.1.2021.