APhEx 20, 2019 (ed. Vera Tripodi)
Ricevuto il: 2/04/2019
Accettato il: 15/07/2019
Redattore: Francesca Ervas & Paolo Labinaz

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI FILOSOFIA
NETWORK
N° 20, 2019

T E M I

Ragionamento in condizioni di incertezza

Margherita Benzi[1]

Il termine "incertezza" assume diversi significati a seconda del contesto scientifico in cui ricorre. Questo saggio si concentra sui diversi modi di concepire e formalizzare il ragionamento con premesse incomplete, vaghe approssimative o tali per cui il legame tra premesse e conclusioni è più debole che nella logica classica. La maggior parte dei metodi qui trattati trova origine, o ha raggiunto piena valorizzazione, a partire dalla ricerca degli anni ottanta in intelligenza artificiale. Il dibattito sul ragionamento incerto ha coinvolto sin dagli inizi contributi provenienti non solo dalla

[1] Desidero ringraziare per i preziosi commenti i revisori anonimi di *APhEx.* La ricerca è stata sostenuta con i fondi dell'Università del Piemonte Orientale.

logica e dall'informatica, ma anche dalla filosofia, dalla matematica e dalla psicologia, e si è esteso ad altri settori, quali quello dell'argomentazione giuridica o della diagnosi medica, provvedendo una adeguata cassetta degli attrezzi per i diversi contesti di applicazione.

Negli ultimi decenni l'espressione "intelligenza artificiale" ha aumentato il proprio carico di ambiguità, dato che viene utilizzata, specialmente dai media, come sinonimo di machine learning, e/o di Big Data. Esplorare le conseguenze della "nuova" IA sulla nozione generale di incertezza richiederebbe un contributo a parte. Se qui ci concentriamo sugli approcci classici è perché: a) non pensiamo che siano obsoleti: di fatto sono tuttora in uso e si rivelano utili anche per capire che cosa sia realmente la 'nuova' IA; b) hanno contribuito a chiarire, o a vedere da nuove prospettive, concetti che in precedenza erano spesso assimilati tra loro - come vaghezza, incertezza, probabilità, imprecisione, approssimazione, default,…, prestandosi a un proficuo scambio di idee tra informatici, logici e filosofi.

INDICE

1. Probabilità e rappresentazione della conoscenza

1.1 Le due anime della probabilità

Secondo una nota metafora di I. Hacking (1975), la storia della probabilità moderna può essere vista come un ciclo di reincarnazioni di due anime, di volta in volta ricomparse sotto diverse spoglie: da un lato, il concetto di *grado di certezza* (o di *credenza*, o di *fiducia*) che è ragionevole assegnare all'ipotesi che afferma il verificarsi di un certo evento; dall'altro, il concetto di *frequenza relativa* con cui un evento si presenta in una classe di riferimento. Una consolidata tradizione filosofica identifica convenzionalmente nel 24 agosto 1654, data di una lettera inviata da Blaise Pascal a Pierre de Fermat, la data di nascita della probabilità moderna. Vi è un certo accordo tra gli storici sull'esistenza di una "svolta probabilistica" iniziata intorno al 1660 e destinata a conquistare il pensiero scientifico moderno e contemporaneo. Vi è accordo, inoltre, nel collocare in tale periodo la nascita del *calcolo delle probabilità*, ovvero di un insieme di principi, che identificano, pur con variazioni, la teoria matematica della probabilità. Vi è accordo, infine, sul fatto che ad uno stesso insieme di principi di calcolo si accompagni una pluralità di concetti e di interpretazioni.

Restano comunque dei dubbi sulle radici storiche della *probabilità* nel pensiero antico e medievale: ci si chiede se il concetto empirico di *ciò che accade per lo più* fosse realmente separato da quello di "opinato dai più" o "dai più saggi" e ci si interroga su quanto i concetti proto-frequentisti dell'antichità e del Medio Evo fossero sovrapponibili alle nozioni proprie del frequentismo contemporaneo. Controverso è inoltre il ruolo svolto dallo sviluppo rinascimentale del calcolo combinatorio nella 'rivoluzione' probabilistica seicentesca (Howson 1978, Romizi 2009, Schuessler 2014).

Quali che siano le radici storiche della probabilità, è un fatto che sin dai primi anni dell'Ottocento l'ambiguità della nozione di probabilità è stata avvertita con intensità crescente, fino a sfociare, nei primi decenni del XX secolo, in una vera e propria "crisi dei fondamenti". Si è a lungo sostenuto che tale crisi sia stata superata grazie al sistema assiomatico proposto da A. N. Kolmogorov nel 1933[2], che sanciva una nitida distinzione tra aspetti

[2] Gli assiomi di Kolmogorov possono avere diverse versioni. Qui assumiamo che S sia un insieme finito di proposizioni A, B, C, ... contenga almeno una tautologia T e sia chiuso rispetto alla negazione, congiunzione e disgiunzione. Una funzione di probabilità P su S è

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI FILOSOFIA
NETWORK

sintattici (il calcolo) e aspetti semantici (le interpretazioni del calcolo). In realtà vi sono diverse assiomatizzazioni della probabilità, e inoltre, come osserva Hájek (2011) alcune delle più diffuse interpretazioni della probabilità non soddisfano tutti gli assiomi di Kolmogorov. Cionondimeno, le diverse definizioni associate a insiemi di assiomi della probabilità mantengono il diritto di chiamarsi "interpretazioni *della probabilità*".

Rimandando, per una esposizione più approfondita, alle numerose introduzioni alla filosofia della probabilità uscite in questi anni (e in particolare a Gillies (2000), Galavotti (2005), Mellor (2005), Hájek (2011) e all'antologia di Eagle (2010), ricordiamo brevemente le interpretazioni più note della probabilità.

Secondo l'*interpretazione classica*, presentata in forma sistematica agli inizi dell'Ottocento da Laplace, la probabilità di un evento è data dal rapporto delle possibilità favorevoli ad un certo evento e il numero delle possibilità totali. Alla base di questa definizione vi è l'idea che, in assenza di evidenza sugli esiti possibili di un fenomeno casuale, essi vadano considerati *ugualmente* possibili (*Principio di indifferenza)*. Nell'interpretazione *logicista*, proposta da Johnson (1921), Keynes (1921), e da Carnap (1950), la probabilità è definita come una relazione di implicazione parziale tra proposizioni. Tale relazione esprime il grado di sostegno che un insieme di dati probatori *E* assegna ad una ipotesi *H*. Di un certo rilievo per gli argomenti che affronteremo è il fatto che Carnap poneva la nozione nozione probabilità logica a fondamento di una teoria dell'inferenza induttiva destinata, nelle sue intenzioni, ad affiancare la logica deduttiva.

Secondo l'interpretazione nota come *soggettivismo* (o personalismo, o *bayesianesimo* soggettivista) la probabilità è intesa come grado di credenza in una certa ipotesi. Poiché l'assegnazione iniziale dei valori di probabilità rispecchia la "forza" della credenza individuale, non vi sono vincoli stretti sulla sua determinazione. Si richiede però che l'agente epistemico sia

---

una funzione che assegna numeri reali non negativi alle proposizioni di *S* in modo che siano soddisfatte le tre condizioni seguenti che corrispondono ai tre assiomi di Kolmogorov:

(A1) $P(A) \geq 0$ per ogni *A* di *S*;
(A2) $P(T) = 1$;
(A3) Se *A* e *B* appartengono a *S* e sono tali che $A \vdash \neg B$ e $B \vdash \neg A$, allora: $P(A \vee B) = P(A) + P(B)$ (*rincipio delle probabilità totali*).

Le probabilità $P(A)$, $P(B)$, ... sono dette *probabilità assolute*. La *probabilità condizionata* (l'espressione "$P(A|B)$" si legge "la probabilità di *A* dato che sappiamo che *B* è vera") descrive la probabilità assegnata ad una proposizione sulla base di una proposizione che descrive un certo corpo di conoscenze ed è definita nel modo seguente: Se *A* e *B* appartengono a *S* e $P(B)$ è diversa da 0, allora $P(A|B) = P(A \wedge B) / P(B)$.

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI FILOSOFIA
NETWORK

razionale, ovvero che il suo sistema di credenze sia coerente in senso probabilistico (un sistema di credenze è coerente se e solo se soddisfa gli assiomi del calcolo delle probabilità). Gli autori che più hanno contribuito alla sistematizzazione teorica di questa interpretazione sono stati F.P. Ramsey e B. de Finetti.

Mentre le interpretazioni che abbiamo fin qui ricordato si configurano come analisi di una nozione epistemica della probabilità, l'interpretazione *frequentista* (Reichenbach 1949, von Mises 1957) e quella *propensionista* inquadrano la probabilità come una caratteristica del mondo esterno, che caratterizza i fenomeni casuali. Nell'interpretazione frequentista la probabilità di un evento è identificata con la proporzione dei casi in cui l'evento occorre all'interno di una appropriata classe di riferimento. A seconda delle versioni del frequentismo considerate, quest'ultima può essere finita o infinita (cfr Hájek 2011). Nelle interpretazioni propensioniste la probabilità è concepita come la disposizione, o la tendenza, di un dispositivo casuale (una qualsiasi situazione fisica che può produrre una pluralità di esiti non prevedibili anticipatamente) a generare un certo risultato, o una certa frequenza. D. Gillies descrive due tipi di interpretazione propensionista: quella del lungo periodo e quella del caso singolo. Una teoria propensionista del lungo periodo è una teoria nella quale le propensità sono associate a condizioni ripetibili, e sono considerate come propensità a produrre, in una lunga serie di ripetizioni di queste condizioni, frequenze che sono approssimativamente uguali alle probabilità. Una teoria propensionista del caso singolo è una teoria nella quale le propensità sono considerate come propensità a produrre un particolare risultato in un'occasione specifica (Gillies 2000, 126).

Una prima esposizione dell'interpretazione propensionista si deve al grande filosofo e logico Americano C.S. Peirce (si veda Miller 1994). Versioni più recenti annoverano tra i contributori Popper (1957) e Gillies (2000) per quanto riguarda le teorie del lungo periodo, Fetzer (1974) e Miller (1994) per quanto riguarda le teorie del caso singolo.

Che cosa accomuna le analisi della probabilità sottese dalle interpretazioni sin qui ricordate? In primo luogo, il fatto di fornire una *misura quantitativa dell'incertezza*; in secondo luogo, una serie di regole di calcolo, corrispondenti, anche se con variazioni, agli assiomi di Kolmogorov; in altri termini, le diverse interpretazioni condividono una sintassi. Inoltre, tutte le interpretazioni storiche della probabilità che abbiamo ricordato si prestano, in misura maggiore o minore, ad essere applicate sia alle scienze della natura, sia alle scienze sociali. Cosa altrettanto importante, la "rivoluzione probabilistica" che in età moderna ha

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI FILOSOFIA
NETWORK

completamente permeato il pensiero scientifico è divenuta un pilastro della razionalità occidentale.

Questo complesso di motivi deponeva a favore di un ingresso trionfale della probabilità in IA, ed in particolare nei settori della rappresentazione della conoscenza e del ragionamento. Tuttavia le cose andarono in maniera diversa: prima che la probabilità stabilisse anche all'interno dell'IA un ruolo predominante, la sua adeguatezza, e perfino la sua applicabilità furono oggetto di un acceso dibattito. Prendendosi qualche libertà, si può parlare di una nuova, e più ampia, crisi dei fondamenti, che coinvolgeva questa volta la stessa sintassi della probabilità. Tale crisi ha favorito una comprensione più ampia dell'incertezza e dunque vale la pena di ripercorrerne i momenti principali. Prima di tutto, tuttavia, conviene preparare il terreno illustrandone il contesto teorico.

1.2 Ragionamento, logica e intelligenza artificiale

In poco più di sessant'anni di vita l'intelligenza artificiale ha già subito una serie di mutamenti di prospettiva – se non proprio di paradigma. Tali mutamenti sono stati enfatizzati nella rappresentazione mediatica e nell'immaginario collettivo, pronti, di volta in volta, a decretare le magnifiche sorti dell'IA, la sua estinzione, e una resurrezione foriera di catastrofi per l'umanità. Questi "cicli" – ben noti agli addetti ai lavori, che parlano di un'alternarsi di *winters* e *springs* - sono legati a successi e insuccessi reali di particolari programmi di ricerca e applicazioni e non coinvolgono la sopravvivenza della disciplina nel suo complesso. E' però vero che attualmente al suo interno convivono due approcci profondamente diversi, legati rispettivamente all'IA simbolica e al *Machine Learning* (e la sua sottoparte più problematica, cioè il *Deep Learning*). Non è compito di questo saggio discutere le relazioni tra i due approcci, ma vale la pena di ricordare che essi sono relativamente connessi a diversi tipi di attività in senso lato "intelligente", quella della capacità inferenziale per l'AI simbolica e quella dell'apprendimento di schemi per il *Machine Learning*. Qui di seguito ci occuperemo soltanto della prima.

Nello storico convegno tenutosi al Dartmouth College di Hanover nel 1956, J.J. McCarthy formulò la congettura che «ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza può in linea di principio essere descritto in maniera tanto precisa da permettere la costruzione di una macchina che lo simuli». Vale la pena di ricordare anche la definizione classica dell'IA di Genesereth e Nilsson (1987, 1), per i quali l’IA è lo «studio del comportamento intelligente, il cui scopo è formulare una teoria dell'intelligenza che renda conto del comportamento di entità intelligenti esistenti in natura e che orienti la costruzione di entità artificiali capaci di comportamento intelligente». Queste definizioni sono abbastanza generali da accogliere le diverse tendenze dell’IA contemporanea. Le cose stanno diversamente se ci concentriamo invece sull’intelligenza come capacità di risolvere problemi attraverso procedure computazionali effettive, dove la computazione è intesa come manipolazione formale di simboli (non interpretati) mediante l’applicazione di regole (Boden 1990, 4)[3].

Qui ci occuperemo soprattutto della IA simbolica, detta anche “Good old fashioned artificial intelligence” (GOFAI), che si occupa di sistemi *knowledge-based,* cioé programmi (o insiemi di programmi) in cui le conoscenze del sistema sono espresse in maniera esplicita, dichiarativa e codificate in forma simbolica.

Questa ipotesi di rappresentazione della conoscenza comporta l’idea che un sistema *knowledge-based* debba essere caratterizzato dai seguenti elementi:

(1) possesso di un linguaggio simbolico interpretabile come un insieme di proposizioni relative a un certo dominio;

(2) possesso di un apparato formale che gli consenta di svolgere inferenze automatiche;

(3) capacità di produrre, servendosi di (1) e (2), risultati che possano essere interpretati come “nuove” proposizioni relative al dominio.

Tali requisiti consentono di evidenziare l’analogia strutturale che sussiste tra i sistemi per la rappresentazione della conoscenza e il ragionamento presentati dalla corrente logicista dell’IA e i tradizionali sistemi logico-formali. I primi proponenti dell’IA simbolica dimostravano notevole dimestichezza con la letteratura corrente in campo logico. E’ tuttavia interessante notare che l’ambito di provenienza dei lavori citati non era tanto la logica matematica (o metamatematica), bensì la *logica filosofica,* come spiega Thomason (1991, 2018). Il motivo va da ricercarsi in un’ampia sovrapposizione tra le tematiche dei suoi settori.

McCarthy, iniziatore di questo approccio, riteneva la logica uno strumento per comprendere i problemi prima di risolverli (come osserva Thomason, questa apparente ovvietà non va data per scontata nell’IA odierna). In seno all’IA, tuttavia, alla logica – o meglio, al suo uso per la formalizzazione del ragionamento di seno comune - venivano richieste

[3] Su questi temi si veda anche il profilo *Alan Turing* su *APhEx* (Frixione e Numerico 2013).

caratteristiche peculiari di flessibilità, rivedibilità e versatilità, intesa come applicabilità a problemi di diversi domini.

Genesereth e Nilsson (1987, 77) osservano che in IA è più appropriato parlare di "insiemi di *credenze*", che non di "conoscenze", per indicare l'insieme delle informazioni che un sistema intelligente possiede circa il suo ambiente: un agente epistemico non è quasi mai certo che tali informazioni siano vere. Il carattere provvisorio della verità attribuita alle premesse dipende dal fatto che nessun sistema di rappresentazione della conoscenza può aspirare a fornire al sistema una descrizione del mondo che sia in qualche senso *completa*. L'assunzione che un certo dominio sia completamente specificato può valere in logica, dove è accettabile un alto livello di astrazione e di idealizzazione, ma sistemi costruiti per risolvere inferenze di senso comune possono contare soltanto su assunzioni sul mondo altamente provvisorie, e passibili di revisione qualora sopraggiungano nuove informazioni.

In poche parole, la formalizzazione del ragionamento di senso comune deve prevedere come trattare le eccezioni, elaborando meccanismi inferenziali che consentano di minimizzare i danni, cioè le conclusioni erronee, che possono derivare dall'esistenza di eccezioni non dichiarate, senza per questo rinunciare a trarre conclusioni dotate di una certa affidabilità. Una volta accertate le insufficienze del calcolo proposizionale e del calcolo dei predicati del primo ordine, la ricerca di impostazione logicista – appunto la GOFAI - ha studiato le potenzialità di calcoli logici alternativi, alcuni dei quali sono stati sviluppati in ambito filosofico, che modellano aspetti particolari del ragionamento di senso comune.

1.3 Probabilità e senso comune

A prima vista, la probabilità appare l'approccio ideale per esprimere il carattere incerto delle nostre credenze e la loro rivedibilità quando otteniamo nuove informazioni: la regola di Bayes ci permette appunto di aggiornare le probabilità di una certa conclusione alla luce di nuove evidenze. Quali problemi, allora, ne resero difficile l'adozione nei primi tempi dell'IA, esponendola agli attacchi dei logicisti? Innanzitutto, i due motivi denunciati da un famoso articolo di McCarthy e Hayes (1969): (i) la (presunta) distanza dal senso comune e (ii) l'inadeguatezza computazionale. Un terzo motivo è dato dal fatto che la probabilità non esauriva *tutti* i sensi dell'incertezza. Affrontiamo qui il primo problema, per occuparci in seguito degli altri.

Nella concezione della razionalità di stampo illuminista la probabilità era associata al senso comune, per lo meno a quello degli *hommes éclairés*. Agli inizi dell'Ottocento, Pierre Simon de Laplace chiamava il "buon senso" a sostegno delle stesse dimostrazioni matematiche, fino ad affermare: «la teoria della probabilità altro non è in fondo che il buon senso ridotto al calcolo; essa fa valutare con esattezza ciò che gli intelletti acuti avvertono per una sorta di istinto, senza che se ne possano rendere conto». (Laplace 1967, 404)

Il buon senso laplaciano si configura come intuizione illuminata dalla ragione. Più di un secolo dopo, de Finetti (1931) assegnava all'intuizione di senso comune un ruolo addirittura prioritario rispetto alle definizioni date dai matematici:

> [...] nel modo che seguiremo, il problema avrà uno svolgimento perfettamente intonato, una risposta precisa ed esauriente, perché si opera direttamente sul grado psicologico di affidamento di un certo individuo rispetto a una certa proposizione. E' appunto quel grado soggettivo di affidamento che nel linguaggio comune si designa col nome di probabilità, e la mia opinione è proprio questa: che il concetto espresso dal linguaggio ordinario abbia, una volta tanto, un valore assolutamente superiore a quello dei matematici, che da secoli si affannano inutilmente per vedervi un significato che non esiste (de Finetti 1931; ed. 1993, 34).

L'idea che il ragionamento probabilistico corretto non sia altro che ragionamento di senso comune debitamente coltivato comincia ad incrinarsi nella seconda metà degli anni '60, quando W. Edwards, uno psicologo impegnato nel confronto tra comportamenti decisionali effettivi e modelli normativi della decisione, pose il quesito se lo *human information processing* sia effettivamente bayesiano. Dopo una serie di esperimenti empirici volti a testare questa ipotesi, Edwards concludeva che, nonostante alcuni evidenti problemi legati alla difficoltà di calcolo, gli umani possono in fondo essere considerati dei *natural-born bayesians*. Pochi anni più tardi, tuttavia, iniziava la pubblicazione di una serie di lavori, dovuti principalmente a D. Kahnemann e A. Tverski[4], destinati a rimettere in dubbio questa conclusione, e a generare un filone della psicologia cognitiva tuttora estremamente prolifico, denominato *heuristics and biases*. Tra i temi che caratterizzano tale filone, quelli che seguono appaiono particolarmente importanti per la nostra trattazione:

[4] A partire dal primo fondamentale saggio Tversky e Kahnemann (1974).

- la mente umana sembra naturalmente predisposta a commettere violazioni sistematiche delle regole della statistica e del calcolo delle probabilità
- da tali "errori sistematici" non sono immuni nemmeno persone dotate di un certo grado di istruzione statistica e probabilistica
- ciò sembra far pensare che esistano *illusioni cognitive* paragonabili alle illusioni percettive - il cui caso paradigmatico è rappresentato dalle note frecce di Müller-Lyer – e tali da viziare alcuno aspetti del ragionamento
- le osservazioni precedenti sostengono in qualche misura l'ipotesi che nel ragionamento di senso comune si ricorra ad euristiche che, pur essendo in contrasto con le norme della probabilità, si rivelano, a lungo andare, altamente economiche e generalmente efficaci.

La ricerca sulle *heuristics and biases*, di grandissimo rilievo per tutte le discipline interessate a definire la razionalità umana, ebbe notevoli conseguenze anche per l'IA. Impegnata nel compito di costruire sistemi che in qualche modo mimassero il ragionamento di senso comune, quest'ultima non poteva ignorare i risultati che sembravano suggerire una incolmabile distanza tra questo ed il ragionamento probabilistico.

2. Alternative alla probabilità

2.1 Indadeguatezza computazionale dei sistemi esperti

Per farsi un'idea delle difficoltà incontrate dalla probabilità in IA è bene ripercorrere la storia dei primi sistemi esperti. Un *sistema esperto* è un programma, o un insieme di programmi, progettato per risolvere problemi che sorgono in un particolare dominio in maniera paragonabile, o persino superiore, a come farebbe un esperto umano. Generalmente un sistema esperto è strutturato in alcune componenti fondamentali: la *base di conoscenze*, che comprende un insieme di assunzioni generali sul dominio considerato, il *motore inferenziale*, un apparato formale che consente di trarre inferenze a partire dalla base di conoscenze, un *insieme di dati specifici*, relativi ad un particolare problema che si intende risolvere, e un *interfaccia*, che consente al sistema di dialogare con l'utente. Una delle applicazioni privilegiate dei sistemi esperti è stato, sin dagli anni '50, il sostegno alla diagnosi medica.

In generale, il giudizio diagnostico era concepito come la determinazione, a partire da un certo insieme di dati che costituiscono l'*evidenza*, dell'*ipotesi* più accettabile come spiegazione dei dati. Il grado in cui una certa evidenza rende plausibile un'ipotesi può essere rappresentato come la probabilità condizionata $P(H_i|E_j)$, dove $\mathbf{E} = \{E_1, E_2, ..., E_m\}$ denota un insieme di possibili osservazioni (l'anamnesi del paziente, sintomi e risultati di analisi, …) e $\mathbf{H} = \{H_1, H_2, ..., H_n\}$ un insieme finito di ipotesi diagnostiche (per semplicità: malattie). Il problema della diagnosi diventa allora quello di trovare l'ipotesi con la probabilità più alta condizionatamente all'evidenza riscontrata.

Nei primi sistemi esperti, valori di $P(H_i|E_j)$ erano ottenuti mediante la già citata regola di Bayes. Questa consente di ricavare la probabilità a posteriori delle ipotesi, date le loro probabilità a priori $P(H_i)$ – ricavate dai dati epidemiologici - e le verosimiglianze $P(E_j|H_i)$. Poiché un paziente può avere più di una malattia, il numero delle possibili diagnosi (combinazioni di malattie) è 2^n; dunque, per ottenere la distribuzione completa della probabilità a priori si devono specificare $2^n - 1$ parametri indipendenti. La specificazione delle verosimiglianze richiede, per *m* osservazioni e *n* ipotesi, la specificazione di $2^n(2^m - 1)$ parametri indipendenti per ciascuna diagnosi [5]. Questo tipo di approccio si rivela pressoché impraticabile quando le ipotesi e i sintomi sono più di due o tre. La complessità computazionale del problema poteva essere ridotta mediante due assunzioni, l'ipotesi che ciascun paziente non avesse più di una malattia e l'ipotesi che, data una certa diagnosi H_i, le diverse osservazioni riscontrate fossero indipendenti tra loro.

Nonostante i buoni risultati ottenuti con questi sistemi (ad esempio, percentuali di diagnosi corrette superiori a quelle ottenute da esperti umani), essi giocarono a sfavore, piuttosto che a favore, dell'applicazione della probabilità in IA. Di fatto, contribuirono a evidenziare due importanti difficoltà legate alla rappresentazione probabilistica dell'incertezza: la prima è che, a meno di non introdurre assunzioni semplificatrici non sempre legittime, il rischio di esplosioni combinatorie è sempre incombente. La seconda è legata alla peculiarità dell'IA simbolica, che richiedeva che i sistemi esperti fossero in qualche modo vicini al modo effettivo di ragionare degli esperti umani, e che potessero facilmente dialogare con essi. La probabilità, con la richiesta di specificare in maniera coerente un gran numero di valori puntuali, non sembrava soddisfare quest'ultimo requisito.

---

[5] Si vedano Henrion, Breese e Horvitz (1991, 69).

2.3. Modi dell'incertezza

Come accennato sopra, il dominio del ragionamento in condizioni di incertezza non era occupato esclusivamente dalla probabilità classica. Si pensi, ad esempio, alle probabilità "baconiane", affiancate da J.L. Cohen (1977) a quelle "pascaliane". Quale prospettiva teorica adottare di fronte a tali alternative? La scelta si riduceva fondamentalmente a quella tra due opzioni: il pieno riconoscimento della dignità teorica di concetti di probabilità regolate da metodi *sintattici alternativi* a quello standard, o la riaffermazione del potere normativo della probabilità come *canone di razionalità*, unita alla convinzione che i limiti apparenti della probabilità possano essere superati mediante una rappresentazione più *user-friendly*.

La comunità IA ha sempre considerato con attenzione il cosiddetto "Teorema di Cox" Nel 1946 fisico americano R.T. Cox elencava una serie di condizioni desiderabili che qualunque misura dei gradi di credenza avrebbe dovuto soddisfare, e dimostrava che qualunque misura della credenza che soddisfaceva tali condizioni era isomorfa al calcolo delle probabilità. Horvitz, Heckerman e Langlotz (1986, 211) presentavano le condizioni di Cox (1946) integrate da quelle aggiunte successivamente da M. Tribus e E.T. Jaines. Esse erano le seguenti: 1. *Chiarezza:* le proposizioni alle quali si assegna un grado di credenza dovrebbero essere ben definite (vale a dire, vere o false): 2. *Continuità scalare*: un singolo numero reale è sufficiente per rappresentare un grado di credenza, 3. *Completezza*: è possibile assegnare un grado di credenza a qualunque proposizione; *4. Dipendenza dal contesto:* la credenza in una proposizione può dipendere dalla credenza in altre proposizioni; 5. *Condizionamento ipotetico*: il grado di credenza nelle proposizioni *A* e *B* deve essere una funzione del grado di credenza in una delle due (ad es. *B*) e del grado credenza nell'altra sapendo che la prima è vera; 6. *Complementarietà*: la credenza nella negazione di una proposizione è una funzione monotonicamente decrescente della credenza della proposizione stessa, e *7. Consistenza*: proposizioni che hanno lo stesso valore di verità dovrebbero ricevere lo stesso grado di credenza.

Il Teorema di Cox è stato visto da alcuni come la dimostrazione che il calcolo delle probabilità era l'unico approccio razionale all'incertezza in un contesto nel quale sembrava ripresentarsi a livello sintattico il problema che nei primi decenni del Novecento si era posto a livello semantico: qual è la formalizzazione giusta? E' possibile, senza essere irrazionali, accettare formalizzazioni dell'incertezza diverse dal calcolo delle probabilità?

Aggravava la situazione una notevole confusione terminologica[6], oggi in parte eliminata grazie al fatto che è invalso l'uso di collegare il termine "incertezza" ai metodi probabilistici, ed altre denominazioni a quelli non probabilistici. Ma più che le questioni terminologiche è interessante vedere quali motivazioni, non necessariamente irrazionali, potevano portare all'abbandono di alcuni dei requisiti *1. - 7.* elencati sopra.

Per affrontare il tema dell'incertezza in AI, bisogna ricordare che fino agli anni ottanta il termine «sembrava essere usato ogniqualvolta non sembrava possibile ragionare in termini di stretta implicazione logica» (Spiegelhalter 1986). Sotto la denominazione "incertezza" rientravano diversi problemi, con cause diverse. Tentiamone una rudimentale tassonomia (si vedano Bathnagar e Kanal 1988 e Dubois e Prade 1988).

Supponiamo di interrogare una base di conoscenze che contenga tutte le informazioni di cui disponiamo. Il problema consiste nel valutare il grado in cui l'evidenza ci consente di ritenere credibile l'ipotesi. Tra le cause principali di una "non piena credenza" possiamo distinguere le seguenti.

1) *Incompletezza*. L'incertezza può derivare dall'incompletezza dell'informazione rilevante. Un tipico caso di incertezza da incompletezza è quella che riguarda i fenomeni aleatori: per esempio, i dati contenuti nella nostra base di conoscenze non ci consentono di specificare l'esito del lancio di una moneta prima che questa venga lanciata. I casi in cui tutta l'informazione rilevante è contenuta nella base di conoscenze, ma il suo recupero è troppo dispendioso o difficile possono essere considerati analoghi a quelli in cui l'informazione rilevante è incompleta.

2) *Approssimazione*. In questo caso la nostra base di dati fornisce tutte le risposte rilevanti, ma queste non sono accurate. Possiamo distinguere diversi casi, alcuni dei quali appaiono strettamente connessi con l'incompletezza, altri con la vaghezza. Qualora l'evidenza sia espressa in termini numerici si parla di informazione approssimata quando la base di conoscenze ci fornisce un intervallo di valori in luogo di un valore puntuale. Nel caso in cui l'evidenza sia espressa in termini non numerici, l'approssimazione può dipendere da un'intrinseca mancanza di precisione del linguaggio. In questo caso, può essere ricondotta alla vaghezza. Si parla di informazione approssimata anche nel caso delle *descrizioni prototipiche*, frequenti nel ragionamento di senso comune. In una descrizione prototipica, un concetto viene definito mediante la

[6] Si veda Benzi (1997, 23-24).

descrizione di un individuo esemplare. Ad esempio, un uccello può essere definito come “Un animale che vola, ha le piume, cinguetta, depone le uova, …” Quando si usa una descrizione prototipica, si ammette che vi possano essere individui che rientrano nel concetto in questione pur presentando differenze significative rispetto al prototipo. Una descrizione prototipica assume dunque il ruolo di una generalizzazione con eccezioni ammesse. Per evitare problemi di incertezza, una base di conoscenze che contenga descrizioni prototipiche dovrebbe includere un elenco completo di possibili eccezioni, cosa che può avvenire solo in casi molto particolari.

3) *Inaffidabilità*. La base di conoscenze può contenere evidenza precisa e completa, ma proveniente da fonti non affidabili. In questo casi si può pensare di associare a ciascun enunciato che riporti evidenza un peso determinato dal grado di affidabilità attribuito alla fonte.

4) *Ambiguità*. L’incertezza deriva in questo caso dal fatto che vi sono più entità che potrebbero costituire il possibile riferimento di un termine; ad esempio, la frase “Ada ha portato la bambina all’asilo” potrebbe creare problemi se vi sono diverse persone di nome “Ada”, o se qualcuna di loro ha più di una bambina, ecc.

5) *Vaghezza*. Un ulteriore tipo di incertezza è quello che deriva dall’intrinseca vaghezza del linguaggio. Termini quali “alto”, “giovane” e “mucchio”, sono intrinsecamente vaghi, nel senso che non è possibile indicare in maniera nitida, indipendente dal contesto, gli individui di cui tali concetti si predicano; si ammette inoltre che la loro predicabilità sia graduata, piuttosto che una questione “tutto o niente”.

Nelle prossime sezioni vedremo come questi problemi sono stati affrontati in IA, a partire dalla prima sfida all’analisi dell’incertezza senza l’uso della probabilità.

2.3 Fattori di certezza come alternativa alla probabilità

Verso la fine degli anni settanta cominciò ad affermarsi un approccio alternativo all’incertezza, consistente in una estensione non deterministica dei sistemi *basati su regole*. In questi sistemi la conoscenza è codificata mediante *regole di produzione*, della forma:

SE (premessa)

ALLORA (conclusione).

Per esempio, un sistema esperto per l'identificazione della famiglia di appartenenza di una pianta, conterrà informazioni del tipo:

SE (una pianta appartiene a una certa classe E le sue foglie hanno una certa forma)
ALLORA (la pianta appartiene a una certa famiglia).

L'uso delle regole dipende dal motore inferenziale. Ad esempio, un sistema può avere un motore inferenziale di tipo *backward-chaining* che, per ogni regola, indica di verificare l'antecedente sulla base dei dati ed eventualmente di altre regole SE-ALLORA. Quando è possibile dare risposte del tipo "vero/falso" alle domande rilevanti per il problema da risolvere e le premesse implicano logicamente la conclusione, le regole SE-ALLORA sono dette *deterministiche*; quando invece le premesse si limitano a conferire alla conclusione una maggiore o minore verosimiglianza, le regole sono *non deterministiche*. La strategia per trattare l'incertezza consiste nel descrivere l'evidenza come antecedente e l'ipotesi come conseguente di una regola SE-ALLORA e di assegnare alla regola un peso numerico che esprime la plausibilità della conclusione in base alle premesse. Questi pesi sono combinati mediante regole diverse da quelle della probabilità.

Il metodo dei fattori di certezza (CF) è stato uno dei primi formalismi sviluppati in questo settore; l'ambito di applicazione era un sistema esperto per la diagnosi di infezioni batteriche, MYCIN. In MYCIN la regole descrivono come la fiducia in una certa ipotesi cambia; per esempio, il fatto che un microrganismo sia Gram-positivo e con conformazione di crescita a catena aumenta la credenza dell'esperto nell'ipotesi che il micro-organismo in questione sia uno streptococco. Il fattore di certezza associato a tale regola è appunto il numero che esprime la misura dei cambiamenti indotti dall'acquisizione di nuova evidenza. Questo metodo di rappresentare l'incertezza ricordava in parte la teoria logicista di Carnap, dove la probabilità era intesa come il grado di conferma di un'ipotesi sulla base di un'evidenza. Tuttavia si differenziava dalla dalla teoria carnapiana in quanto si limitava ad esprimere il cambiamento complessivo della credenza apportato dall'evidenza e non richiedeva di indicare esplicitamente le probabilità a priori e a posteriori di ciascuna ipotesi [7]. In poche parole,

[7] Henrion, Breese e Horvitz (1991, 70).

MYCIN sembrava offrire un modello non probabilistico per misurare la forza dell'evidenza.

MYCIN ammetteva CF negativi per i casi nei quali l'evidenza era a sfavore dell'ipotesi e pertanto consentiva una rappresentazione distinta tra le ipotesi disconfermate (a causa di evidenze sfavorevoli) e quelle non confermate (a causa dell'assenza di evidenza: tale distinzione assecondava un'esigenza diffusa nel campo del ragionamento medico. Due ulteriori vantaggi del sistema erano la facilità dei calcoli richiesti e – in assenza di dati statistici - la possibilità di assegnare i CF a casi individuali, sulla base delle intuizioni degli esperti. Infine, MYCIN consentiva una attribuzione di valori numerici modulare, che evitava di tener conto delle relazioni tra le diverse regole e tra queste e le evidenze. Proprio la modularità, però, generava caratteristiche del sistema non desiderabili. Essa deriva infatti da una proprietà caratteristica della logica deduttiva, la proprietà di monotonicità, che vieta di modificare il valore di verità di una proposizione asserita come vera alla luce di altre proposizioni del sistema. Tuttavia il ragionamento incerto, come vedremo meglio più avanti, richiede la ritrattabilità delle nostre assegnazioni di valori di verità alla luce di nuove evidenze, e tale ritrattazione non si ottiene agevolmente con i CF. In particolare, il sistema rende complicata l'inferenza intercausale.

Per comprendere quest'ultimo problema, consideriamo la regola "SE in un certo punto del mare troviamo petrolio, ALLORA in quel punto vi è un giacimento petrolifero (con certezza *b*)". Se appuriamo la verità dell'antecedente, il nostro grado di certezza che in quel punto vi sia un giacimento dovrebbe aumentare; tale incremento sarebbe tuttavia destinato a venir meno se la nostra base di conoscenze contenesse l'informazione che nelle vicinanze è affondata una petroliera. E' evidente, quindi, che le relazioni rappresentate mediante CF possono essere fuorvianti se considerate separatamente, senza ispezionare l'intera base di conoscenze. Le regole locali potrebbero essere rese più affidabili specificando tutti i possibili controesempi, per esempio scrivendo "SE ((In un certo punto del mare troviamo petrolio) E (Non è affondata una petroliera) E (In quel punto non è stato sversato un carico di petrolio) E …), ALLORA (In quel punto vi è un giacimento petrolifero)", ma nella maggior parte dei casi l'obiettivo di formulare una lista esauriente dei possibili controesempi non sembra praticabile; anche se lo fosse, richiederebbe un notevole sforzo computazionale e la conseguente perdita dei vantaggi derivanti dalla modularità.

Degno di nota è il fatto che la prima formulazione di MYCIN dava origine a conseguenze non consistenti; una sua riformulazione successiva volta a rimuoverne le inconsistenze equivaleva ad un *rafforzamento*, o una *specializzazione*, del calcolo delle probabilità, ovvero ad una teoria che soddisfaceva i requisiti imposti da R.T. Cox con l'aggiunta di esplicita imposizione di una ulteriore regola relativa all'indipendenza tra elementi di evidenza[8].

2.4 La teoria dell'evidenza di Dempster-Shafer

La *teoria dell'evidenza*, o teoria di *Dempster-Shafer*, o *teoria delle funzioni di credenza*, ha avuto origine da una serie di ricerche condotte negli anni '60 da A.P. Dempster, e da una successiva rielaborazione da parte di G.R. Shafer.

La teoria delle funzioni di credenza, a differenza delle teorie classiche della probabilità, non produce una valutazione diretta della probabilità delle ipotesi, bensì una valutazione dell'attendibilità dell'evidenza relativa a tali ipotesi. Questo sembra renderla particolarmente interessante in quelle situazioni dove non sono disponibili conoscenze sufficienti per la formulazione di un modello probabilistico completo, ma si disponga di insiemi di osservazioni sparsi e indipendenti tra loro.

Shafer nega che il calcolo delle probabilità costituisca una teoria normativa della razionalità, nel senso che una assegnazione coerente di gradi di credenza debba forzatamente rispettare le leggi del calcolo delle probabilità. Il suo rifiuto è legato alla critica di due delle regole fondamentali del bayesianesimo: il principio di additività (che afferma che $P(\neg A) = 1 - P(A)$) e la regola di aggiornamento delle probabilità mediante condizionalizzazione alla nuova evidenza.

Per quanto riguarda il primo, egli ritiene che se si attribuisce alla proposizione A la probabilità $P(A) = p$, sia perfettamente legittimo assegnare alla sua negazione A una probabilità q diversa da $1 - p$. In particolare, il principio di additività va rifiutato nel caso in cui un certo agente epistemico dispone di informazioni che giustifichino l'attribuzione di una probabilità p all'ipotesi A e di nessuna informazione in merito all'ipotesi $\neg A$; in questo caso Shafer ritiene corretto dire che $\neg A$ riceve dall'evidenza un sostegno pari a zero, non a $1 - p$.

Per quanto riguarda la regola della seconda, Shafer osserva che la condizionalizzazione della probabilità di un'ipotesi H rispetto a un'evidenza E richiede la predisposizione di un modello probabilistico che già contenga ogni possibile E, cioè tutti gli eventuali elementi dell'evidenza di cui

[8] Heckerman (1986). Si veda anche Horvitz, Heckerman e Langlotz (1986).

possiamo venire in possesso. Questo equivale a prevedere in anticipo tutti i possibili futuri corsi dell'esperienza. In molti ambiti di applicazione - si pensi al ragionamento di senso comune - appare difficile costruire un modello esaustivo delle possibili evoluzioni della nostra conoscenza. Inoltre, la formulazione di un modello probabilistico completo richiede che a ciascun corpo di evidenza E sia assegnata una probabilità *a priori*, operazione altamente controversa. Alla condizionalizzazione bayesiana Shafer sostituisce una *regola di combinazione* che consente di aggiungere o eliminare liberamente dati osservativi nel corso dell'analisi di una funzione di credenza, ponendo come unica condizione che i diversi corpi di evidenza che vengono combinati tra loro siano indipendenti. Tale regola consente effettivamente di aggiornare i gradi di credenza evitando la specificazione a priori di un modello probabilistico completo.

La teoria di Dempster-Shafer prevede che a ciascuna proposizione A possano essere assegnate due misure di credenza, denominare rispettivamente *credenza* (Bel (A)) e *plausibilità* (Plaus (A)). La *plausibilità* di A è la misura del grado in cui l'evidenza non è contraria all'ipotesi ed è data dalla formula

$$\text{Plaus}(A) = 1 - \text{Bel}(\neg A).$$

L'intervallo [Bel (A), Plaus (A)] è detto *intervallo di credenza*. Esso fornisce una misura della informatività dei valori Bel ottenuti sulla base dell'evidenza: intuitivamente, quanto più piccolo è tale intervallo, tanto più conclusiva è da considerarsi l'evidenza.

La teoria viola le proprietà di continuità scalare e di complementarietà ricordate nel paragrafo precedente e poiché ammette l'esistenza di ipotesi ben definite alle quali non può essere assegnato un grado di credenza, prevede un indebolimento della proprietà di *completezza*. Tuttavia, essa non rappresenta una rottura drastica rispetto alla teoria classica della probabilità, della quale può essere considerata una generalizzazione.

2.5 La Probabilità vendicata: le reti bayesiane

Mentre la teoria di Dempster-Shafer ha costituito una alternativa alla probabilità classica, un altro filone di ricerca ha recuperato in modi nuovi anche la visione bayesiana della probabilità. A partire dalla seconda metà degli anni '80 l'approccio probabilistico noto col nome di *reti bayesiane* ha segnato il rientro a pieno titolo della probabilità nella rappresentazione della conoscenza e nel ragionamento in IA (Pearl 1988).

I sistemi probabilistici più “primitivi”, o “bayesiani ingenui”, riuscivano a limitare la complessità computazionale rinunciando alla garanzia della coerenza probabilistica o, in alternativa, ricorrendo ad assunzioni semplificatrici. Oltre a queste limitazioni, pesava su di essi l’idea che la probabilità fosse lontana dal senso comune e inapplicabile in assenza di dati statistici affidabili. La sfida che i sistemi probabilistici dovevano affrontare consisteva dunque nel rispettare la coerenza probabilistica, snellendo nel contempo gli oneri computazionali derivanti da un uso ortodosso della probabilità e fornendo calcoli il più possibile trasparenti, cioè comprensibili in ogni passaggio per utenti privi di un particolare addestramento matematico o statistico.

Un aspetto innovativo della teoria delle reti bayesiane è che essa si avvale, oltre che della probabilità classica, di elementi della teoria dei grafi. Una rete bayesiana è la rappresentazione di un insieme di variabili causali e le relazioni di dipendenza condizionale che tra queste sussistono mediante un grafo aciclico orientato nel quale i nodi rappresentano le variabili, e gli archi le relazioni di dipendenza. Per esempio, in una rete bayesiana per la diagnosi medica le variabili (che possono assumere, per semplicità, i valori “presente/assente”) rappresentano malattie e sintomi, e gli archi le relazioni che legano i sintomi alle malattie. Se un arco va dal nodo *A* al nodo *B*, si dice che *A* è “genitore” di *B*. Ad ogni nodo è associata una funzione di probabilità che calcola la probabilità della variabile rappresentata da un nodo sulla base di particolari valori di probabilità dei genitori.

Quali sono i fattori che hanno garantito il successo delle reti bayesiane? Oltre alla soddisfazione dei requisiti di Cox ricordati sopra, le reti bayesiane alleviano enormemente i problemi computazionali che tanto avevano pesato sui primi sistemi probabilistici; infatti, la rappresentazione delle relazioni di dipendenza consente un calcolo locale dei valori di probabilità, riducendo il numero dei parametri da considerare. Altrettanto importante, come rileva Pearl (1988) è il fatto che la rappresentazione grafica delle relazioni di dipendenza consente una visione immediata di quali fattori siano rilevanti per ciascun altro fattore, ed è pertanto più vicina al ragionamento di senso comune.

La teoria delle reti bayesiane riveste una notevole importanza, sia tecnica che filosofica, anche per avere dato origine ad una teoria formalizzata dell’inferenza causale. Attualmente la teoria del ragionamento causale basata su modelli grafici, che ha riscosso anche una notevole attenzione filosofica (Hitchcock 2001, Woodward 2003), si compone di due programmi principali, dedicati rispettivamente alla estrazione di relazioni causali a partire da basi statistiche di dati (Glymour, Scheines et. al. 2014) e l'altra che assegna alle relazioni causali priorità epistemologica rispetto a quelle probabilistiche, e intende rappresentare anche la cosiddetta "causalità singolare" (Pearl 2009), che riguarda eventi individuali che talvolta si discostano dalle regolarità statistiche.

3. Altri approcci: Logica fuzzy e logica non monotona

3.1 Incertezza, vaghezza e valori di verità

La teoria dei *fuzzy sets* [*insiemi vaghi*] è stata presentata per la prima volta nel 1965 da L.A. Zadeh, un ricercatore noto per i suoi contributi nel campo della teoria dei sistemi e dei controlli automatici. Una prima, grossolana, linea di demarcazione tra l'approccio *fuzzy* e i formalismi per la rappresentazione dell'incertezza che abbiamo fin qui presentato può essere tracciata considerando un particolare tipo di incertezza, che potremmo denominare *imprecisione linguistica*, o meglio, un particolare tipo di imprecisione, la *vaghezza*, intesa come assenza di confini nitidi per il significato di una parola (si pensi al termine "mucchio" e al noto paradosso del sorite).

La vaghezza è divenuta un tema molto discusso in filosofia a partire dagli anni '90. Qui non approfondiremo il tema filosofico della vaghezza[9] e ci rifaremo soprattutto a contributi in campo logico e informatico.

In logica classica è sempre possibile determinare con precisione se un oggetto è membro di un insieme che costituisce l'estensione di un certo predicato o non lo è; in altri termini, la *funzione di appartenenza* di un certo insieme può assumere esclusivamente i valori 0 o 1. Nel linguaggio ordinario, invece, la maggior parte dei termini non consente determinazioni così precise: in generale, un individuo di dieci anni è sicuramente giovane, un individuo di novanta sicuramente non lo è, ma per una larga parte di persone di età intermedie l'appartenenza all'insieme dei giovani sarà parziale.

È abbastanza facile vedere intuitivamente come la vaghezza non si limiti ad essere una proprietà di singoli termini del linguaggio comune, ma investa anche situazioni complesse e articolate. Supponiamo, con Machina (1976), che un individuo di nome Orazio deponga delle piantine di petunia

[9] Si vedano ad es. Paganini (2008), Gaio (2010), Moruzzi (2012).

in un aiuola del proprio giardino, senza curarsi di interrarle o concimarle e che alcune piantine sopravvivano: potremmo asserire che Orazio ha piantato le petunie? Se paragoniamo le azioni di Orazio con quelle che compie un giardiniere qualificato nel piantare le petunie, avremo qualche difficoltà a riconoscere come vera la proposizione "Orazio ha piantato le petunie"; d'altro canto, se sapessimo che Orazio ha gettato le petunie nel bidone della spazzatura, essa ci apparirebbe decisamente falsa. Potremmo dunque pensare che in questo caso la vaghezza delle proposizioni che descrivono le azioni di Orazio ci conduca a considerare valori di verità intermedi tra "assolutamente vero" e "assolutamente falso", che Zadeh chiama "gradi di verità".[10]

Fin qui abbiamo visto come la vaghezza possa essere messa in relazione da un lato con l'estensione dei predicati, dall'altro con il valore di verità delle proposizioni. Ora cercheremo di approfondire le differenze tra la vaghezza e le forme di incertezza trattate nei capitoli precedenti (d'ora in poi: *incertezza* tout-court). Una distinzione frequente nella letteratura – anche se, come vedremo, non del tutto corretta - afferma che l'incertezza può essere ricondotta alla *incompletezza della informazione*, mentre ciò non avviene nel caso della vaghezza. In un classico problema posto da una situazione aleatoria (per esempio, quando ci domandiamo se un certo lancio di un moneta produrrà il risultato "testa" o il risultato "croce"), siamo di fronte a un insieme di alternative definito con precisione, ma l'informazione di cui disponiamo non è sufficiente per sceglierne una. Per eliminare l'incertezza, ed assegnare un valore di verità alla proposizione "Il risultato del lancio è croce" abbiamo bisogno di ulteriore informazione, quale quella che ci dà il lancio effettivo della moneta. Viceversa, nel caso del nostro Orazio, possiamo immaginare di dare delle azioni da lui compiute nell'appoggiare le piantine sul terreno una descrizione estremamente meticolosa e particolareggiata, e tuttavia rimanere incerti sulla verità della frase "Orazio ha piantato le petunie".

L'incertezza riguarda dunque, enunciati o frasi linguisticamente precisi che non possiamo giudicare veri o falsi, poiché disponiamo di una informazione incompleta, cioè lacunosa o mal definita; la vaghezza riguarda enunciati che non possiamo giudicare (in senso assoluto) veri o falsi, anche se disponiamo di una informazione completa, poiché ci appaiono non del tutto veri, né del tutto falsi.

A questo punto si potrebbe pensare che i formalismi quali teoria della probabilità, teoria delle funzioni di credenza, ecc. trattino dell'incertezza,

[10] Si veda anche Machina (1976, 53).

mentre l'approccio *fuzzy* tratti (esclusivamente) della vaghezza. (S)fortunatamente, la situazione non è così semplice, e l'approccio *fuzzy* accoglie una pluralità di logiche elaborate per rendere conto delle relazioni tra ipotesi ed informazioni nelle quali uno dei due relata, o entrambi, sono di volta in volta vaghi, incerti, o *nitidi* (cioè non vaghi). Rimandando a Dubois e Prade (1980), Dubois, Nguyen e Prade (2000) per un'illustrazione più approfondita, presentiamo nel prossimo paragrafo uno dei primi formalismi basati sulla teoria degli insiemi *fuzzy*.

3.2 Teoria della possibilità e logica della possibilità

Teoria della possibilità. La teoria della possibilità è stata elaborata da L.A. Zadeh sul finire degli anni'70 (Zadeh 1978) come estensione della sua teoria degli insiemi vaghi (*fuzzy*), una teoria finalizzata alla rappresentazione dell'intrinseca vaghezza del linguaggio naturale. Mentre nella teoria degli insiemi classica l'appartenenza è una questione tutto-o-niente, nella teoria dei *fuzzy sets* l'appartenenza è graduata. L'appartenenza ad un insieme vago *F* è misurata dalla funzione di appartenenza μ_F. Se *u* è un elemento dell'universo del discorso U, $\mu_F(u)$ specifica il *grado di appartenenza* di *u* a *F* (intuitivamente, il *grado di compatibilità* di u *con* F). Il valore 1 denota la totale appartenenza e il valore 0 la totale non appartenenza. Per esempio, se l'universo del discorso è l'insieme dei possibili anni di età, è abbastanza naturale porre per U = 20 una compatibilità pari a 0.9 con l'insieme vago associato al predicato "Giovane" (in simboli: $\mu_{\text{Giovane}}(20) = 0.9$), per U = 40 un grado di appartenenza minore, e così via.

Zadeh traccia una connessione tra la nozione di appartenenza ad un insieme vago e quella di *possibilità*: semplificando, la possibilità che un certo individuo sia giovane coincide con il grado di appartenenza della sua età all'insieme vago associato al predicato "Giovane". Questa connessione è posta alla base della costruzione di *funzioni di distribuzione di possibilità* numericamente equivalenti alle funzioni di appartenenza.

La differenza tra una distribuzione di possibilità e una distribuzione di probabilità è illustrata in maniera intuitiva da un celebre esempio di Zadeh.

Consideriamo l'enunciato "Hans ha mangiato *u* uova a colazione", dove *u* è una variabile che può assumere i valori {1, 2, ... 8}.

Supponiamo che 'Poss(*u*)' denoti la distribuzione di possibilità associata a *u*: può grosso modo essere interpretata come il grado di facilità con cui Hans mangia *u* uova; con l'espressione 'Pr(*u*)' indichiamo invece la *probabilità* che Hans mangi *u* uova. Supponiamo che le due distribuzioni siano le seguenti:

Uova	1	2	3	4	5	6	7	8
Poss(*u*)	1	1	1	1	0.8	0.6	0.4	0.2
Pr(*u*)	0.1	0.8	0.1	0	0	0	0	0

Un alto grado di possibilità non comporta un alto grado di probabilità, ma un evento impossibile deve anche essere improbabile. Zadeh illustra la propria nozione di possibilità affermando che essa esprime una generica "facilità di realizzazione" o "facilità di conseguimento".

Logica della possibilità. La teoria della possibilità di Zadeh è stata utilizzata da Dubois e Prade (1980, 1988) come base per una teoria del ragionamento in condizioni di incertezza, la *logica della possibilità*. Il suo ambito di applicazione inteso è ancora una volta quello in cui l'incertezza deriva dalla mancanza di precisione dell'informazione disponibile.

Nella logica della possibilità a ciascuna proposizione *p* sono assegnati un *grado di possibilità* e un *grado di necessità* che *p* sia vera, ove il grado di possibilità è indicato con $\Pi(p)$. Gli assiomi fondamentali della logica della possibilità richiamano alla mente quelli della teoria della possibilità. La combinazione di diversi corpi di evidenza avviene mediante particolari connettivi. Se i simboli "¬", "∧", e "∨" denotano rispettivamente gli usuali connettivi logici di negazione, congiunzione e disgiunzione, avremo:

$$\Pi(\neg p) = 1 - \Pi(p)$$
$$\Pi(p \vee q) = \max\,(\Pi(p), \Pi())$$
$$\Pi(p \wedge q) = \min\,(\Pi(p), \Pi(q)).$$

Non ci soffermeremo ulteriormente sulla logica possibilistica se non per inserirla nel contesto *q* generale dei metodi per il trattamento dell'incertezza. In generale, la questione delle relazioni tra probabilità e approccio *fuzzy*, intesa come questione se vi siano problemi che possano essere modellati soltanto con un approccio *fuzzy* e non con un approccio probabilistico è stata molto dibattuta. Le critiche che i sostenitori dell'approccio *fuzzy* (in particolare Zadeh) rivolgevano contro la probabilità sono le seguenti:

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI FILOSOFIA
NETWORK

a) la probabilità non modella adeguatamente la dipendenza dal contesto;
b) una distribuzione di possibilità è qualcosa di diverso da una distribuzione di probabilità;
c) la probabilità richiede sempre che si possano definire con precisione valori di incertezza;
d) la probabilità richiede una determinazione precisa di mondi possibili disgiunti (e questo richiede che sia nitido il numero delle alternative in questione e il loro significato); ma questo non sempre è facile.

Grazie alla mole di ricerche volte ad integrare i due approcci e il progresso dei metodi probabilistici in AI, oggi la *querelle* ha perso mordente. Oggi, più che di rivalità, si può parlare di prospettive differenti e complementari, che di volta in volta possono essere scelte sulla base di considerazioni pragmatiche e che spesso vengono combinate. L'approccio *fuzzy* è visto con particolare favore in ambito medico, per la difficoltà di dare una caratterizzazione nitida di condizioni quali "malato", "sano", "prediabetico", "osteopenia", ecc. (si veda Sadegh-Zadeh 2012) e in campo giuridico, dove la probabilità trova notevoli difficoltà di applicazione.

3.3 Il commonsense reasoning e i suoi problemi

Nelle sezioni precedenti abbiamo esposto i principali approcci quantitativi all'incertezza. In questi due ultimi paragrafi ci dedichiamo a una breve disamina del programma logicista di formalizzazione del ragionamento di senso comune. Ma che cosa è veramente il ragionamento di senso comune? In un certo senso è fin troppo facile definirlo: si tratta di quell'insieme di conoscenze e inferenze che consentono ad una persona che abbia superato la prima infanzia di muoversi nello spazio, formulare e realizzare progetti, prevedere l'effetto di eventi e azioni nella vita quotidiana, senza fare ricorso a particolari nozioni scientifiche. In realtà, una definizione formale adeguata di questo tipo di ragionamento è molto complessa. Lo si mostra con un noto esempio: una persona vuole recarsi da casa propria all'aeroporto; costui sa che la propria automobile è davanti a casa; il senso comune gli dice che l'obiettivo sarà raggiunto se va a piedi fino alla macchina e poi guida la macchina fino all'aeroporto. E' possibile esprimere l'asserzione che questo piano gli consentirà di raggiungere il suo obiettivo con una formula espressa in un linguaggio della logica matematica e successivamente dimostrare tale formula sulla base di assiomi che descrivono alcuni fatti comunemente noti riguardo al camminare e al guidare?

Per McCarthy[11], l'abilità di dedurre automaticamente le conseguenze immediate di ciò che è già noto ha «molto in comune con ciò che ci consente di descrivere alcuni esseri umani come persone dotate di *commonsense*», dove il *commonsense* è la capacità di orientarsi in diverse situazioni; è quindi strettamente dipendente dal contesto e dal tipo di situazioni in cui si opera. Per formalizzare questa dipendenza McCarthy elaborava un particolare calcolo logico, il *calcolo delle situazioni*, dove una *situazione* era descritta, intuitivamente, come «lo stato di cose completo del mondo in un certo istante di tempo». [12] Il problema era che una situazione è, almeno in linea di principio, infinitamente ricca e pertanto non completamente descrivibile; i suoi tratti salienti possono essere individuati mediante opportune *qualificazioni*, ma anche questa strategia comporta, come vedremo, notevoli problemi. Lo studio del ragionamento comune era dunque destinato ad estendersi in nuove direzioni. Nella sua analisi del *commonsense* McCarthy distingue tra il problema di rappresentare la *conoscenza* di senso comune e quello di formalizzare il *ragionamento* di senso comune. Rappresentare la *conoscenza* di senso comune richiede la capacità di rappresentare il fatto che gli oggetti cambiano posizione e sono talvolta creati o distrutti, il fatto che alcune situazioni cambiano nel tempo come conseguenza di eventi, la capacità di rappresentare la metaconoscenza (conoscenza circa la propria conoscenza), di "dialogare" con la conoscenza altrui, di guardare da una prospettiva epistemologicamente diversa ai fenomeni studiati dalle scienze esatte (per esempio, nel prevedere il comportamento dell'acqua in un bicchiere che cade, non ricorriamo generalmente alle equazioni sui corpi in caduta o a quelle di Navier-Stokes sui fluidi)[13].

Il *ragionamento* di senso comune è caratterizzato dal fatto che le sue inferenze non rispettano alcune proprietà della logica classica, in particolare la *proprietà di monotonicità*. Ciò non comporta che la logica sia del tutto incapace di formalizzare il *commonsense reasoning* e che vada abbandonata in favore di approcci alternativi: occorre piuttosto «arricchirla con metodi aggiuntivi di ragionamento non monotòno, i quali sono tanto formali quanto

---

[11] Qui accenneremo solo al calcolo delle situazioni e alla logica della *circumscription*. Per una visione complessiva dell'importanza di McCarthy e la sua influenza in IA si veda Morgenstern-McIlaith (2011).

[12] Su un aggiornamento del calcolo delle situazioni di McCarthy vedi Lakemeyer-Levesque (2011).

[13] Da qui nasce anche l'idea di formalizzare una "fisica ingenua" che corrisponde alla visione intuitiva dei cambiamenti nello spazio e nel tempo (Hayes 1979).

lo è la deduzione e altrettanto suscettibili di studio matematico e di implementazione su calcolatore» (McCarthy 1990, 194). [14]

Queste parole ben sintetizzano la prospettiva generale dell'approccio logicista. Da un punto di vista storico, lo studio delle logiche non monotòne fu consacrato a settore di punta dell'IA nel 1980, con la pubblicazione di un numero monografico della rivista *Artificial Intelligence* nel quale erano messi a confronto i principali formalismi per il ragionamento non monotòno. La strategia dell'*orientamento logicista* in IA era sintetizzabile in tre passi:

(1) i programmi devono basarsi su una grande quantità di conoscenze;
(2) le conoscenze devono essere in qualche modo rappresentate nei programmi;
(3) dobbiamo specificare la conoscenza in formule logiche prima di scrivere i programmi stessi.

Quale soluzione individua McCarthy per trattare il ragionamento di senso comune? Innanzitutto, ne identifica una caratteristica distintiva, l'uso di generalizzazioni flessibili. Nei nostri ragionamenti quotidiani trattiamo come premesse universali certi enunciati pur sapendo che in realtà non lo sono: ne conosciamo infatti le possibili eccezioni, ma decidiamo di ignorarle e *saltiamo alle conclusioni*; accettiamo così quelle che *sarebbero* le conseguenze delle premesse *se* queste fossero autentici enunciati universali. Le conclusioni così ricavate sono, per così dire, accettate *con la riserva* che, qualora si scopra una eccezione, esse possano essere ritrattate, senza che ciò comporti l'invalidazione delle credenze iniziali.

L'esempio immancabilmente utilizzato in letteratura per chiarire questa forma di ragionamento, detto *ragionamento di default* (cioè in difetto di informazione contraria), è quello in cui dalle premesse "Normalmente gli uccelli volano" e "Titti è un uccello" vorremmo trarre la conclusione "Titti vola". Quando affermiamo "Normalmente (tipicamente, di solito, in generale) gli uccelli volano" sappiamo bene che vi sono uccelli che non volano: gli struzzi, i pinguini, gli uccelli finti da richiamo, e così via. Tuttavia, se di un certo individuo, Titti, sappiamo soltanto che è un uccello, *saltiamo alla conclusione* che esso vola, conclusione che sarà ritrattata se entreremo in possesso dell'informazione che Titti è un uccello di tipo particolare, ad esempio un pinguino. E' possibile dare una rappresentazione formale di questo tipo di ragionamento? La logica classica non è adeguata, poiché la proprietà di monotonicità comporta che la validità delle

[14] Sullo sviluppo della logica nel ragionamento non monotono si veda Sandewall (2011).

APhEx
PORTALE ITALIANO DI FILOSOFIA ANALITICA
GIORNALE DI FILOSOFIA
NETWORK

conclusioni non può essere messa in questione da ulteriori informazioni. Nel ragionamento di senso comune il riferimento a individui ‘normali’ può essere visto come basato su prototipi. Come abbiamo visto nel paragrafo 2.3, l’uso di prototipi può essere ricollegato in parte ad una forma particolare di incertezza, l’approssimazione, ed in parte all’incompletezza delle informazioni. Le logiche non monotone rappresentano uno dei tentativi di rispondere alla domanda sulla possibilità di dare una rappresentazione formale, ma devono affrontare due ulteriori problemi.

I due importanti problemi che ogni logica per il ragionamento di senso comune deve affrontare riguardano le azioni. In generale, il compimento di un’azione modifica in parte una situazione e in parte la lascia immutata: si può dire, schematicamente, che le cose rimangono invariate, tranne alcune eccezioni rappresentate da ciò che cambia; si possono pertanto intuire le connessioni tra ragionamento non monotòno e tentativi di formalizzare le nozioni di *azione*, di *cambiamento* nel tempo e di *causalità*, intesa come manipolazione di un sistema da parte di un agente esterno al sistema. Queste connessioni si evidenziano sui due problemi considerati come il banco di prova canonico per ogni formalizzazione del ragionamento plausibile: il *frame problem* e il *qualification problem*.

Il *frame problem* è legato ai tentativi di rappresentare la conoscenza di un mondo dinamico. In conseguenza di azioni o eventi alcuni stati di cose cambiano ed altri rimangono inalterati; alcune delle proposizioni che descrivono tali stati mutano il loro valore di verità, mentre altre lo conservano. Come prevedere quali stati di cose permarranno immutati e quali cambieranno nel corso del tempo? L’introduzione di una serie di assiomi (*frame axioms*) che specifichino, per ogni situazione, quali stati del mondo non vengono modificati da un certo evento sembra necessaria, ma impraticabile. Come ricorda P. Hayes, «se vi sono 100 azioni e 500 proprietà temporanee, potremmo avere bisogno di 50.000 di questi stupidi *frame axioms*». Occorre dunque trovare una maniera economica e sensata di descrivere in maniera succinta quali sono i cambiamenti introdotti da un’azione «senza dovere elencare esplicitamente tutte le cose che essa lascia immutate». (Hayes 1987, 125). Una via di uscita dal *frame problem* è quella di sostituire i *frame axioms* con un più economico “Principio di persistenza”, che postuli che un certo evento o azione lasci immutate tutte quelle cose di cui non si afferma esplicitamente che vengono mutate. Ancora una volta ci troviamo nell’ambito di un ragionamento non monotòno, in cui si assume che “normalmente” le cose rimangono come sono, tranne le eccezioni, ossia gli oggetti che subiscono un cambiamento.

Il *qualification problem* riguarda invece il fatto che è sempre possibile aumentare il livello di dettaglio della base di conoscenze di un sistema. Tuttavia l'abbondanza di dettagli potrebbe ostacolare seriamente la capacità del programma di impostare in maniera efficiente la soluzione dei problemi: se infatti la rappresentazione del mondo è abbastanza dettagliata da includere la conoscenza di eccezioni, condizioni invalidanti, malfunzionamenti, eccetera, il programma controllerà ogni volta se esse valgano o meno e sarà costretto a esaminare ogni volta un numero ingestibile e inverosimile di *qualificazioni*. Occorre dunque formalizzare l'assunzione che, in mancanza di informazioni specifiche, l'enunciazione del problema contiene tutti e soli gli elementi rilevanti per la sua soluzione.

3.4 Le logiche non monotòne

Il programma di ricerca logicista in IA consiste nel tentativo di formalizzare, mediante opportune estensioni della logica classica e di quella modale, quelle inferenze che, pur non rispettando i canoni di validità della logica classica, si rivelano nondimeno uno strumento *efficiente* del ragionamento di senso comune in quanto permettono di trarre conclusioni (provvisorie) quando l'informazione è incompleta o, viceversa, troppo vasta per essere esaminata in tempi ragionevoli. Il fatto di rimanere all'interno di un orizzonte, in senso lato, logico rispecchia l'esigenza di garantire comunque una certa *affidabilità* alle "inferenze fallibili": l'intero programma di ricerca logicista può essere visto come il tentativo di raggiungere un equilibrio tra un'esigenza di efficienza e un'esigenza di affidabilità[15].

La proprietà di monotonicità è strettamente connessa con la nozione sintattica di *prova* (o *dimostrazione*) e con quella semantica di *conseguenza logica*; è comprensibile pertanto che le varie formalizzazioni dell'inferenza non monotòna comportino l'abbandono o la modifica di uno di questi due aspetti. Per quanto riguarda il primo aspetto, la monotonicità richiede che in una dimostrazione standard la validità di ciascun passo della dimostrazione dipenda esclusivamente dalle formule che compaiono negli stadi precedenti o nell'insieme degli assiomi di partenza e il fatto di aggiungere ulteriori assiomi non può invalidare una formula precedentemente derivata. Dal punto di vista semantico, la monotonicità può essere qualificata come segue: una formula α è conseguenza logica di

[15] Si veda Israel (1994).

un insieme di formule Γ se e soltanto se α è vera in tutti i modelli di Γ. Se anche si aggiungesse un insieme di formule Δ poiché tutti i modelli di $\Gamma \cup \Delta$ sono anche modelli di Γ, allora, se α è conseguenza logica di Γ, allora è conseguenza logica anche di $\Gamma \cup \Delta$.

Una caratterizzazione dell'inferenza non monotòna in termini di modelli porterà a modificare la nozione standard di modello o quella conseguenza logica o entrambe. Distinguiamo le seguenti logiche non monotòne in due gruppi, a seconda della prospettiva da cui affrontano il problema dell'inferenza non monotòna:

(i) logiche basate sulla coerenza (*default logics*, *logiche modali* e *logiche autoepistemiche*, *circumscription logics*) privilegiano gli aspetti sintattici; esse partono dalla considerazione di una teoria contenente enunciati (o *regole)* di *default* e si concentrano sulla determinazione di quali conclusioni siano derivabili da essa; schematicamente, l'idea è che una conclusione di *default* è corretta se la sua negazione non è derivabile dalle premesse;

(ii) logiche *a modelli preferenziali* (la denominazione è di Brewka, Dix, Konolige 1993) sono più specificamente semantiche, e perseguono l'obiettivo di minimizzare l'estensione di determinati predicati o relazioni. La nozione di conseguenza logica è diversa da quella classica: una formula α è conseguenza logica (non-monotòna) di un insieme di formule Γ se e soltanto se α è vera in tutti i modelli "preferiti" di Γ. Nel caso delle logiche circoscrittive, tali modelli sono identificati con i modelli minimali di una teoria[16]. Variando la definizione di preferenza si ottengono di volta in volta semantiche diverse per le diverse logiche considerate.

La ricerca sulla semantica a modelli preferenziali ha puntato alla costruzione di un quadro unificato, all'interno del quale i sistemi fossero derivabili come casi particolari. Tra i contributi in questo senso vanno ricordati il sistema C di Kraus, Lehman e Magidor), il sistema P dei modelli preferenziali (Kraus et al. 1990), le *ranking functions* di W. Spohn (1988), la ε-semantica di Geffner e Pearl. Friedman e Halpern (1995) propongono un approccio all'incertezza basato su misure della plausibilità, dove una misura di plausibilità associa a ciascun evento un elemento di un insieme parzialmente ordinato. Tale approccio è una generalizzazione di altri approcci all'incertezza, quali quello probabilistico, le funzioni di

[16] Limitiamo a questi brevi cenni la trattazione delle logiche non monotone; per una esposizione più ampia rinviamo, oltre ai già citati Sandewal (2011) e Benzi (1997), a Gabbay (1985, 1994), Strasser e Antonelli (2018).

credenza e le misure di possibilità; inoltre presenta interessanti relazioni tra queste e le logiche non monotòne.

Un aspetto interessante delle logiche non monotone è che si prestano alla formalizzazione del *defeasible resoning*, ovvero quelle forme di ragionamento nelle quali si vuole essere liberi di poter ritrattare conclusioni precedentemente affermate. Questo le rende particolarmente adatte allo studio e all'automazione dell'argomentazione in campo legale[17].

4. Conclusioni

Molti dei metodi per il ragionamento in condizioni di incertezza che abbiamo presentato – ai quali si affiancano numerosi altri – sono nati intorno agli anni'80 e, con i debiti sviluppi, sono tuttora applicati. E' ancora presto per fare previsioni sulla loro sopravvivenza a fronte degli stupefacenti processi del *machine learning*. Ci sembra però che valga la pena di ricordare brevemente l'opinione di uno dei più autorevoli guru dell'incertezza, Judea Pearl. In (Pearl 2018) ci ricorda che le *learning machines* utilizzano quasi esclusivamente informazione di carattere statistico. Vi sono tuttavia aspetti del ragionamento per i quali la statistica non è sufficiente. Pearl, naturalmente, ha in mente il ragionamento causale, che rende necessario immaginare che cosa può accadere se si *interviene* sulla realtà, o di immaginare «ipotetici scenari alternativi per l'apprendimento e la pianificazione». Pearl suggerisce di integrare i sistemi di *learning* con strumenti per il ragionamento causale, quali le reti bayesiane causali e i modelli strutturali. Ci piace credere che considerazioni analoghe possano essere fatte per situazioni nelle quali si debba decidere, o agire, sulla base di *possibilità*, di nozioni vaghe, o ambigue, o in penuria di dati statistici, e che gran parte del lavoro dedicato all'incertezza possa ancora svolgere, magari accanto al *machine learning,* un *ruolo essenziale*.

Bibliografia

Bathnagar R.J., Kanal L.N., 1986, «Handling uncertain information: a review of numeric and non-numeric methods», in Kanal e Lemmer 1986, pp. 3-26.

Bayes T., 1763, «An essay towards solving a problem in the doctrine of chance» *Philosophical Transactions of the Royal Society of London,*

[17] Si veda ad esempio Walton (2011).

53, pp. 370-418. Riproduzione facsimilare con commento in Molina E. C. e Deming E., *Facsimiles of Two Paper by Bayes*, New York, Hafner, 1963.

Benzi M., 1997, *Il ragionamento incerto*. *Probabilità e logica in intelligenza artificiale*, Milano, Franco Angeli.

Boden M., 1990, *The Philosophy of Artificial Intelligence*, Oxford, Oxford University Press.

Brachman R., Levesque H. (eds), 1985, *Readings in Knowledge Representation*, Los Altos, CA, Kaufmann.

Brewka G., Dix J., Konolige K., 1993, *A tutorial on nonmonotonic reasoning*, Lisboa, 5th European Summer School in Logic, Language and Information.

Carnap R., 1950, *Logical Foundations of Probability*, Chicago, Chicago University Press, 1950.

Cohen L. J., 1977, *The Probable and the Provable*, Oxford, Clarendon Press.

Cox R.T., 1946, «Probability, frequency and reasonable expectation», *American Journal of Physics* 14, 1, pp. 1-13, anche in Shafer e Pearl 1990, pp. 353-365.

de Finetti B., 1931, «Sul significato soggettivo della probabilità», *Fundamenta Mathematicae* 17, pp. 289-329, anche in P. Monari e D. Cocchi (a cura di), *Bruno de Finetti: Probabilità e induzione (Induction and Probability)*, Bologna, CLUEB, 1993, pp. 31-61.

Dubois D., Prade H., 1980, *Fuzzy Sets and Systems*: *Theory and Applications*, New York, Academic Press.

Dubois D., Prade H., 1988, «An introduction to possibilistic and fuzzy logics», in *Non-standard Logics for Automated Reasoning*, New York, Academic Press. Anche in Shafer e Pearl 1990, pp. 742-761.

Dubois D., Nguyen H. T., Prade H., 2000, «Possibility theory, probability and fuzzy sets misunderstandings, bridges and gaps», in *Fundamentals of Fuzzy Sets*, Boston, MA, Springer, pp. 343-438.

Eagle A. (ed), 2010, *Philosophy of Probability*, London, Routledge.

Fetzer J. H., 1974, «A single case propensity theory of Explanation», *Synthese*, 28, 2, pp. 171-198.

Friedman N., Halpern J. Y., 1995, «Plausibility measures: a user's guide», in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, Los Altos, CA, Morgan Kaufmann, pp. 175-184.

Frixione M., Numerico T., 2013, «Alan Mathison Turing», *APhEX*. On-line: http://www.aphex.it/public/file/Content20141117_19.APhEx7,2013ProfiliTuringFrixione-Numerico.pdf

Gabbay D. M., 1985, «Theoretical foundations for nonmonotonic reasoning in expert systems», in Apt K. R. *et al.* (eds) *Proceedings NATO Advanced Study Institute on Logics and Models of Concurrent Systems, La-Colle sur-Loup, France*, Berlin, Springer, pp. 439-457.

Gabbay, D. M., Hogger C. J., Robinson J. A. (eds), 1994, *Handbook of Logic in Artificial Intelligence and Logic Programming*, Oxford, Clarendon Press.

Gaio S., 2010, «Vaghezza», *APhEX*. On-line: http://www.aphex.it/public/file/Content20100621_gaioconpagine.pdf

Galavotti M. C., 2005, *Philosophical Introduction to Probability*, Stanford, CSLI.

Geffner H., Pearl J., 1987, *A Framework for Reasoning with Defaults*, Technical Report R-66, Los Angeles, Cognitive Systems Laboratory, University of California.

Genesereth M. R., Nilsson N. J., 1987, *Logical Foundations of Artificial Intelligence*, Los Altos CA, Morgan Kaufmann.

Gillies D., 2000, *Philosophical Theories of Probability*, London, Routledge.

Glymour C., Scheines R., Spirtes P., Kelly K., 2014., *Discovering Causal Structure: Artificial Intelligence, Shilosophy of science, and Statistical Modeling*, Orlando, FL, - London, Academic Press.

Hacking I., 1975, *The Emergence of Probability*, Cambridge, Cambridge University Press. Trad. it.: *L'emergenza della probabilità*. Il Saggiatore, Milano 1987.

Hajek A., 2011, «Interpretations of Probability», *Stanford Encyclopedia of Philosophy*. On-line: https://plato.stanford.edu/entries/probability-interpret/

Hayes P., 1987, «What the frame problem is and isn't», in Pylyshyn Z. (ed), *The Frame Problem in Artificial Intelligence*, Norwood, NJ, Ablex, pp. 123-138.

Hayes P., 1979, «The naive physics manifesto», in Michie D. (ed), *Expert Systems in the Micro-Electronic Age*, Edinburgh, Edinburgh University Press, pp. 242-270.

Heckerman D., 1986, «Probabilistic interpretations for MYCIN's certainty factors», *Machine Intelligence and Pattern Recognition*, vol. 4, pp. 167-196. North-Holland.

Henrion M., Breese J. S., Horvitz E. J., 1991, «Decision analysis and expert systems», *AI magazine*, 12, 4, pp. 64-91.

Hitchcock C., 2001, «The intransitivity of causation revealed in equations and graphs», *The Journal of Philosophy*, 98, 6, pp. 273-299.

Horvitz E., Heckerman D., Langlotz C. 1986, «A Framework for Comparing Alternative Formalisms for Plausible Reasoning», In *AAAI. 5th National Conference on Artificial Intelligence*, Philadelphia, pp. 210-214.

Howson C., 1978. «On the prehistory of chance», *British Journal of Philosophy of science* 29, 3, pp. 174-280.

Israel D. J., 1994, «The Roles of Logic in Artificial Intelligence», in Gabbay D. M., Hogger C. J., Robinson J. A. (eds), *Handbook of Logic in Artificial Intelligence and Logic Programming*, Oxford, Clarendon Press, pp. 1-28.

Johnson W. E., 1921, *Logic*, Cambridge, CambridgeUniversity Press.

Kanal L. N., Lemmer, J. F. (eds), 1986, *Uncertainty in Artificial Intelligence*, Amsterdam, Elsevier.

Keynes J. M., 1921, *A Treatise on Probability*, London, Macmillan.

Kraus S., Lehman D., Magidor M., «Nonmonotonic reasoning, preferential models and cumulative logics», *Artificial Intelligence* 44, 1-2, pp. 167-207.

Lakemeyer G., Levesque H. J., 2011, «A Semantic Characterization of a Useful Fragment of the Situation Calculus with Knowledge», *Artificial Intelligence*, 175, 1, pp. 142-164.

Laplace P.S. de, 1967, *Opere*. A cura di O. Pesenti Cambursano, Torino, UTET.

Lifschitz V. (ed), 1991, *Artificial Intelligence and Mathematical Theory of Computation. Papers in Honor of John McCarthy*. Boston, MA, Academic Press.

Machina K. F., 1976, «Truth, belief and vagueness», *Journal of Philosophical Logic*, 5, 1, pp. 47-78.

Mamdani E. H., Gaines B. R. (eds), 1981, *Fuzzy reasoning and its applications*, New York, Academic Press.

McCarthy J., 1990, *Formalizing Common Sense*. *Papers by John McCarthy*, edited by V. Lifschitz, Norwood, NJ, Ablex.

McCarthy J., Hayes P., 1969, «Some philosophical problems from the standpoint of artificial intelligence», In Melter B., Michie D. (eds), *Machine Intelligence 4*, Edinbugh, Edinburgh University Press, pp. 463-502. Anche in McCarthy 1990, 21-63.

Mellor D. H., 2005, *Probability: A Philosophical Introduction*, London, Routledge.

Miller D. W., 1994, *Critical Rationalism: A Restatement and Defence*, Lasalle, Il., Open Court.

Morgenstern L., McIlraith S.A., 2011, «John McCarthy's legacy», *Artificial Intelligence,* 175, SI, pp. 1-24.

Moruzzi S., 2012, *Vaghezza, confini, cumuli e paradossi*, Roma, Laterza.

Paganini E., 2008, *La vaghezza*, Roma, Carocci.

Pearl J., 1988, *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA, Morgan Kaufmann.

Pearl J., 2000 *Causality: Models, Reasoning, and Inference*. New York, Cambridge University Press. 2*nd* ed. 2009.

Popper K. R., 1957, *The Poverty of Historicism*, London, Routledge.

Pearl J., 2018, *The Seven Pillars of Causal Reasoning with Reflections on Machine Learning*, Technical Report R-481, Department of Computer Science, University of California. On-line: https://ftp.cs.ucla.edu/pub/stat_ser/r481.pdf

Reichenbach H., 1949, *Philosophical Foundation of Probability*, Berkeley and Los Angeles, University of California Press.

Romizi D., 2009, *Fare i conti con il caso*, Bologna, CLUEB.

Sadegh-Zadeh K., 2012, *Handbook of Analytic Philosophy of Medicine Philosophy and Medicine series*, vol. 113, Berlin, Springer.

Sandewall E., 2011, «From systems to logic in the early development of nomonotonic reasoning», *Articial intelligence* 175, 1, pp. 416-427.

Schuessler R., 2014. «Probability in medieval and Renaissance Philosophy», *Stanford Encyclopedia of Philosophy.* On-line: https://plato.stanford.edu/entries/probability-medieval-renaissance/

Shafer G., Pearl J. (eds), 1990, *Readings in Uncertain Reasoning*, San Mateo, CA, Morgan Kaufmann.

Spiegelhalter, D. J., 1986, «Probabilistic reasoning in predictive expert systems», in Kanal e Lemmer 1986, pp. 47-68.

Spohn W., 1988, «Ordinal Conditional Functions: a dynamic theory of epistemic states», in Harper W. L., Skyrms B. (eds), *Causation in Decision, Belief Change and Statistics*, Dordrecht, Reidel, pp. 105-134.

Strasser C., Antonelli A., 2016, «Nonmonotonic Logic», *Stanford Encyclopedia of Philosophy*. On-line: https://plato.stanford.edu/entries/logic-nonmonotonic/

Thomason R. H., 1991, «Logicism, AI, and common sense: John McCarthy's program in philosophical perspective», in Lifschitz 1991, pp. 449-466.

Thomason R. H., 2018, «Logic and Artificial Intelligence», *Stanford Encyclopedia of Philosopy*. On line:
https://plato.stanford.edu/entries/logic-ai/

Turing A., 1936-1937, «On computable numbers, with an application to the *Entscheidungsproblem*» *Proceedings of the London Mathematical Society*, 2, 42, pp. 230-365, 43, pp. 544-46.

Tverski A., Kahneman D., 1974. «Judgement under Uncertainty: Heuristics and Biases», *Science* 185.4157, pp. 1124-1131, anche in Shafer e Pearl 1990, pp. 32-39.

von Mises R., 1957, *Probability, Statistics and Truth*, edizione inglese riveduta, New York, McMillan.

Woodward J., 2003, *Making Things Happen: A Theory of Causal Explanation*, Oxford, Oxford University Press.

Zadeh L. A., 1965, «Fuzzy sets», *Information and Control*, 8, pp. 338-353.

Zadeh L. A., 1978, «PRUF: a meaning representation language for natural languages», *International Journal of Man-Machine Studies*, 10, 4, pp. 395-460. Anche in Mamdani e Gaines 1981, 1-65. Le indicazioni di pagina nel testo si riferiscono a quest'ultima edizione.