

ALFONSO DI PROSPERO

**INTELLIGENZA ARTIFICIALE E INFERENZA NON-MONOTONA:
MODELLI CULTURALI E QUESTIONI EPISTEMOLOGICHE**

1. *Modelli di comunicazione e condizionamento delle vite*
2. *L'I.A. tra implicazioni culturali e problemi epistemologici* 3. *Circoli di prossimità epistemica*

**ABSTRACT: ARTIFICIAL INTELLIGENCE
AND NON-MONOTONIC INFERENCE:
CULTURAL MODELS AND
EPISTEMOLOGICAL ISSUES**

In the research in A. I. John McCarthy has shown the interest of the non-monotonic reasoning. My attempt is to connect such topic with Piaget's genetic epistemology, within an inductivistic approach. The hypothesis is that this theoretical framework can be adapt to deal with both some features of the logical form of human thought - and this can entail relevant consequences also for A. I. - and for the cultural and social issues due to the widespreading of the digital communication.



1. *Modelli di comunicazione e condizionamento delle vite*

Luciano Floridi ha ripetutamente insistito che l'intelligenza artificiale riesce a ottenere successi e applicazioni sempre più consistenti non grazie a un progresso nei tentativi di *imitare* la mente umana *globalmente*, in un modo che consentirebbe di parlare di "intelligenza" artificiale nel senso in cui agli albori di questo tipo di ricerche - o nei romanzi di fantascienza - si prefiguravano robot "simili" agli esseri umani, ma piuttosto attraverso la combinazione da un lato della scomposizione delle funzioni di cui è capace la mente umana (creando sistemi esperti

in grado di effettuare ciascuno una sola delle innumerevoli operazioni di cui è capace la mente umana) e dall'altro del modificarsi progressivo che le attività umane hanno - per esempio nelle produzioni industriali - in vista di un *adattamento* sempre più perfezionato alle modalità di funzionamento dell'intelligenza artificiale stessa¹. Il "formato" delle nostre attività - e alla lunga delle nostre esistenze - si adatta alla struttura delle attività implementate con mezzi artificiali: da un diverso punto di vista, usando le parole di Éric Sadin, si può vedere in questo una «svolta ingiuntiva della tecnica»².

Il reciproco integrarsi delle esigenze imposte dalle tecnologie e delle forme che le nostre azioni vengono ad avere, suggerisce in modo naturale l'idea (che lo stesso Floridi prende in considerazione) che processi dello stesso genere possano essere riferiti non solo al presente e al futuro, ma anche al passato, in particolare se si pensa alle celebri tesi di Marshall McLuhan e degli autori della cosiddetta Scuola di Toronto sul rapporto tra mezzi di comunicazione e forme di pensiero nel corso della storia umana.

Questo vorrebbe dire che anche le diffuse resistenze che si riscontrano contro il diffondersi delle tecnologie digitali per la comunicazione, potrebbero essere correttamente paragonate alle critiche che un filosofo come Platone ha rivolto contro la tecnologia della scrittura nel periodo in cui in Grecia essa andava diffondendosi. Un tale paragone ancora non basta a screditare le posizioni dei più pessimisti: le scelte e i giudizi da discutersi qui sono di carattere storicamente contestuale. Si può considerare confutata una formulazione "alla lettera" delle idee di Platone, in quanto smentita dall'evoluzione successiva. Si dà però la necessità di valutare in che modo un sistema di

¹ L. Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo* (2014), tr. it. Raffaello Cortina, Milano 2017.

² É. Sadin, *Critica della ragione artificiale. Una difesa dell'umanità* (2018), Luiss University Press, Roma 2019, p. 29.

convinzioni si inserisca e interagisca con i fatti che si danno nel momento storico in cui esso viene prodotto. Da ciò deriva che è impossibile trarre l'implicazione che si sia automaticamente in errore se si esprimono riserve di fronte a un troppo rapido diffondersi delle nuove tecnologie della comunicazione nel mondo attuale, dato che si tratta di contesti del tutto diversi.

Inoltre è da osservare che, anche se oggi la tecnologia della scrittura ci appare come una acquisizione irrinunciabile, vi sono però studiosi che hanno condotto contro le conseguenze dovute alla sua introduzione delle critiche interessanti. Per Walter Ong la diffusione della scrittura ha creato *alienazione* e distanza tra le persone³. Non si tratta naturalmente di negare quanto l'invenzione della scrittura abbia potuto modellare - in meglio - le condizioni del vivere umano. È opportuno però cercare di elaborare un modello analitico globale che, tenendo conto in una certa misura anche dei nessi individuati da Ong, sia in grado di dar conto di alcuni significativi aspetti degli sviluppi attuali.

L'idea di fondo è che in un certo senso la *prima* tecnologia che ha condizionato e modellato la mente umana sia non la scrittura, ma lo stesso linguaggio (secondo un'idea che è presente per esempio in Bergson).

Radu Bogdan ha mostrato la plausibilità di un modello per cui il bambino non dispone di una struttura linguistica e cognitiva fondamentale: la forma tema-commento⁴. Applicando anche l'infrastruttura teorica dell'epistemologia genetica⁵, per la quale il bambino alla nascita non comprende le nozioni di spazio e tempo, di oggetto semi-permanente, né è capace di coordinare punti di vista diversi, per cui diventa in effetti quasi impossibile che possa essere presente la forma tema-commento, l'autore sostiene

³ W. Ong, *Interfacce della parola: studi sull'evoluzione della coscienza e della cultura* (1977), tr. it. il Mulino, Bologna 1989.

⁴ R. Bogdan, *Predicative Minds. The Social Ontogeny of Propositional Thinking*, MIT Press, Cambridge 2008.

⁵ J. Piaget, *La costruzione del reale nel bambino* (1937), tr. it. La Nuova Italia, Firenze 1973.

che questa si sviluppa progressivamente nei primi anni di vita, facendo da asse principale che rende possibile l'acquisizione delle più diverse conquiste della cultura.

L'ipotesi che ho cercato di sviluppare maggiormente in dettaglio in precedenti pubblicazioni⁶ è che la teoria di Jean Piaget possa essere re-interpretata in termini induttivistici: la condizione di egocentrismo che Piaget attribuisce al bambino sarebbe da associarsi con il suo essere un soggetto epistemico iper-induttivista, che procederebbe nello sviluppo applicando nelle sue operazioni cognitive generalizzazioni induttive che partono dai contenuti della sua propria esperienza - è in ciò appunto che si manifesta quello che possiamo vedere come "egocentrismo". In questo quadro è essenziale però sottolineare che l'inferenza induttiva deve essere caratterizzata come non-monotona: il bambino, facendo nuove esperienze, *non* è affatto vincolato a conservare come validi e veri i risultati delle generalizzazioni precedenti, ottenuti a partire da premesse induttive differenti. È da considerare che qui si vede l'induzione nella sua valenza psicologica, per cui è da attendersi ragionevolmente che giochino un ruolo importante i meccanismi dell'attenzione selettiva: perciò gli *script* e gli schemi di comportamento che si costituiscono e si stabilizzano non dovranno necessariamente essere respinti, con la comparsa di nuove esperienze che, in una logica non-monotona, potrebbero invalidarne il funzionamento (il che renderebbe impossibile il definirsi e lo stabilirsi in modo più o meno permanente di strutture cognitive man mano più complesse). Il fissarsi di schemi che continueranno a essere "preferiti" anche in seguito, quando la disponibilità di nuovi contenuti di esperienza formalmente renderebbe possibile il loro abbandono, può spiegarsi con il fatto che la dinamica dell'attenzione selettiva (anche al

⁶ A. Di Prospero, *La forma del significato. Saggio sui costituenti del pensiero*, Aracne, Roma 2020; Id., *Linguaggio e ordine del mondo*, Il Sileno, Rende (CS) 2020.

livello senso-motorio e di *embodied cognition*) continua ad alimentarne il funzionamento.

Anche il senso della reciprocità razionale, che per Piaget si acquisisce solo dopo anni di vita (non si insista qui eccessivamente su scansioni troppo rigide e prevedibili nella sequenza dello sviluppo, che sono state oggetto di critica da parte di molti studiosi) sarebbe ottenuto a partire da esperienze che hanno creato l'aspettativa che un ordinamento intersoggettivo e condiviso socialmente nelle funzioni del credere sia degno di rispetto, interesse e fiducia.

Un autore che è importante accostare qui a Piaget, per quanto poco citato, è Ash Gobar⁷, che ha cercato di dare un'interpretazione dell'epistemologia genetica che ne sottolinea i punti di contatto con la *Gestalt Psychology*. Riguardo al tema dell'intersoggettività l'autore scrive che

the degree of the intersubjectivity of knowledge varies directly with the degree of object constancy and inversely with the amount of critical interference: Thus objective knowledge (e.g. logic) builds its home on the bedrock of permanent constancy, and where critical interference prevails there also prevail conflicting viewpoints (e.g. aesthetics)⁸.

In termini intuitivi, l'ipotesi che si può proporre a partire da questa idea è che il bambino sia mosso da un fondamentale empirismo, che venendo attuato rigorosamente obbligherebbe però il soggetto a non poter condividere con altri segni e significati: la comunicazione - posti i vincoli dell'empirismo e la differenza tra le esperienze individuali - presuppone una misura di *indeterminazione* (è a questo riguardo che si dovrebbe inserire un collegamento - da svilupparsi più in dettaglio - con le idee sopra riferite di Ong).

Nel dibattito sulle comunicazioni digitali, Michael Lynch si richiama a un problema (simile a quello che Quine definisce "trazione oggettiva") che segue un'euristica analoga:

⁷ A. Gobar, *Philosophical Foundations of Genetic Psychology and Gestalt Psychology*, Martinus Nijhoff, The Hague 1968.

⁸ *Ibid.*, p. 281.

The problem of distinguishing the real from the unreal, or the true from the untrue, is hardly the result of the digital age [...] Take a coin out of your pocket and hold it in your hand before you. Now look at the coin: what shape does it look like? [...] If requested] A child might draw a circle, but a more skilled artist would draw the ellipse. Why? Because that's what we are perceiving. But if so, then we have a puzzle. The coin is circular. What we perceive is not circular. Therefore, what we perceive is not the coin⁹.

Anche C. Thi Nguyen¹⁰, parlando di *echo chamber* e *epistemic bubble*, torna a più riprese a osservare che sono fenomeni che esistevano in forma diversa anche in contesti del tutto differenti (per esempio nei regimi politici che praticavano il culto della personalità).

Il punto è che la realtà "virtuale" non fa che prolungare dinamiche e logiche presenti già in tutta la storia della socializzazione umana: si può comunicare - anche con chi ci è fisicamente di fronte - solo perimetrando i significati in un modo che li renda accettabili da tutti i partner della comunicazione coinvolti.

2. L'I.A. tra implicazioni culturali e problemi epistemologici

Questo schema teorico fornisce alcune premesse interessanti per affrontare le questioni poste dall'Intelligenza Artificiale e in generale dalla "cultura del digitale". In particolare: 1) i flussi di informazione sono sempre più caotici e violenti - salvo poter essere controllati solo in modo meccanizzato e attraverso algoritmi da soggetti non umani, data la quantità enorme di dati da considerare, e da forze economiche in grado di finanziare tutto il processo di elaborazione che si rende necessario: quella che si profila è la necessità di modificare il nostro modello di intersoggettività e di reciprocità razionale, allargandolo quantitativamente, secondo le esigenze imposte dalle dinamiche della globalizzazione, ma forse dovendolo anche trasformare qualitativamente, nella sua struttura epistemologica. Il rischio

⁹ M. Lynch, *The Internet of Us*, Norton, New York 2016.

¹⁰ C.T. Nguyen, *Echo Chambers and Epistemic Bubbles*, in «Synthese», 197, 7, 2020, pp. 2803-2821.

con cui ci si deve confrontare, entro questa prospettiva, non è tanto legato ai nuovi supporti tecnologici della comunicazione (come potrebbe sembrare leggendo per esempio Manfred Spitzer), ma piuttosto che l'ambiguità e l'equivocità *costitutive* della comunicazione in ogni epoca storica, dovendo oggi confrontarsi con identità e relazioni su di una scala che è planetaria, possano di fatto esplodere, rendendo i significati trasmessi nella comunicazione inservibili per l'individuo. In altre parole la persona, dovendo partecipare a pratiche di comunicazione la cui semantica astrae sistematicamente da quelle che sono le sue esperienze, deve rinunciare a dare un ruolo motivante nell'azione o a questa semantica *oppure* alle proprie esperienze¹¹.

2) Uno dei terreni su cui si sono combattute alcune delle battaglie per la realizzazione dell'intelligenza artificiale è proprio quello delle forme di ragionamento non-monotono: John McCarthy, lo scienziato che ha introdotto l'espressione "intelligenza artificiale", organizzando il convegno di Dartmouth del 1956 che ha dato inizio al dibattito su di essa, ha cercato di provvedere a creare delle formalizzazioni del ragionamento non-monotono, con il fine dichiarato di poter arrivare a permettere a dei computer di ragionare secondo forme che riproducano quelle degli esseri umani¹². Sullo sfondo naturalmente vi sono le speranze e problemi legati al test di Turing: l'idea è che gli esseri umani, a differenza dei computer, sappiano orientarsi - in una conversazione o in un ragionamento - riconoscendo le situazioni di ambiguità e di indeterminazione delle informazioni che ricevono, per cui un computer, per poter riprodurre e per potersi interfacciare con le modalità di ragionamento di un umano, dovrebbe poter essere programmato in modo da attenersi a una logica non-monotona. McCarthy propone quella che chiama *circumscription Logic*. L'idea è che

¹¹ Cfr. J. Arditi, *Simmel's Theory of Alienation and The Decline of the The Norational*, in «Sociological Theory», XIV, 2, 1996, pp. 93-108.

¹² J. McCarthy, *Formalizing Common Sense*, Ablex, Norwood, NJ 1990.

It seemed that in order to fully represent the conditions for the successful performance of an action, an impractical and implausible number of qualifications would have to be included in the sentences expressing them. For example, the successful use of a boat to cross a river requires, if the boat is a rowboat, that the oars and rowlocks be present and unbroken, and that they fit each other. Many other qualifications can be added, making the rules for using a rowboat almost impossible to apply, and yet anyone will still be able to think of additional requirements not yet stated¹³.

Lo strumento proposto dall'autore è quindi il seguente:

Circumscription is a rule of conjecture that can be used by a person or program for jumping to certain conclusions. Namely, *the objects that can be shown to have a certain property P by reasoning from certain facts A are all the objects that satisfy P*. More generally, circumscription can be used to conjecture that the tuples $\langle x, y, \dots, z \rangle$ that can be shown to satisfy a relation $P(x, y, \dots, z)$ are all the tuples satisfying this relation. Thus, we *circumscribe* the set of relevant tuples.

We can postulate that a boat can be used to cross a river unless something prevents it. then circumscription may be used to conjecture that the only entities that can prevent the use of the boat are those whose existence follows from the facts at hand¹⁴.

Hubert Dreyfus aveva scritto:

le macchine mancano di intelligenza pratica. Esse sono "esistenzialmente" stupide nel senso che non possono fronteggiare situazioni specifiche. Così esse non possono accettare l'ambiguità e la rottura delle regole, fino a che le regole per trattare le deviazioni siano state così completamente specificate che l'ambiguità sia scomparsa¹⁵.

Il problema è che la realtà degli esseri umani è contestuale e in divenire, mentre «le leggi della scienza sono universali e senza tempo, perché trattano tutta l'esperienza come se potesse essere allo stesso modo che nel passato»¹⁶: con una sorta di contraddizione intrinseca «gli studiosi di I.A. dovrebbero sviluppare una teoria non temporale e non spaziale di un'attività umana che si sviluppi in una situazione»¹⁷. Una macchina dovrebbe necessariamente «trattare con fenomeni che appartengono al mondo situazionale degli esseri umani come se questi fenomeni appartenessero all'oggettivo formale universo della scienza»¹⁸.

¹³ *Ibid.*, p. 142.

¹⁴ *Ibid.*

¹⁵ H. Dreyfus, *Che cosa non possono fare i computer* (1972), Armando, Roma 1988, p. 285.

¹⁶ *Ibid.*, p. 284.

¹⁷ *Ibid.*, p. 285.

¹⁸ *Ibid.*

Come si vede, la contrapposizione è tra logica classica - atemporale - e ragionamento non-monotono, e di questa distinzione McCarthy intende appunto avvalersi per rispondere a Dreyfus¹⁹.

Le logiche non-monotone sviluppate in seguito da altri hanno suscitato notevole interesse. Per il nostro discorso si può ricordare quella di Raymond Reiter, che prevede una “close world assumption”, per la quale «if no proof of a positive ground literal exists, then the negation of that literal is assumed true. This can be viewed as equivalent to implicitly augmenting the given data base with all such negated literals»²⁰.

Donald Gillies è un autore che si avvicina al problema da me posto, pur senza definirlo negli stessi termini. Il senso del suo contributo è di portare a riflettere sul significato dei successi che l’I.A. ha conseguito nell’ambito dell’apprendimento induttivo meccanico. L’autore sostiene che così si è “provato” che l’anti-induttivismo radicale di Popper è scorretto. È possibile che nella sua formulazione questa idea sia fuorviante: anche *prima* che l’*inductive learning* ottenesse nell’I.A. dei risultati significativi, l’anti-induttivismo di Popper poteva essere già fatto oggetto di critiche. In ogni caso l’autore afferma con forza che il ragionamento induttivo deve essere riconosciuto come non-monotono.

Un atteggiamento più corretto nei confronti dell’anti-induttivismo di Popper è difeso da Isaac Levi, che sospetta che molti studiosi impostino le proprie analisi dell’induzione in modo da evitare «a form of subjectivity in inquiry that will give comfort to the excess of cultural relativists and antirationalists»²¹, mentre

Others, as Popper, Carnap, and their antiinductivist followers, nurse and protect the objectivity of science by denying the legitimacy of inductive expansion in science. To counsel abstinence from induction

¹⁹ McCarthy, *Artificial Intelligence, Logic and Formalizing Common Sense*, in R. Thomason, ed, *Philosophical Logic and Artificial Intelligence*, Springer, Dordrecht, p. 4.

²⁰ G. Reiter, *On Closed World Data Bases*, in «Readings in Artificial Intelligence», 1981, pp. 119-140, p. 119.

²¹ I. Levi, *For The Sake of The Argument*, Cambridge University Press, Cambridge, Mass., 1996, p. 160.

in this manner is an admission of failure to bring the fixation of belief under critical control. Since inquirers will extrapolate from data and add theoretical conjectures to background information as resources in subsequent inquiry regardless of what antiinductivists say, the admission of failure concedes to the social relativists and antirationalists that induction is a topic for psychology, sociology and history to study and that there is no point in attempting to identify prescriptive standards for assessing its legitimacy²².

Per proseguire in questa indagine, vediamo in che modo Carl Gustav Hempel aveva provato a fornire «a purely syntactical definition» della conferma induttiva²³. Se S è una proposizione che deve confermare induttivamente M , vuol dire che M

determines a certain class of individuals [e.g.] $\{a, b, c\}$; and we may say that M confirms S because from the information contained in M it can be inferred that in $\{a, b, c\}$ S is completely satisfied; or, to use a metaphor: in a world containing exclusively the individuals a, b and c , the sentence S would be true, according to the information contained in M ²⁴.

Nelson Goodman ricapitola così il significato della definizione di Hempel:

a hypothesis is genuinely confirmed only by a statement that is an instance of it in the special sense of entailing not the hypothesis itself but its relativization or restriction to the class of entities mentioned by that statement. The relativization of a general hypothesis to a class results from restricting the range of its universal and existential quantifiers to the members of that class. Less technically, what the hypothesis says of all things the evidence statement says of one thing (or one pair or other n-ad of things)²⁵.

Sia Goodman sia Hempel cercano di perfezionare questa definizione, abbandonandola infine anche per i celebri paradossi da loro stessi formulati. Si vede comunque come la logica soggiacente ai concetti di “circostrizione” o di “mondo chiuso” operi in effetti anche qui in modo abbastanza chiaro.

Si può sostenere inoltre che, considerando l’inferenza induttiva come non-monotona, il paradosso di Hempel sembra che non disponga più dello spazio per poter sorgere.

Il problema viene enunciato dall’autore così:

²² *Ibid.*

²³ C.G. Hempel, *A Purely Syntactical Definition of Confirmation*, in «The Journal of Symbolic Logic», VIII, 4, 1943, pp.122-143, p. 122.

²⁴ *Ibid.*, p.130.

²⁵ N. Goodman, *Fact, Fiction, and Forecast*, 4th ed., 1983, pp. 69-70.

Consider, for example, the hypothesis “All swan are white”, which we symbolize by “ $(x)(Sw(x) \supset Wh(x))$ ”, and the molecules “ $Sw(a). Wh(a)$ ” and “ $\sim Sw(b) \sim Wh(b)$ ”. It might be felt—and would probably be argued by many—that under a materially adequate definition of confirmation, the first molecule should be confirming, the second neutral with respect to the hypothesis. However, according to the same intuitive standards, the second molecule would constitute confirming evidence with respect to the hypothesis “ $(x)(\sim Wh(x) \supset \sim Sw(x))$ ”. But the two hypotheses here mentioned are equivalent [per contrapposizione], and adherence to the intuitive standards exemplified in this illustration would therefore make the confirmation of a hypothesis by a certain body of evidence a matter not only of the content of the hypothesis, but also of the form in which it is expressed. This feature appears intuitively intolerable²⁶.

Il paradosso sta nel fatto che se osservo un corvo nero, che non è bianco e non è un cigno, per questa via posso confermare induttivamente che tutte le cose non-bianche sono non-cigni, e quindi – per contrapposizione – che tutti i cigni sono bianchi. Se però l’inferenza induttiva è riconosciuta come non-monotona, il passaggio logico da “Questo è un corvo nero” a “Questo è un oggetto non-bianco che è un non-cigno” diventa in effetti scorretto: per quanto ovviamente la deduzione in se stessa sia valida, una volta effettuato questo passaggio si è modificato di fatto l’insieme delle premesse disponibili, rendendolo estremamente più povero, per cui è come se si stesse facendo direttamente una *diversa* inferenza induttiva. Un’idea simile a questa è stata difesa da Branden Fitelson²⁷, che osserva come l’inferenza induttiva sia non-monotona, e per questo il paradosso *non* dovrebbe sorgere, per ragioni analoghe a quelle da me difese, ma ugualmente si impegna in una formulazione di una strategia di soluzione che, ispirandosi all’approccio seguito da molti a partire dalle analisi di Janina Hosiasson-Lindenbaum²⁸, utilizza

²⁶ C.G. Hempel, *op. cit.*, p. 128.

²⁷ Tra i vari articoli, si può vedere B. Fitelson, J. Hawthorne, *The Wason Task(s) and The Paradoxes of Confirmation*, in «Philosophical Perspectives», XXIV, 1, 2010, pp. 207-241. Un’analisi che si serviva di questa strategia è stata da me presentata in A. Di Prospero, *Un’analisi dei paradossi della conferma proposti da Carl Gustav Hempel*, in «Itinerari», 2003, 3, pp. 121-134; Id., *Art, Knowledge and Induction*, in «Rivista Italiana di Filosofia del Linguaggio», 2017, 2, pp. 147-163; Id., *Esperienza estetica, giustizia e inferenza induttiva*, in «Discipline filosofiche», 2021, 31, pp. 219-242.

²⁸ J. Hosiasson-Lindenbaum, *On Confirmation*, in «The Journal of Symbolic Logic», 1940, V, 4, pp. 133-149.

la teoria della probabilità per l'analisi del paradosso. L'impressione che si può avere - su cui però sarebbe necessario discutere con molta più ampiezza - è che questo studioso voglia presentare la sua strategia di analisi del paradosso come un successo dell'approccio bayesiano. Formalmente però è già il solo riferimento al carattere non-monotono dell'induzione che basta a confutare il paradosso²⁹.

In generale è interessante notare che viene a offrirsi una spiegazione del fatto che molti studiosi abbiano infine rifiutato lo strumento dell'induzione: in realtà si può ipotizzare che un modo fuorviante di riferirsi a essa - determinato dall'esigenza, presente in Popper o nel neopositivismo, di definire il metodo scientifico in modo da preservare *formalmente* il presupposto dell'intersoggettività del sapere scientifico - abbia portato a non dar peso al carattere non-monotono dell'induzione (che intrinsecamente oppone problemi urgenti a un tale principio di intersoggettività, se fatto valere *a priori* e non come oggetto di apprendimento *formalmente* contingente), spingendo infine al suo abbandono. Se invece ci si richiama all'epistemologia genetica, si può mostrare che il risultato di una sostanziale - e pragmaticamente fondamentale - intersoggettività del sapere (in particolare in ambito scientifico), può essere ottenuto con strumenti concettuali diversi.

L'intuizione di fondo è che il bambino, applicando momento per momento una logica non-monotona che obbedisce a un principio simile a quello della "circostrizione", e accogliendo in memoria gradualmente prima frammenti e poi segmenti di cognizione sempre meglio sistematizzata, ottenuti appunto con questi mezzi, edifica l'intera immagine del mondo, nella quale - in un senso che in

²⁹ La strategia euristica di fondo mi sembra che sia descritta bene da McCarthy: «Some people think it is possible to try to save monotonicity by saying that what was in your mind was [...] a probabilistic rule. Indeed it seems to me that when probabilistic reasoning (and not just the axiomatic basis of probability theory) has been formalized, it will be formally nonmonotonic» (J. McCarthy, *Artificial intelligence...*, cit., p. 11).

realtà è lapalissiano - le diverse prospettive degli altri soggetti epistemici sono contenute, ma in quanto *oggetto* del sapere del soggetto di riferimento.

Gillies riflette sulle conseguenze del fatto che il *Prolog* ha un comportamento di tipo non-monotono, dato che dispone solo di una negazione caratterizzata come “fallimento”. La posizione del bambino è in effetti analoga: seguendo Piaget, il bambino inizialmente non dispone del concetto di negazione, ma si limita a compiere azioni che - a volte - falliscono. Nella prospettiva di un induttivismo rigoroso, anche l’adulto in realtà, cercando di comprendere nel suo proprio esatto contenuto il punto di vista di un’altra persona, va incontro a un analogo fallimento, nel senso che i contenuti *esatti* di un *altro* punto di vista non possono essere per principio da me afferrati, perché allora non sarebbero più i contenuti di un *altro* punto di vista³⁰. Per questo nel riferirmi ai punti di vista altrui devo necessariamente ricorrere alla nozione di *indeterminazione*, in un senso per cui essi a rigore risulteranno dei *non*-contenuti. Si consideri anche che per le caratteristiche del ragionamento non-monotono, che nel dibattito sull’induzione trovano espressione in quello che Carnap chiama “principio dell’evidenza totale”, se io ho anche solo una minima parte dei miei contenuti di esperienza diversa da un altro soggetto, le applicazioni del principio dell’evidenza totale mi porteranno a generalizzazioni induttive a rigore diverse, producendo quello che può essere visto come un effetto di campo gestaltico, in grado in ogni caso di contraddire la possibilità di un atomismo rigido e non-contestualizzante.

³⁰ Gli autori che qui vorrei almeno rapidamente ricordare sono il Wittgenstein del *Tractatus Logico-philosophicus* e Nicholas Rescher (*Rationality*, Oxford University Press, Oxford 1988). Sul ragionamento non-monotono si deve anche C. Cellucci (*Le ragioni della logica*, Laterza, Roma-Bari 1998) mentre da una diversa prospettiva vorrei almeno menzionare D. Kincaid, *The Convergence Theory of Communication, Self-Organization, and Cultural Evolution*, in D. Kincaid, ed, «Communication Theory. Eastern and Western Perspectives», Academic Press, San Diego 1987, pp. 209-221.

Dreyfus è anche alla psicologia cognitiva dello sviluppo del bambino che fa riferimento per sostenere la sua critica all'intelligenza artificiale³¹, ma il modo in cui qui si impiega l'epistemologia genetica potrebbe aprire degli spazi per contestare anche proprio su questo punto specifico la sua argomentazione.

3. *Circoli di prossimità epistemica*

Per procedere in queste riflessioni, può essere interessante suggerire un confronto con l'analisi del concetto di "post-verità" proposta da Filippo Ferrari e Sebastiano Moruzzi³², che introducono a completamento le nozioni di "post-indagine" e "post-prova", richiamando anche la nozione di "pensiero collettivo" di Ludwik Fleck³³, che viene posta in analogia con quelli che gli autori chiamano "circoli di prossimità epistemica", composti

da tutti quegli agenti la cui indagine è regolata da uno stesso filtro epistemico [...] Questo filtro contribuisce a determinare una bolla epistemica. È plausibile pensare che ogni nostra indagine particolare sia regolata dalla presenza di filtri epistemici di questo tipo. Questo fatto non va però valutato necessariamente in modo negativo da un punto di vista epistemico - anche se nel caso dei terrapiattisti pensiamo certamente che lo sia³⁴.

È da notare che Fleck ha un'immagine abbastanza positiva del carattere intersoggettivo e (almeno in senso debole) consensuale della conoscenza scientifica. Ferrari e Moruzzi invece (prevalentemente) utilizzano il concetto di circolo di prossimità epistemica per evidenziare le fallacie e i paradossi di gruppi, come i "terraplattisti", che cercano unicamente le conferme - quali che possano essere - per le proprie convinzioni, sottoponendo le obiezioni a un trattamento epistemologico del tutto penalizzante. Il confronto con questi due autori, che ricapitolano e rappresentano in maniera chiara un punto di vista filosoficamente diffuso ma le cui implicazioni sono forse ancora

³¹ H. Dreyfus, *op. cit.*, p. 91.

³² F. Ferrari, S. Moruzzi, *Verità e post-verità*, 1088 Press, Bologna 2020.

³³ *Ibid.*, p. 85.

³⁴ *Ibid.*

da esplorare sotto vari aspetti, può essere interessante perché a una prima apparenza può sembrare che le considerazioni fin qui portate sulle funzioni dell'induzione siano in difficoltà nel cogliere la distinzione tra un consenso sociale "sano" e uno distorto e deviante (si pensi anche alle riflessioni sul concetto di fiducia condotte da Alvin Goldman³⁵).

In realtà, nell'indicare nell'esperienza personale e nell'induzione le basi del procedere cognitivo, si deve considerare che le nostre esperienze, in quanto tali, non sono affatto da noi scelte: per questo non vi sono margini per un relativismo che legittimi le scelte più arbitrarie e stravaganti. È vero però che formalmente la nostra convinzione che alcune credenze (oggi, con le nostre conoscenze) siano assurde, deve essere analizzata come dovuta a esperienze il cui statuto logico è quello della contingenza - il che non pregiudica la nostra certezza: se vedo che piove, ne sono sostanzialmente certo, anche se avrebbe potuto benissimo essere bel tempo.

Questo tipo di riflessione è apparentemente ovvia, ma pure nella sua formulazione generale si scontra con dei presupposti largamente fatti valere - spesso implicitamente - nella filosofia della scienza più diffusamente accettata, che accoglie il principio dell'intersoggettività del sapere scientifico come un presupposto che viene *trattato* (anche quando non si usano esattamente queste espressioni) come una verità o una necessità di tipo logico e *a priori* (si pensi a titolo di esempio a Popper).

Può essere interessante vederne alcune possibili implicazioni anche per una questione classica del dibattito sull'I.A., il test di Turing.

Dreyfus critica uno dei presupposti dell'I.A., sostenendo (come Searle) che per principio un computer non può avere un'intelligenza perché non può avere una semantica, posto che si

³⁵ A. Goldman, *Experts: Which One Should You Trust?*, in «Philosophy and Phenomenological Research», LXIII, 1, 2001, pp. 85-110.

utilizzi una definizione di informazione come quella data da Shannon in *The Mathematical Theory of Communication*³⁶. Si può però osservare che, se si vuole tener conto del carattere non-monotono dell'induzione - e del ruolo che l'induzione ha nel costituirsi dei significati - l'impostazione di Shannon è adatta proprio perché calibrata su una misura del tutto minima di condivisione univoca del significato, elaborata in modo da potersi dire che la "semantica" (se tale ancora può definirsi) sia riducibile alla sintassi (Wittgenstein nel *Tractatus*, e Moritz Schlick in modo più semplice, ma ispirandosi apertamente a Wittgenstein, attraverso le nozioni di "forma" e "contenuto", cercano di definire una teoria del significato di questo tipo). Essendo così impostata, una teoria "semantica" dovrebbe in effetti per principio potersi adattare a tutto il *range* delle possibili variazioni del significato, che il ragionamento non-monotono consente (quindi da un *minimo* a un *massimo* di possibile condivisione del significato), invece di dover presupporre l'accettazione di un codice univoco e condiviso.

Questo però non vuol dire che entro un simile quadro analitico la semantica, intesa come centrata sulle nostre esperienze in prima persona, scompaia. Probabilmente l'obiezione principale che può farsi al test di Turing è che fa riferimento a un modello della comunicazione che astrae dal fatto che nel comunicare ciò che i parlanti vogliono è sapere qualcosa che è ancora *al di fuori* del campo del loro attuale sapere: qualcosa che si può esprimere dicendo che le esperienze altrui - per il principio di non-monotonia - sono irriducibili alle nostre. Gli indizi di cui mi servo per valutare se un testo è prodotto da una macchina o no, corrispondono a un repertorio di conoscenze che *già* ho, non a quelle che il testo stesso dovrebbe permettermi di acquisire. Se chiedo a un amico il suo parere su una certa questione, il fatto che riesca a farmi credere che la sua risposta è pertinente può

³⁶ H. Dreyfus, *op. cit.*, p. 244.

essere una condizione necessaria ma non sufficiente di un'attribuzione di intelligenza. Quando - per esperienza - posso valutare che in una certa misura le esperienze (e quindi i significati) di un interlocutore "combaciano" con le mie, estendo a questo interlocutore la posizione di partecipe di ciò che chiamo intersoggettività - o "circolo di prossimità epistemica". Le mie esperienze però mi spingono a cercare *informazione* (attendibile), non le forme esteriori di una pratica del comunicare che astrae dai contenuti di verità (si deve considerare da questo punto di vista l'influenza della psicologia comportamentista su Turing). In questo senso, la non-condivisione di significati univocamente definiti tra me e un computer è un fatto formalmente di tipo empirico (come in un certo senso è ovvio che sia), così come lo è in generale il costituirsi di una dimensione della condivisione intersoggettiva tra me e gli altri esseri umani³⁷.

Anche l'utilizzo che McCarthy fa del concetto di ragionamento non-monotono risente molto (come peraltro era prevedibile) dell'impostazione data al problema da Turing, che d'altra parte è molto più vicina alle teorie epistemologiche più diffuse di quanto non lo siano le proposte che qui invece sono state discusse. Per motivi di spazio non è possibile approfondire come sarebbe necessario in questo scritto queste tematiche. Come si è cercato di mostrare, è però possibile concepire un quadro concettuale alternativo, dal quale deriverebbe la necessità di formulare in modo diverso (1) le questioni dell'I.A. associate in genere al test di Turing (e anche a quello che in filosofia della mente è il cosiddetto *hard problem*, anche per la rilevanza fondamentale che entro questo quadro viene ad avere una prospettiva in prima persona), (2) il problema, presentato in precedenza, dei modelli

³⁷ Sarebbe importante qui confrontarsi almeno con le posizioni formulate da R. French (*Subcognition and the Limits of the Turing Test*, in «Mind», XCIX, 393, 1990, pp. 53-65), da S. Bringsjord, P. Bello e D. Ferrucci (*CreatIbid.ty, the Turing Test, and the (Better) Lovelace Test*, in «Minds and Machines», XI, 2001, pp. 3-27) e da J. Moor (*The Status and Future of the Turing Test*, in «Minds and Machines», XI, 2001, pp. 77-93). Per ragioni di spazio si è obbligati a rimandare a una diversa occasione questa discussione.

di comunicazione tra *identità distanti*, così come oggi viene imposto dalle dinamiche della globalizzazione e delle comunicazioni digitali.

Queste due famiglie di problemi – per quanto fenomenologicamente molto lontane – forse possono in effetti essere affrontate all'interno di un quadro analitico almeno in parte comune.

ALFONSO DI PROSPERO è Dottore di Ricerca in Scienze Sociali e svolge attività di Docenza nelle Scuole Secondarie di secondo grado

alfonsodiprospero@yahoo.com