IM@GO
A JOURNAL OF THE SOCIAL IMAGINARY

Coscienza e Intelligenza Artificiale: due ontologie dell'immaginario[1]

Maurizio Bonolis
Maurizio.bonolis@uniroma.it
Dipartimento di Comunicazione e Ricerca Sociale | Sapienza Università di Roma

Barbara Sonzogni
Barbara.sonzogni@uniroma1.it
Dipartimento di Comunicazione e Ricerca Sociale | Sapienza Università di Roma

Abstract

Consciousness and Artificial Intelligence: two ontologies of the imaginary

The juxtaposition of the category of the Imaginary with what is commonly understood by the term Artificial Intelligence evokes representations of temporal dynamics based on the 'feeling of infinity'. However the dimensions of the Imaginary that come into play in this way are two and unmistakable. The first is that of the imaginary concerning the intangible horizon of intelligent systems design and processing programmes. The second meaning is that in itself obscure, and therefore no less indeterminate, which refers to the ability to simulate the existence of subjects endowed with decision-making autonomy. The concept of Artificial Intelligence refers here, in a word, to the ability to simulate consciousness. This is the real 'problem' of Artificial Intelligence. This is why the most authoritative sources of philosophical reflection and computer neurology continue to emphasise the disconcerting indefiniteness of consciousness, from a natural rather than an artificial point of view. The robotic simulation of attractiveness, for example in the form of sex toys, is both plausible and current, but less so the simulation of the aesthetic experience on the basis of which an attractive humanoid is axiomatically designed.

Keywords

AI | Metaphors | Narratives of AI | Imaginary | Children

[1] Pur essendo frutto di condivisione, a Maurizio Bonolis sono attribuibili i paragrafi 1, 2, 5 e 6; a Barbara Sonzogni i paragrafi 3 e 4; a entrambi gli autori il paragrafo 7.

ISSN: 2281-8138

N. 23 - Year XIII / July 2024 pp. 87-108
www.imagojournal.it

CC BY

1. La parola «immaginario»

Prima di accostare il concetto di *immaginario* al tema dell'Intelligenza Artificiale, nell'intento di mostrare in quale senso il primo possa designare una qualità della seconda, sembra opportuno convenire su una minima definizione di base del concetto, se non altro per via delle accezioni, con qualche esempio, alle quali esso si presta in rapporto ai diversi oggetti che ne sono *complemento di specificazione*. Lo scopo è porre l'accento su alcuni interrogativi che si ritiene sollevi, in particolare, la questione dell'immaginario dell'Intelligenza Artificiale. Si cercherà di sottolineare, a tale riguardo, che due, inconfondibili e reciprocamente irriducibili, sono le linee di tematizzazione che supportano questa ipotesi di lavoro: quella della *artificialità* della conoscenza e quella della *artificialità* della coscienza.

Per cominciare, «immaginario», può essere senz'altro un personaggio teatrale, mentre «immaginario» può essere ritenuto il piano logico sul quale si pongono le rappresentazioni eidetiche che, in riferimento a oggetti percepiti precedentemente, riescono a riprodurli con un dettaglio non riducibile alla nozione di memoria fotografica. E «immaginaria» può essere la *visio* inerte, semicosciente, che compare nelle esperienze ipnagogiche connesse all'affievolirsi degli stati di veglia, quando sta per arrivare il sonno. «Immaginaria», infine, può essere la raffigurazione di un oggetto del pensiero, il contenuto di una «intenzionalità» in senso husserliano, nello stesso senso in cui Wittgenstein sosterrà che il *pensabile* comporta sempre potersene fare un'immagine (*Tractatus, 3001*). Ma già nel *De Anima*, per Aristotele si dà unità di pensiero e immagine: pensare una cosa è renderla presente 'in immagine', nella modalità di un contenuto «pre-riflessivo» che pervade la nostra mente (Crespi, 1993: 14, 19).

Quest'ultima accezione vale anche a dar conto del fatto che intenzionalità condivise possono costituire «immaginario collettivo». Di ciò si sono occupati a suo tempo, pure nel contesto di prospettive disciplinari diverse, Gustav Jung (1934-1954), com'è noto, ma anche Edward Sapir (1929), in antropologia culturale. Poi, con non inferiore autorevolezza, Gadamer si è riferito a qualcosa di simile in prospettiva storicizzata, lamentando la crescente incidenza di produzioni culturali alimentate dalla comunicazione mediatica, cioè la dilagante incapacità di esprimere significati individuali spontanei, di scegliere parole non contaminate dai tormenti e dalla voracità delle simbolizzazioni di massa, ove tutto appare nella desolante luce del 'già detto': evoluzione grottesca della definizione heideggeriana del linguaggio come «casa dell'essere» (1946: 5). È chiaro che il fenomeno dell'opinione pubblica attinge esemplarmente a questa dimensionalità collettiva dell'immaginario: un secolo fa (era il 1922) Walter Lippman ne ha parlato in termini di «quadri mentali» (Pratkanis e Aronson, 1992: 64); Neil Smelser, più tardi, in termini di «credenze generalizzate» (*generalized beliefs*) (1962: 15-17).

Scopo di questo saggio è innanzitutto porre in luce come la dimensione immaginaria che investe il fenomeno dell'Intelligenza Artificiale rappresenti un caso esemplare del modo in cui oggetto di questo investimento simbolico è il progresso tecnico in generale. Tale esemplarità è esaltata dal fatto che ciò che denominiamo

ISSN: 2281-8138

col termine Intelligenza Artificiale, in quanto evoluzione della cibernetica e della teoria generale, dei sistemi alimenta le aspettative che nei più diversi ambiti applicativi il dinamismo tecnologico promette di soddisfare e sta mostrando di conseguire. Secondo obiettivo delle pagine che seguono, è sottolineare come questo processo espansivo ed evolutivo, proprio per le sue crescenti proporzioni renda evidente l'abisso che lo separa da ciò che non fa parte di esso ma di esso è presupposto teleologico, ossia la coscienza, che è qualcosa di diverso delle acquisizioni dei saperi 'intelligenti' della tecnica. Un certo spazio, nel corso del procedimento espositivo, sarà pertanto dedicato alla distinzione tra *conoscenza* e *coscienza*. Si mostrerà come il cammino dirompente dell'Intelligenza Artificiale rimarchi in questo senso il carattere di «problema difficile» che le scienze cognitive attribuiscono da tempo al fenomeno della coscienza e, conseguentemente, le esitazioni con le quali si guarda all'idea della progettabilità di una Coscienza Artificiale.

2. Immagine e oggetto

Ora, è chiaro che queste diverse modalità di impiego del concetto di *immaginario* (ma sono solo esempi), in tanto condividono la sua idoneità a designare entità in immagine, in quanto *immaginare*, innanzitutto, non è *percepire*. *Percepiamo*, condizionatamente, per via dei nostri sensi, il viso di Marylin ma non tutto il resto che quel viso richiama alla nostra mente e che recuperiamo subito grazie all'immaginazione, o meglio, grazie all'immaginario che quel volto attiva nella nostra memoria conservativa. In senso banale, qualcosa di simile accade quando percepiamo sensibilmente il lato di un solido ma non gli altri lati che non si vedono.

Può accadere che questo recupero non sia facile, né condiviso; in tal caso l'immaginario si riempie di oggetti diversi, a seconda della risorsa euristica che si pone al servizio dell'attività immaginativa, come nell'esempio dell'apologo dell'*Ignaro*, mediante il quale Galileo, nel *Saggiatore*, descrive i ripetuti stupori di chi, 'ascoltando' il ripetersi di un suono, scopre come esso provenga di volta in volta da fonti diverse, dal canto di un uccelletto allo zufolo di un pastore, dagli arpioni di sostegno del portale di una chiesa alle vibrazioni di un marranzano, dallo strofinio del bordo di un bicchiere al sibilo di un grillo (1622 [XXI]: 280-281). La 'morale della favola' è che, data la complessità del mondo, ognuno finirà con l'*immaginarsi* qualcosa di diverso, a seconda del proprio criterio («euristico») di autopersuasività (Kahneman e Tversky, 1984: 344-345).

Ne discende un importante risvolto ontologico, comune alle accezioni del termine «immaginario». Esso consiste in una specie di paradosso, illustrabile a partire da ciò che Sartre intende per «illusione di immanenza». Il fatto, cioè, che, seguendo Husserl, la coscienza sia sempre diretta a un oggetto, non è però da confondersi con una reificazione di quest'ultimo; perché l'oggetto resta del tutto assente, irreale, comparendo solo sottoforma di rappresentazione analogica, resa possibile da un rimando a familiarità cognitive che precedono e rendono possibile l'atto di coscienza

ISSN: 2281-8138

come loro sintesi (Sartre, 1940: 97). L'aspetto paradossale del prodotto di immaginazione è proprio in questo suo essere finzione di una verità ontica: ribadiamo, l'immagine non è conoscenza dell'oggetto, semmai si può dire la presupponga, ne erediti il sapere – quale esso sia – che lo riguarda e che consente di istituirlo neutralizzando la sua *assenza*. Sartre, seguendo Husserl, fa l'esempio del «Centauro che suona il flauto», precisando che in nessun modo esso può sostituirsi a un «*Erlebnis* psichico» (*esperienza vissuta*): esso – egli scrive – «non esiste né nell'anima, né *nella* coscienza, né altrove: esso è niente, è del tutto immaginazione» (1936: 137, corsivo aggiunto). La coscienza, insomma, non *contiene* l'immagine di qualcosa: essa *è* l'immagine.

Ora, di esempi del tipo del Centauro, se ne potrebbero fare innumerevoli, ma è importante considerare il presupposto cognitivo dell'immaginario, quale esso sia, cioè la sostanziale irriducibilità che separa la *coscienza* dalla *conoscenza* dei suoi oggetti (Sartre, 1943: 17), opponendo la capacità della coscienza di attribuire «senso» a entità sprovviste di «riferimento», sebbene il materiale del quale sono fatti questi oggetti, né più né meno come la materia di cui sono fatti i sogni, abbia come tale un «riferimento»[2]. Il Centauro chiamato in causa da Husserl resta una figura mitologica, reale letterariamente quanto Biancaneve; il flauto, poi, non è affatto privo di «riferimento»: è la raffigurazione associata dei due elementi ad apparirci fantastica. Già immaginario è un corpo metà uomo, metà cavallo; se poi lo si raffigura alle prese con uno strumento musicale, non resta che sorridere. Questa raffigurazione: la comprendiamo, le diamo un senso, che è il senso del suo 'non-senso', ma non è uno sguardo nel buio. E ciò si rende possibile perché essa si compone di cose conosciute, sebbene combinabili in un'immagine che può benissimo essere un nulla irreale (Parisi, 2012: 73). Per Freud, così è per i sogni, fatti di 'materia' reale, ricordi, espressioni, vicende, persone, ciononostante talvolta uniti in un miscuglio vertiginoso, grottesco, comico oppure angoscioso, ma sempre irreale. Nei sogni come nelle attività immaginative, insomma, compaiono solo frammenti di esperienza, come tali 'reali' (persone, cose, ambienti), ma il loro insieme non è quasi mai 'reale' come invece può esserlo un ricordo (sebbene i ricordi, a volte appaiano sconcertanti per la loro parziale attendibilità).

3. Intelligenza Artificiale e *infinito*

Su questo aspetto di indeterminatezza empirica e di connotazione fantastica dell'immaginario ha insistito a lungo Max Weber ravvisando, nel tema dell'infinito, il carattere intrinseco del progresso tecnico (1919: 17) quale movente di un'attesa, in parte riferita a evidenze note, in parte riferita alla loro oscurità futuribile. William Ogburn ha così ravvisato nel principio di continuità cumulativa il tratto specifico della

[2] La distinzione fra «senso» (*Sinn*) e «riferimento» (*Bedeutung*) è quella legata al tema fregeano (1892) che oppone gli enunciati categoriali (funzione *connotativa*), per richiamare attributi e proprietà descrittive associate a un termine rendendolo comprensibile, agli enunciati referenziali, che servono a *denotare* entità extralinguistiche, anche semplicemente denominandole (si veda a riguardo Dummet, 1973: xii-xiv).

ISSN: 2281-8138

modalità di mutamento della «cultura materiale» (1922). La tesi, che per comodità denominiamo di Weber-Ogburn, è che ogni condizione di stato della cultura materiale subirà le sorti recitate dal paradosso di Teseo: il suo essere, la sua identità, saranno sempre legati al tempo ma non si sa in che cosa si realizzeranno di volta in volta: ogni conquista della tecnica nasce già morta e il principio di tale indeterminatezza investe l'impossibilità di porre fine all'immaginario nel nome dell'idea di un 'compimento'. Il progresso tecnico è senza fine e ogni sua conquista è a sua volta priva di un fine che la giustifichi intrinsecamente. Questa sorta di assurdo destino si rivela mediante una irreversibile successione di stadi, ognuno dei quali supera esaustivamente e nello stesso tempo incorpora tutti quelli che lo hanno preceduto, gettandoli nell'oblio della loro ragione strumentale e consentendo loro, semmai, di sopravvivere come reperti storici, a disposizione del gusto estetico e dell'apprezzamento testimoniale del tempo perduto (Sonzogni, 2005).

Per questa via, com'è noto, Weber ha dato fiato alla sua presa di distanza da ogni filosofia della storia, incluse ovviamente le ingenuità escatologiche delle quali essa è stata investita, a esse contrapponendo la sua celebre definizione di «infinità priva di senso» (1904: 96). Di qui, la crescente attrattività prospettica di ciò che viene designato, oggi, con il termine Intelligenza Artificiale, quale evidente incarnazione dell'itinerario, ad un tempo promettente e oscuro, della modernizzazione tecnologica, del suo procedere 'senza fine'. Stupore, fascinazione estetica e preoccupazioni etiche sono le componenti principali (necessariamente multidirezionate) del travolgente immaginario, individuale e collettivo, che si appella, dilatandolo, a ogni esito incrementale della sua operatività e alla diffusa percezione dei più diversi ambiti nei quali si avvertono le sue ricadute, né più né meno in modo simile a quanto paventava Weber stesso al riguardo della sconfinata progressione culturale del «disincantamento» (1919: 20). Le sue sfere di designazione, i mondi – chiamiamoli così – della Artificialità, comprendono tutto ciò che promana in origine dall'universo dell'ICT e della tradizione cibernetica: l'ambito del *Machine Learning*, delle *Chatbot* e delle funzioni di *Virtual Assistant*, del marketing, della telechirurgia e della sanità in senso lato, della bioingegneria, della robotica industriale, dei controlli di sostenibilità e progettazione economico-ambientale, degli armamenti, della propulsione elettromeccanica.

L'apporto teoretico di Weber e di Ogburn alla comprensione di questo processo vale a mettere in luce, in particolare, la proprietà autocatalittica (autoreplicante) dell'immaginario associato alla percezione del dinamismo tecnologico, ossia alla percezione di indeterminatezza che alimenta i diversi orizzonti applicativi dell'innovazione artificiale. Ogni acquisizione ne annuncia un'altra, alimentando una promessa senza fine. La dimensione immaginaria, addirittura fantastica, è inevitabile perché il dinamismo tecnologico in atto è doppiamente anomico e intemperante: da un alto le singole innovazioni non hanno un fine, nel senso che rispondono compulsivamente a istanze performative incrementali direzionate ma prive di

N. 23 – Year XIII / July 2024 pp. 87-108
imagojournal.it
ISSN: 2281-8138

definizioni di stato[3]. A sua volta, l'intero cammino evolutivo della tecnica resta conseguentemente senza fine, in senso weberiano refrattario a qualsivoglia idea di destino. Ignoto e ricco di promesse: buio e luminoso al tempo stesso.

4. Intelligenza Artificiale e coscienza

Come in parte anticipato sopra, questo ordine di considerazioni riguarda però solo una delle due linee di tematizzazione inerenti all'immaginario del progresso tecnico in termini di percezione di infinito e indeterminatezza: quella propriamente denominata della «Intelligenza». In sintesi, guardando al panorama dei contesti applicativi dell'Intelligenza Artificiale, è chiaro che essi si presentano come terreno di sviluppo di programmi di lavoro di tipo simulativo finalizzati alla progettazione e alla costruzione di soluzioni capaci di abilità altrimenti irraggiungibili, per grado di complessità e quantità di informazione presa in carico (*Big Data*). La connessione tra obiettivi di programma e definizione preliminare dei compiti simulativi è data dall'impiego di algoritmi, di esecuzione e di apprendimento statistico, che rientrano nella struttura dei programmi di lavoro sviluppando la loro trasformazione digitale e accrescendone le abilità applicative: essi ampliano la potenza computazionale di tali programmi, senza ovviamente modificarne gli sviluppi attesi, se non entro alternative previste dal linguaggio strutturale del programma stesso.

Resta senz'altro dirompente e suscettivo di incredule rappresentazioni immaginifiche, il richiamo alla specificità tecnica dell'Intelligenza Artificiale, consistente nel suo superamento della condizione di ripetitività meccanica, attraverso forme di apprendimento riconducibili ai modelli denominati di *Machine Learning* o *Deep Learning*, a seconda che il procedimento artificiale riguardi la ricorsività di un algoritmo oppure le opzioni sequenziali legate alle sue variazioni parametriche. In pratica, la logica dell'algoritmo è isomorfica a quella di un sistema di equazioni strutturali, fra loro collegate in ragione di valori-soglia dal rispetto dei quali dipende il grado di ottusità, o meno, del sistema. Sta al linguaggio e alla versatilità del software variare sequenze e formato del programma in relazione alle variazioni dei parametri, superando l'ottusità e reindirizzando lo sviluppo di un algoritmo come se fosse intervenuto l'apprendimento di una nuova modalità di applicazione, adeguata alle modifiche del contesto. Ma si tratta di potenzialità già contenute nella logica del procedimento, anche se può trarre in inganno la portata della loro realizzazione (*performance*). Ciò che discrimina è l'evidenza che la mente individuale, da sola, non ci arriverà mai, come non arriverà mai a leggere – per esempio – tutto ciò che a livello

[3] Secondo Parsons è su queste basi che si rende palese la convergenza tra la nozione weberiana (sopra ricordata) di «infinito», associata al tema del progresso tecnico e quella che Durkheim impiega come sinonimo di *anomia* («passione dell'infinito») (1897: 310). L'autore americano menziona in questo senso le ossessive aspettazioni che, come nella competizione tecnologica, governano, invece della definizione dello «stato finale» degli atti compiuti, la loro *direzione*: i competitori dell'innovazione, non sanno che cosa vogliono, ma sanno solo qual è l'orizzonte senza fine della loro operosità (1951: 192-193, corsivo aggiunto). Una sorta di 'mitologia della nuova frontiera'.

ISSN: 2281-8138

planetario afferisce a un argomento, poi traendone l'ottimizzazione sintattico-lessicale 'incarnata' in un timbro vocale definito (*Chatbot*). Mediante un software sintatticamente evoluto e un sistema di sensori ad alta versatilità, anche il più analfabeta dei parlanti, pur mantenendo il suo inconfondibile timbro vocale, si esprimerà in forma video-plastica digitale molto attendibile, mediante un inglese oxfordiano di rara bellezza. Niente è da ritenersi non ottenibile, che non sia programmaticamente ottenibile. È insomma chiaro che l'innovazione e l'apprendimento stesso hanno effetti tanto più inattesi quanto più ampia sia la capacità del programma di prevedere e rimodulare la propria morfogenesi, entro i limiti di un dinamismo strutturale che, comunque, non può andare – insistiamo – al di là dei confini del suo perimetro simulativo. Sempre più largo, certo, ma sempre perimetro.

Veniamo così alla seconda linea di tematizzazione dell'immaginario. Si è fatto giustamente osservare, infatti, come razionalità e adeguatezza tecnica rivelino di essere ispirate a fini e motivazioni «che non possono essere il prodotto delle elaborazioni interne di un sistema chiuso» (perimetro algoritmico) (Manzotti e Tagliasco, 2002: 93). In altri termini, la definizione di contenuto degli scopi che dettano le operazioni strumentali a essi afferenti sono proprietà semantiche intenzionali, riferibili agli estremi di «un soggetto unitario, capace di fare esperienza di sé stesso e del mondo circostante (…), in grado di produrre motivazioni e fini, valori soggettivi e sensazioni. In poche parole, la coscienza» (*ibidem*)[4]. Ecco il secondo ambito dell'Immaginario. In ultima analisi, si sostiene al riguardo, «l'intelligenza ha cioè bisogno di una *volontà* che la guidi» (*ibidem*) e tale ragione teleologica non può che farsi valere al di là di ciò che in essa è comprovata adeguatezza strumentale e accurata istruzione protocollare. Seguendo il sentiero lungo il quale Weber ha sottolineato come «il corso di azione» ("l'agire reale") si svolga «nella gran parte dei suoi casi in una oscura semicoscienza o nell'incoscienza del suo senso intenzionato (…) istintivamente o in base all'abitudine», sicché «L'agire effettivamente, cioè pienamente e consapevole e chiaro, è in realtà sempre soltanto un caso limite» (1922b: 19), si può concedere che tale consapevolezza (*awareness*) resti una mera possibilità; il che significa, però, che questo secondo atto del pensiero «è già diverso dal puro pensare». Esso «è qualcos'altro. È un altro processo» (Leonardi, 1957: VIII). Un neurochirurgo sa ovviamente come procedere a una craniotomia e questa abilità competenziale può, in versione 'debole' essere supportata da risorse tele-endoscopiche o, in versione 'forte', sostitutivamente da un impianto robotizzato. Ma questa dimensione della decidibilità prognostica è estranea al suo eventuale interrogarsi sull'esito dell'operazione, quale ne sia il responso introspettivo, o sul giudizio relativo alla propria affidabilità.

[4] La distinzione concorda con la tesi di Searle («esperimento della stanza cinese»), secondo la quale i procedimenti dell'Intelligenza Artificiale sono capaci di sintassi, mentre la mente umana possiede anche una semantica, ossia è in grado di *comprendere* quello che sta facendo, manipolando simboli (come una macchina), però nello stesso tempo producendo significati" per questi simboli, cosa che la macchina non può fare (1992: 61).

È stato Sartre, su basi teoretiche non dissimili da quelle appena richiamate, a sottolineare così, ripetutamente, l'inconfondibilità del concetto di *coscienza* con quello di *conoscenza*, proponendo a carico del primo il ricorso all'esperire autoriflessivo, in luogo della categorizzazione oggettivante, che appartiene al secondo (Farina, 2013: 94)[5]. E c'è voluto più di un secolo, da Mary Shelley (1818) a Collodi (1883), a Isaac Asimov (1950) e Ridley Scott con *Blade Runner* (1982), poi Spielberg (2001), prima che la letteratura e la filmografia di fantascienza abbandonassero l'illusoria convinzione che gli esseri artificiali siano sempre dotati di coscienza: cioè, non di intelligenza ma – prima di tutto – della «capacità di essere soggetti autonomi di decisioni e obiettivi» (Manzotti e Tagliasco, 2002: 89). Solo con Kubrick (1968), il rifiuto del postulato antropomorfista, sostituito dall'idea di un essere artificiale, privo di corporeità e però dotato di coscienza ed emozioni similumane, solleverà il problema, ancora attuale, che spinge a ritenere l'annessione delle categorie del modello teleologico (scopi, valori), all'Intelligenza Artificiale e alla robotica, ammissibile soltanto metaforicamente (ivi: 94). Si tratta di argomento che va di pari passo con quello della legittimità di revocare in dubbio le interpretazioni naturaliste o internaliste della vita mentale (Sonzogni, 2022), da Chalmers segnalate come il «problema difficile» della coscienza: ovvero l'arduo compito di spiegare come processi meramente fisici siano in grado di dar conto della complessità e della autonomia ontologica dell'esperienza cosciente, nell'inevitabile rispetto della dualità cartesiana-hegeliana, magari poi con l'intento di dissolverla («eliminativismo»)[6].

Gilbert Ryle, tra gli antesignani di questo orientamento critico, per rimarcare l'insuperabilità di Cartesio, in un celebre contributo di settant'anni fa faceva rilevare – a titolo esemplificativo – che un conto sono le idee che uno 'ha in testa' (pura metafora), un conto i suoni che turbano l'udito e che si possono confermare con uno stetoscopio (significato letterale 'materialista') (1949: 33-35).

5. Immaginario «esternalista». Con alcune domande

Tutto ciò non vuol in nessun modo sollecitare allarmismi o – all'opposto – entusiastiche rivalutazioni di preminenza dell'umano sulla macchina, ma semplicemente invitare a un'ipotesi di lavoro che abbia come punto di partenza una distinzione: quella tra Intelligenza Artificiale e Coscienza Artificiale, tra *conoscenza* e *coscienza*. È chiaro che il cammino non è agevole: se è difficile immaginare che cosa sarà fra tre secoli un lap top, se ancora potrà chiamarsi così, quasi umanamente impossibile sembra immaginare una coscienza artificiale, una condizione di libero arbitrio, costruita (forse), ma senz'altro in nessun senso dipendente, dal *nostro* libero

[5] Le parole di Sartre sono in armonia anche con ciò che sostiene Freud quando, *nell'Interpretazione dei sogni*, afferma che la coscienza non è «nient'altro che un organo di senso per la percezione di qualità psichiche» (1899: 560). Su queste basi, è chiaro che la distinzione tra *conoscenza* e *coscienza* si ritrova anche nella irriducibilità di quest'ultima al concetto di «introspezione» (Ryle, 1949: 157-159).

[6] Si vedano, al riguardo, i termini della discussione a suo tempo enunciati in Agazzi (1967) e recentemente richiamati in Lavazza e De Caro (2013).

ISSN: 2281-8138

arbitrio. Difficoltà, questa, che ha suscitato autorevoli pronunciamenti sulla coscienza *tout court*: in termini di entità «ineffabile» (Rorty, 1982), di una 'cosa' che, sulle orme di Rainer Rilke, si rileva sia «più grande del cielo, più profonda del mare» (Edelman, 2004), o semplicemente «un mistero» (Quine, 1987; Fodor, 1998). Difficoltà, che resiste nel tempo, visto che, a quasi cinquant'anni dal celebre saggio di Warren Mc Culloch, che si domandava «perché la mente fosse nella testa» (1951), veniamo informati che il filosofo americano Stephen Stich ha ritenuto di rilanciare il problema dell'osservabilità della coscienza citando la vecchia storia dell'ubriaco che sta cercando le chiavi dell'auto sotto la luce di un lampione e risponde «laggiù nel parcheggio» a chi gli chiede dove le ha perse, mentre quando gli vien chiesto perché allora lui le cerchi proprio qui, risponde «Perché qui c'è la luce» (Carli, 1997: 211). Probabilmente, un buon punto di partenza è il dubbio, quanto meno il dubbio, che la coscienza abbia un *dove*.

Ora, un indirizzo teoretico che si muove in questa direzione, cioè dubitando della ammissibilità di un posizionamento spaziale della coscienza, ha preso forma muovendo proprio dalle posizioni assunte da David Chalmers a favore di una riformulazione interpretativa della vita mentale, nota come prospettiva «esternalista». Di fatto, il problema di un *dove*, chiedersi quale sia la sede (fisica) del pensiero, significa ammettere il primato materialista e quindi monista di una *res extensa*. Rifiutare l'interrogativo vuol dire accettare la prospettiva cartesiana e affermare che la mente non ha un *dove*, rinunciando all'ipotesi del posizionamento spaziale e abbracciando la prospettiva della cosiddetta «mente estesa» (Clark e Chalmers, 1998: 9; Amoretti, 2013: 228-229). In definitiva, mentre l'internalismo sostiene che la vita della mente dipende dal sistema nervoso centrale, l'esternalismo sostiene che essa dipende da qualcosa che non appartiene in esclusiva alla regione cerebrale, ma dinamicamente e unitariamente anche all'ambiente fisico e al mondo sociale, sicché non si può assumere una rigida separazione tra la mente stessa e ciò che è – in senso lato – il suo *contesto*. Anche se non appartiene alla schiera degli esternalisti, una concessione ad argomenti che si aprono a questa prospettiva è ravvisabile – sebbene con qualche oscurità – anche in Rorty, nella sua triade della modernità, centrata su atteggiamenti capaci di «solidarietà, ironia, contingenza» (1989).

In effetti, Chalmers perviene al modello della mente estesa dopo aver introdotto, in precedenza, una distinzione antinaturalistica (*The Two Mind-Body Problems*) che, accogliendo i progressi diagnostico-strumentali della neurofisiologia clinica, propone di riservare loro il campo dell'indagine sulla coscienza e sui suoi «correlati neurali» come *easy problem*, mentre disegna la sfera di un dominio di *hard problems*, ossia l'ambito della spiegazione scientifica inerente agli aspetti qualitativi e soggettivi dell'esperienza cosciente, da ritenersi incompatibili con le categorie di qualsivoglia monismo materialista (1996: 24-26). Tale problema, da alcuni autori menzionato come inerente alla categoria dei *qualia*, si basa sull'evidenza, posta in risalto dal filosofo John Searle, che «ogni stato cosciente è caratterizzato da una sensazione qualitativa», sia che si tratti di una percezione esterna, sia che si tratti di una rappresentazione cognitiva o del pensiero astratto (1998: 39).

N. 23 – Year XIII / July 2024 pp. 87-108
imagojournal.it
ISSN: 2281-8138

Coloro i quali, con Chalmers, condividono questo orientamento, in un certo senso si può dire che cerchino una via di mezzo, separando, da un lato la questione dei correlati neurali che si accompagnano alla effettualità delle idee, dall'altro le idee stesse che, nel loro contenuto, nulla hanno a che vedere con le proprietà della 'scatola cranica' (per usare l'espressione di Schopenhauer). Per questa via, volendo fare un esempio, è indiscutibile che l'eccitazione (o il suo contrario) produca effetti associati a un riscontro neurologico, ma resta fermo che il movente soggettivo dell'eccitazione stessa, l'attrazione contemplativa di un'immagine piuttosto che un'altra, appartiene a ordini di dominio fenomenico estranei a quel riscontro. Ad esempio, sarebbe veramente arduo ricondurre la percezione della bianchezza a una azione del cervello o il fascino estetico di un volto, confrontato con quello sconcertante di un altro, a qualcosa che abbia origine da entità sinaptiche[7].

Fatta salva questa inevitabile ma non dogmatica riaffermazione del dualismo classico (Bonolis, 2022: 68-69), restano ugualmente in gioco alcuni interrogativi, riguardanti il tema dei correlati analitici, chiamato in causa dalla nozione di «mente estesa». Se il richiamo alla ipotesi di una funzione di costitutività da parte del mondo esterno (ambiente fisico, mondo sociale) (Clark e Chalmers, 1998: 18) nei confronti della coscienza è degna di vera considerazione, l'eredità sociologica classica non consente di eludere la questione inerente al livello logico dell'opzione esternalista.

Durkheim ha convincentemente insistito sulla natura sociale – i suoi epigoni hanno aggiunto *culturale* – dell'individuo in quanto entità storicizzata (Bonolis, 2020: 48) e, in questo senso, evolutivamente formata: una sensibilità soggettiva, che non contrappone l'idea del libero arbitrio, l'idea paretiana di «integrità dell'Io» al mondo esterno, ma coglie un piano di indistinguibilità tra esterno ed interno: siamo nel nostro tempo, siamo 'il' nostro tempo, *siamo* storia. Ora, applicando il suggerimento tipologico di Susan Hurley, filosofa britannica, a proposito delle varietà referenziali dell'esternalismo (2010), questo livello *formale* del ragionamento durkheimiano, è altra cosa rispetto al piano *sostanziale* sul quale il grande autore francese, mediante la sua celeberrima elaborazione di una teoria della solidarietà («organica») corrispondente alla modernizzazione individuale, poneva l'esigenza di salvaguardia dell'individuo nei confronti di se stesso (*anomia*). Elaborazione, poi da Durkheim stesso abbandonata per via della palese impercorribilità del modello (Hinkle, 1960: 271). Il problema è che, sebbene formalmente e ineludibilmente l'individuo sia socialmente *costituito* (livello 'formale'), ciò non toglie che egli possa esprimere un atteggiamento di ostilità e di indifferenza nei confronti del suo mondo sociale, ossia che questo stesso mondo (livello 'sostanziale'), pur essendo *in* lui, alimenti desideri di

[7] Se si ignora questa barriera ontologica si va incontro a un'antinomia simile a quella narrata in *Das Parfüm* (Süskind, 1985). Il protagonista del romanzo rivela la condotta di un individuo (Jean-Baptiste Grenouille) che vive un rapporto con il mondo di tipo esclusivamente olfattivo, ma questo vincolo decade nel momento in cui egli coglie in una emanazione genitale femminile un sentore completamente avulso dalla sostanza organolettica (batterica) che attiva quella emanazione e che – invece – evoca in lui l'Immaginario di una fanciulla, ossia 'sapori' e 'odori', irrimediabilmente connessi a un'esperienza estetica. Potremmo dire: i profumi immaginari dell'amore, non della corporeità. Il protagonista, improvvisamente 'umano', testimonia della intrascendibilità del «sesto senso» al quale si riferisce la Arendt allorché, richiamandosi a Tommaso D'Aquino, denuncia i limiti dell'empirismo britannico (1971-1978: 133).

ISSN: 2281-8138

repulsione quando non di odio o efferato cinismo: tutte cose che hanno diversi nomi, da Auschwitz alla *acrasia* aristotelica, dalla «banalità del male», alla letale apatia e alla «indifferenza» umana, narrate in letteratura da Alberto Moravia (1949; 1960; 1965).

Non dovrebbero servire altri esempi.

D'altro canto, l'illusione solidaristica di Durkheim contiene un elemento irriducibile all'argomento della sua stessa fragilità teoretica o quanto meno non avulso dalla legittimità di un sentimento di speranza: a ciò, in ultima analisi, si ispirava la già accennata posizione di Richard Rorty e, prima di lui ma anche prima di Durkheim, la laconica esortazione di Dostoevskij, quando, nella metafora del «sottosuolo» narra del vissuto appartenente alla condizione dell'uomo solo, escluso da ogni vita di relazione e condannato alla propria indifferenza e alla propria misoginia. Un autentico inferno, che peraltro è sempre in agguato visto che in generale, nella solitudine, anche l'uomo dotato delle più alte virtù e qualità – egli scrive – e delle migliori fortune materiali «unicamente per il gusto di sbagliare, sarà capace di commettere una mascalzonata (...) e a bella posta desidererà la più dannosa delle sciocchezze» (1864: 229).

6. Una coscienza artificiale? L'immaginario impossibile

Resta però aperto un interrogativo, in sé poco rasserenante. Traendo spunto dalla inequivocabile sentenza di Dostoevskij, vi sarebbe ragione di chiedersi infatti se, allora, non sia proprio la solitudine, nella finzione narrativa di Kubrick, l'origine del proposito omicida di Hall 9000. L'assunzione di una condotta moralmente esecrabile verrebbe a dipendere dall'isolamento della coscienza, dalla sua condizione di naturale impossibilità di autogenerarsi. Seguendo le citate ispirazioni di Dostoevskij, di Rorty, di Durkheim, potremmo condividere con le insuperate parole di quest'ultimo grande autore la convinzione che l'individuo, «preso a sé, è troppo poca cosa, non è un fine sufficiente alla sua attività. (...) Quando non si hanno altri obiettivi all'infuori di noi stessi, non possiamo sfuggire all'idea che i nostri sforzi siano destinati, in fondo, a perdersi (...) Ma l'annullamento ci terrorizza e in tali condizioni non sapremmo trovare il coraggio di vivere, di agire, di lottare quando di tanta fatica nulla deve restare»» (1897: 250).

Nello stesso tempo, sottolineare così l'importanza del legame sociale non può significare che l'immaginario di una coscienza artificiale confidi sui benefici di un'emulazione, che annullerebbe la natura stessa dell'autonomia fenomenica della coscienza. Da un lato la coscienza è inseparabile dal suo presupposto dinamico, quello del libero arbitrio, ossia dal principio della libertà di scelta di fronte al bene e al male, come più volte ribadito da Sartre (1946). Dall'altro, l'insegnamento durkheimiano vale a sottolineare che è nell'esperienza dei legami e delle interazioni della vita sociale che si riempie di significati e simbolizzazioni la dimensione potenzialmente vuota della coscienza stessa, con esiti dinamicamente connessi a questa apertura 'estesa', nel senso di Chalmers. È chiaro che, date queste premesse, diventa difficile immaginare la via della costruzione di una coscienza artificiale e la

ISSN: 2281-8138

nostra riflessione, come quella di altri autori al riguardo, non può che proporre il terreno di una ricerca degna di questo nome: il postulato del libero arbitrio è più che un'opera aperta. Per definizione, può portare ovunque e c'è da sperare che l'intersoggettività della vita associata proponga nei fatti, evolutivamente, abiti e modelli di pensiero intanto condivisi, in quanto desiderabili. Ma senza che, antinomicamente, il libero arbitrio venga insegnato e appreso, dando luogo ai paradossi già impietosamente illustrati da Hegel (del tipo: 'amami spontaneamente, non perché te lo chiedo io'). Dell'impossibilità di comporre questo dilemma, ha dato conto proprio Durkheim nella delineazione dell'ipotesi di una morale laica, capace di sostituire la ragione all'assoluto delle ontologie religiose, così residuando però la spinosa questione di capire come potesse prendere forma una siffatta nuova eticità, oltretutto a fronte delle tetre nuvolosità culturali dei primi del Novecento (la Francia dell'affare Dreyfus e, non molto tempo dopo, la nascita del Nazionalsocialismo in Germania).

L'interrogativo è sinteticamente lontano da ogni possibilità di immaginazione.

È forse difficile ma non più di tanto impossibile coltivare l'immaginario di una piena sostituibilità tecnologica e di una vera scomparsa di intere branche della medicina chirurgica, così come della viabilità stradale orizzontale. Ben diverso è l'immaginario che abbia come oggetto la decidibilità, su basi metacognitive, degli atti che danno luogo e ottemperano al pensiero del progresso tecnico: diverso, cioè, è l'immaginario di una coscienza, artificiale, che guardi – in quanto tale – ma non appartenga, agli orizzonti dell'intelligenza – essa sì – plausibilmente artificiale.

Gli enunciati richiamati nel precedente paragrafo sottolineano gli estremi di una distinzione che può allora essere accolta rilevando che il livello logico della nozione di *coscienza* stia a quello della nozione di *conoscenza* (che coincide con il dominio dell'Intelligenza Artificiale) come la funzione metalinguistica sta a quella della funzione referenziale, quali proprietà del linguaggio articolato. La separazione delle due funzioni, che risale a Karl Bühler, ripresa da Roman Jakobson e poi da Popper (1978: 42), ammette ovviamente che in entrambi i casi il linguaggio sia veicolo istitutivo del suo oggetto, rappresentandolo ai parlanti anche quando non è materialmente nella loro prossimità (come quando nel cuore di una distesa sahariana si discutesse di orsi polari) o addirittura per negarne l'esistenza[8]. Solo che quando la funzione è metalinguistica o, se vogliamo, autoreferenziale, l'oggetto è il linguaggio stesso; il che accade, ad esempio, per *spiegare* il significato di una parola o di una regola grammaticale. Il linguaggio qui *parla* del linguaggio. In questo senso, similmente ai termini di questa distinzione, la logica della conoscenza è invece di tipo osservativo, anche quando il suo scopo è pratico, come nell'Intelligenza Artificiale, perché il suo oggetto è extralinguistico, si trattasse pure di analisi, non del linguaggio, ma del comportamento linguistico o della comunicazione verbale. L'aspetto extralinguistico si riferisce al fatto che il rapporto tra linguaggio e oggetto è qui

[8] Non c'è al riguardo espressione simbolica migliore di quella, celeberrima, che risale ad Heidegger, che Heidegger stesso considera quale paradosso di riconoscimento ontico di ciò che si afferma ne sia privo per statuto semantico (esempio estremo del *niente*) (1946: 21-23).

attuato mediante nesso referenziale. La logica della coscienza resta invece di tipo normativo-riflessivo: non persegue, né postula un'imputazione causale.

Come abbiamo sottolineato in precedenza, a proposito della distinzione tra il pensiero e la consapevolezza del pensare (paragrafo 5), in senso figurato conoscenza e coscienza rimandano all'uso incomparabile di due vetri davanti ai nostri occhi: l'istanza della conoscenza richiede di guardare il mondo attraverso trasparenze, più o meno nitide, più o meno deformanti, anche se trattasi solo di comprendere le origini del proprio mal di testa. L'istanza della coscienza richiede un rispecchiamento: non di attraversare il vetro per andare oltre, analizzando un oggetto, ma di dedicarsi alla propria immagine riflessa, in una *visio* esclusivamente interiore, alla ricerca del senso che attribuiamo al guardare, mettendo tra parentesi ciò che si vedrebbe sostituendo lo specchio con una finestra per guardare al di là.

Sul piano analitico, essi (i due concetti) si oppongono omologicamente a ciò che rende inconfondibili giudizi di fatto da giudizi di valore, dove, per definizione, l'evidente simulabilità computazionale di ciò che appartiene all'ambito dei primi non è attuabile nel campo dei secondi (Sonzogni, 2022). Quando si tratta di spiegazione scientifica o di progettazione empirica conseguente (progettare un ponte richiede disporre di saperi concernenti l'aerostatica, i calcoli del cemento armato, informazioni geofisiche e mille altre acquisizioni afferenti), i risultati, attesi necessariamente nel contesto del *probabile*, hanno validità accettata fino a prova contraria, cioè condivisa e intersoggettivamente plausibile, ma mai assoluta. L'argomento rientra notoriamente nelle categorie della nozione kuhniana di «paradigma» (1962). Il consueto esperimento didattico della rifrazione della luce non è in nessun senso una verità ontica ma resta quanto di meglio si reputa consolidato, nei limiti weberiani di un sapere necessariamente «attuale» (Weber, 1904: 59), per comprendere l'arcana percezione del bastoncino spezzato che affiora dall'acqua.

Quando, invece, ci si muove sul terreno del *certo*, dell'«etica della convinzione» e del dominio dell'imperativo sintetizzabile nel domandarsi, non «come posso conoscere» bensì «come devo agire»[9], è invece del tutto improprio (e inutile) servirsi delle connessioni illustrate da una o più equazioni strutturali (vedi *supra*): qui ci si sta chiedendo altro e cioè se sto sentendo il suono delle campane o no, se percepisco sensibilmente una sollecitazione corporea o, ancora, se faccio mia una credenza, un precetto normativo, un determinato criterio etico al quale fare riferimento per agire, per orientare miei determinati atti. È chiaro che, nel caso dei giudizi di valore, il *certo* della consultazione riflessiva e, analogamente, dell'intento che è alla base dell'introspezione, in ultima analisi non è più incontrovertibile del *probabile* che caratterizza la conoscenza, empirica o teorematica[10], ma sconta tale limite dal lato

[9] Sull'ispirazione kantiana della distinzione analitica dei concetti di fatto e valore, e delle due nozioni di «etica» (della *responsabilità* e della *convinzione*) ad essi omologhe, il riferimento è ovviamente a Weber (1917), fatte salve le riserve espresse al riguardo di quella distinzione, dai più credibili esponenti del pragmatismo americano (Putnam, 1990: 61). La distinzione, concettualmente affine, tra logica del *certo* e logica del *probabile* è estesamente esposta da Sartre ne *L'imaginaire* (1940).

[10] Mentre una *verità sintetica* trova il suo limite nella eventualità di una prova contraria, i confini di una *verità analitica* (ad esempio nel campo della geometria) restano quelli assiomatici del procedimento che istituisce l'assioma stesso (Agazzi, 1961; Boudon, 1968: 22-23).

ISSN: 2281-8138
CC BY

della sua non esaustività in termini di coincidenza con l'azione conseguente. Il problema centrale della coscienza è in gran parte tutto qui, in questa non coincidenza, quando il *certo* si infrange contro la debolezza della volontà, l'inganno muto e desolante dell'acrasia. Voci autorevoli della filosofia analitica ne hanno dato un'esemplificazione inimitabile, al riguardo, citando la figura di Pietro, che sa perfettamente che cosa vuol dire tradire e cionostante non sa astenersi dal farlo, sicché 'prima che il gallo canti', per tre volte lo farà (Anscombe, 1957: 154; Davidson, 1980: 88).

In generale, la tradizione filosofica ha dato più volte conto di questa inefficacia pragmatica dei giudizi di valore utilizzando l'apologo dei disturbi che, nel modello aristotelico dell'inferenza pratica, interrompono la relazione di continuità tra le premesse del sillogismo e le sue *conclusioni*, quando queste sono di ordine pratico o, ancora, quando ciò che conclude l'inferenza non è una *consecutio* logico-proposizionale ma un'azione (*Etica Nicomachea*, Libro VII, 3). La pretesa che il sapere inerente a un campo d'azione possa dar luogo a una condotta d'azione ad esso ottemperante è il più delle volte destinata al fallimento. Le cose non accadono come quando l'idraulico, convocato in emergenza, mette in atto il suo sapere per sturare un lavandino. Spesso intercorrono fattori di «disturbo» (*Etica Nicomachea*, Libro VII, 3) del tipo di quelli emozionali e passionali che per Weber depongono a favore della impossibilità di non essere irrazionali (1906: 157) e, all'estremo, a favore della ricorrente eventualità che, sapendo distinguere tra il Male e il Bene, si imbocchi ciononostante la prima via, del tutto «acratica», sempre nel senso di Aristotele.

Sta di fatto che, così, tra premesse di conoscenza competenziale e agire conformemente a esse, possono intercorrere distorsioni della conformità, attesa o desiderata, che la vanificano totalmente o in parte. D'altra parte, è chiaro che le fonti di irrazionalità o di mancata «adeguatezza» nel rapporto mezzi-fini, per usare un'altra espressione di Weber, sono potenzialmente innumerevoli: vanno da quelle concernenti gli stati mentali del soggetto agente a quelle che investono le condizioni meramente pratiche, sopravvenienti, dell'adeguatezza., Inoltre, si rende altresì discutibile l' attendibilità dei saperi che intrinsecamente sarebbero guida dell'azione: chi fuma, compresi i medici che fumano, sa del danno cancerogeno cui va incontro, ma qual è l'esaustività di tale sapere? In certi momenti il fumo aiuta, sul piano nervoso e su quello psicologico e questo, a suo modo, è un altro sapere. È vero che poi le cose si complicano se entrano in gioco condizioni di dipendenza neurale, per cui il sapere di prevenzione diventa un pallido ricordo. Ma per arrivare a questo stadio bisogna averne superati alcuni nei quali il sapere era già pallido, quanto meno di debole efficacia, il che porta ritenere che la conoscenza del danno cardiocircolatorio non esaurisca tutto quello che c'è da sapere per decidere di fumare o di continuare a farlo.

E torniamo così allo spettro delle cose cattive che però, seguendo Mandeville, producono effetti non necessariamente esecrabili: ad esempio, cose che, come il fumo, attenuano il 'male di vivere', del quale ha parlato Montale, con toni non dissimili da quelli della descrizione dell'ossessione del dolore della vita, in Schopenhauer.

ISSN: 2281-8138

Da queste considerazioni, il tema dell'autonomia del concetto di coscienza guadagna evidenti ragioni di supporto e può ulteriormente rafforzare le espressioni di indeterminatezza che – come si è visto – da varie fonti si accompagnano al tentativo di identificarne lo statuto fenomenico. D'altra parte, le incertezze che investono l'argomento e delle quali l'ipotesi della mente estesa fornisce un'indubbia e feconda rappresentazione, ma non un contributo risolutivo dal lato della progettazione della artificialità della coscienza, restano insuperate anche per via di una plausibile lettura del concetto di *acrasia* come chiave del dinamismo enigmatico della coscienza e come fondamento esemplare del libero arbitrio. Prima dell'esplicita posizione assunta in questo senso da John Searle (1980) e ancora prima da Sartre (1946), merita che si rammenti al riguardo l'ostentata paura nei confronti della dimensione del *certo*, ostentata da Nietzsche (1965: 19) con un atteggiamento chiaramente precursore del fascino dell'*indeterminatezza* nella teoria matematica di Shannon e Weaver: del loro richiamo all'attrattiva dell'informazione, tanto più 'vera' ove densa di ciò che si potrebbe dire, in luogo della noia di ciò che si sa e si dice (1949: 8). Per questo, sembra lecito ribadire che non c'è alcuna possibilità, al riguardo, di ricorrere a una qualche forma di sapere consolidato, essendo il problema principale proprio quello della identità fenomenica della coscienza, ossia di guardare a una connettività in grado di coglierne l'indeterminatezza. Difficile, al punto da ritenersi inimmaginabile, salvo fantasie cinematografiche che saremmo ansiosi di prendere in considerazione, ipotizzare il modo in cui la coscienza entri in gioco nelle costituenti di un profilo robotico, da un lato rendendosi chiara la sua autonomia dai saperi 'intelligenti', dall'altro mostrandosi intelligibile la sua 'condanna alla libertà' e il suo paradossale diritto di abbandonarsi al Male.

7. Epilogo e dilemmi

In conclusione, è sostenibile che il problema della coscienza, in relazione al suo posizionamento nell'ambito delle neuroscienze, come nello specifico delle voci dell'immaginario sull'Intelligenza Artificiale, ribadiamo, resti quello che, condivisibilmente, ormai vent'anni fa è stato definito da David Chalmers «un problema difficile» (1995: 201). Il sentiero di ricerca della naturalizzazione è sempre attraente, nel senso che «se si riuscisse davvero a chiarire quali processi cerebrali siano sottesi alla comparsa della coscienza, sarebbe del tutto irragionevole continuare a credere che la coscienza stessa non sia una funzione biologica attivata, al pari di molte altre, da normali processi cerebrali» (Nannini, 2014: 332). Fino ad allora, restano però tanti i motivi per non abbandonare la diagnosi, per quanto laconica, di Chalmers: qui si vuole sottolineare quello che tra essi si presenta in una particolare forma dilemmatica e che, allo stato attuale e che allo stato attuale del dibattito su questi temi, non sembra consentire qualcosa di più di un'ipotesi di ricerca. Va tenuto conto che, come in tutti questi casi, il dilemma è cornuto, anche se l'espressione, ispirata a Giovanni Buridano, quasi volgare, resta molto suggestiva: è da rimarcare che, in sostanza, la difficoltà, il 'difficile problema' di concepire una

ISSN: 2281-8138

coscienza artificiale riposa sulla proclamata evidenza di non riuscire a offrire una accettabile definizione del fenomeno della coscienza *tout court*, preliminarmente alla questione del perseguimento o della progettazione della sua artificialità. Il che significa che la coscienza stessa potrebbe rivelarsi un'entità esitante, in balia dello spettro acratico del tipo di quello che un altro detto popolare, coniato da un giornalista italiano vissuto nella prima metà del Novecento, menziona la degenerazione della debolezza umana nei termini di un soggetto 'buono a nulla e capace di tutto' e che Dante, forse anche Shakespeare, non riterrebbe poi tanto esagerato al di là della ricorrente mania di pensare che solo gli italiani siano così.

Sono le incertezze e le edulcorate santificazioni della categoria di cui ci stiamo occupando, tra filosofi e neuroscienziati eminenti, a sollecitare i termini di una referenzialità cosmopolita: anche un'ipotetica coscienza artificiale porterebbe con sé lacrime e sangue, analogamente alla storia della presenza dell'uomo nel mondo. Al cospetto di un'ipotesi del genere, già Durkheim aveva abbandonato la fiducia nell'idea, di origine comtiana, in una «morale laica» che, in epoca moderna, fosse in grado di sostituirsi allo spirito religioso (credenze nel soprannaturale), in virtù dell'educazione intellettuale e della sua persuasività circa l'importanza della disciplina e della solidarietà umana disinteressata (1925). Com'è noto, Durkheim lascerà poi cadere l'argomento, peraltro molto fecondo nella continuità dei suoi epigoni, in base alle stesse riserve con le quali in precedenza egli stesso aveva preso le distanze dall'utopismo delle tesi hobbesiane sul Leviatano e da quelle rousseauiane sulla *volonté générale* (1895: 114): in definitiva, il senso delle sue riserve era che, passi l'idea di educare un popolo, ma quali garanzie offrono al riguardo gli educatori?

In alternativa, le strade (dilemmatiche) sono due. Cercare di superare la distinzione tra coscienza e conoscenza, incorporando la prima nella seconda (Brooks, 1991), andando però incontro ad aporie delle quali s'è detto e che è lecito dubitare riescano a trattare la coscienza in termini di programmazione cognitiva (equazioni e *switch* di logica deontica). Oppure, pur non sapendo che cosa essa sia, nel naturale come nell'artificiale, la coscienza resta un enigma, ossia il regno dell'indeterminatezza e non c'è allora altra strada che affidare l'artificiale all'umano-naturale, alla guida di chi governa ogni programmazione dei suoi algoritmi (Intelligenza Artificiale), senza ambizione alcuna di ordine formativo o didattico. Ma ciò non può convincere più di tanto: forse finirà così, come per i conflitti bellici, gli armamenti, i rapporti di forza tra nazioni e governi, le efferatezze dittatoriali. La coscienza non artificiale, umana, non può farci dormire tranquilli. Il fatto è che, nel nostro Immaginario, non trova posto qualcosa di diverso.

ISSN: 2281-8138

Bibliography

Agazzi E. (1961), *Introduzione ai problemi dell'assiomatica*, Milano, Vita e Pensiero.

Agazzi E. (1967), "Alcune osservazioni sul problema dell'intelligenza artificiale", in *Rivista di Filosofia neoscolastica,* 59 (1): 1-34.

Amoretti M. C. (2013), "Sui correlati fisici della coscienza: neurali o estesi?", in Nannini S e Zeppi A. (a cura di), *La mente può essere naturalizzata?* Colle di Val d'Elsa, Salvietti & Barabuffi, 2013, 217-233. ISBN: 978-88-97082-49-1.

Anscombe G. E. M. (1957), *Intention*, Oxford, Basil Blackwell; tr. it., *Intenzione*, Roma, EDUSC, 2016.

Arendt H. (1971-1978), *The Life of the Mind*, New York, Harcourt, Brace, Jovanovich, 2 voll.; tr. it., *La vita della mente*, Bologna, il Mulino, 1987-2009.

Aristotele (IV a.C.), *Etica Nicomachea*, C. Mazzarelli (a cura di), Milano, Bompiani, 2000.

Bonolis M. (2020), "La mente umana e la volontà di credere", in Bonolis M. (a cura di) (2020), *Lo «spirito» nell'anima. Indagine sulla spiritualità giovanile*, Milano, Angeli, 17-77.

Bonolis M. (2022), "La naturalizzazione della mente e il suo contrario", in Bonolis M. e Lombardo C. (a cura di) (2022), *Sociologia degli stati mentali (Teoria e ricerca)*, Angeli, Milano, 19-81.

Boudon R. (1968), *"A quoi sert la notion de 'Structure'?"*; tr. it., *Strutturalismo e scienze umane*, Torino, Einaudi, 1969.

Brooks R. A., (1991), "Intelligence without representations", in *Artificial Intelligence Journal*, 47(1-3) (January), 139-159.

Carli E. (a cura di) (1997), *Cervelli che parlano. Il dibattito su mente, coscienza e intelligenza artificiale*, Milano, Mondadori.

Chalmers D. (1995), "Facing Up the Problem of Consciousness", in *Journal of Consciousness Studies*, 2 (3), 200-219.

Chalmers D. J. (1996), *The Conscious Mind. In Search of a Fundamental Theory*, Oxford, Oxford University Press.

Clark A. (2003), *Natural Born Cyborgs: Technologies and the Future of Human Intelligence*, Oxford, Oxford University Press.

N. 23 – Year XIII / July 2024 pp. 87-108
imagojournal.it
ISSN: 2281-8138

Clark A., David J. C. (1998), "The Extended Mind", *in Analysis*, 58: 7-19, DOI:10.2307/3328150.

Crespi F. (1993), *Evento e struttura. Per una teoria del mutamento sociale*, Bologna, il Mulino.

Davidson D. (1980), *Actions and Events. Philosophical essays*, Oxford, Oxford University Press; tr. it., *Azioni ed eventi*, Bologna, il Mulino, 1992.

Dostoevskij F. M. (1864), *Zapiski iz podpol´ja*, Letchworth, Bradda Books; tr. it., *Memorie dal sottosuolo*, in (Id.) *Il romanzo del sottosuolo*, Torino, Einaudi, 1974, 203-322.

Dummett M. A. E. (1973), *Frege: philosophy of language*. New York, Harper & Row.

Durkheim É. (1925), *L'Education morale*, Paris, Félix Alcan; tr. it., *L'educazione morale*, in *Il suicidio e l'educazione morale*, Torino, Utet, 1969.

Durkheim É. (1895), *Les règles de la méthode sociologique*, Paris, Félix Alcan; tr. it., *Le regole del metodo sociologico – Sociologia e filosofia*, Milano, Comunità, 1963.

Durkheim È. (1897), *Le Suicide. Etude de sociologie*, Paris, Félix Alcan; tr. it., *Il suicidio*, in *Il suicidio e L'educazione morale*, Torino, Utet, 1969.

Edelman G. M. (2004), *Wider than the Sky: The Phenomenal Gift of Consciousness*, New Haven, Connecticut, Yale University Press; tr. it., *Più grande del cielo. Lo straordinario dono fenomenico della coscienza*, Torino, Einaudi, 2004.

Farina G. (2013), "Coscienza e conoscenza: Sartre lettore di Freud", in *Bollettino Studi sartriani*, IX: 89-109.

Fodor J. A. (1998), *In Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind*, Cambridge, Massachusetts, MIT Press.

Frege G. (1892), *«Über Sinn und Bedeutung», Zeitschrift für Philosophie und philosophische Kritik*, n. 100, 25-50; tr. it., *Senso, funzione e concetto: scritti filosofici 1891-1897*, Roma, Laterza, 2001.

Freud S. (1899), *Die Traumdeutung*, Lipsia-Vienna, Franz Deuticke Verlag; tr. it., *L'interpretazione dei sogni* (*Opere*, vol. III), Torino, Boringhieri, 1966.

ISSN: 2281-8138

Gadamer H. G. (1966), *Die Universalität des hermeneutischen Problems*, Tubingen, Mohr; tr. it., *L'universalità dell'ermeneutica*, in A. Babolin (a cura di), *Filosofi tedeschi d'oggi*, Bologna, il Mulino, 1967.

Galilei G. (1622), Il *Saggiatore* (1890-1909), in Galilei G. (1890-1909), *Le opere di Galileo Galilei* (a cura di Antonio Favaro), Edizione Nazionale sotto gli auspici di S. M. il Re d'Italia, Barbera, Firenze (20 volumi). VI: 197-372.

Heidegger M. (1946), *Brief über den «Humanismus»*, Frankfurt am Main, Klosterman; tr. it., *Lettera sull'umanismo,* Milano, Adelphi, 1995.

Hinkle R. C. jr. (1960), "Durkheim in American Sociology", in Kurt H. Wolf (ed. by) (1960), *Essays on Sociology and Philosophy (by Emile Durkheim et al.)*, New York-London, Harper & Row (1964), 267-295.

Hurley S. (2010), "The varieties of externalism", in Richard Menary (ed. by) (2010), *The extended mind*, Cambridge, Massachusetts, MIT Press.

Jung C. G. (1934-1954), "Über die Archetypen des kollektiven Unbewussten", in Jung C.G., *Gesammelte Werke Band 9.1*, Olten-Zürich, Walter Verlag; tr. it., *Gli archetipi dell'inconscio collettivo*, Torino, Bollati Boringhieri, 1977.

Kahneman D., Tversky Amos (1984), "Choices, values, and frames", *American Psychologist*, 39: 341-350.

Kuhn T. (1962), *The Structure of Scientific Revolutions*, Chicago, The University of Chicago Press; tr. it., *La struttura delle rivoluzioni scientifiche*, Torino, Einaudi, 1969.

Lavazza A., De Caro M. (2013), "Neuroetica e neuroscienza: tra spiegazione della vita e cura della mente", *Rivista Internazionale di Filosofia e Psicologia*, 4(3): 249-251. DOI:10.4453/rifp.2013.0026

Leonardi F. (1957), *Introduzione allo studio del comportamento sociale*, Milano, Giuffré.

Manzotti R., Tagliasco V. (2002), "Si può parlare di coscienza artificiale?", *Sistemi intelligenti*, 1: 89-107.

Mc Culloch W. S. (1951), "Why the Mind is in the Head", in Jeffress, Lloyd Alexander (*ed. by*.), (1951), *Cerebral Mechanisms in Behavior*, New York, Wiley; tr. it., *Perché la mente è nella testa?*, in Karl S. Lashley *et al.* (1969) *La fisica della mente, Torino,* Boringhieri, 112-136.

Moravia A. (1949), *Gli indifferenti*, Milano, Bompiani.

ISSN: 2281-8138

Moravia A. (1965), *L'attenzione*, Milano, Bompiani.

Nannini S. (2014), "La seconda svolta cognitiva in filosofia della mente", *Reti, saperi, linguaggi - Italian Journal of Cognitive Sciences*, I (2), 320-334.

Nietzsche F. (1965), *Epistolario (1865-1900)*, a cura di Barbara Allason, Torino, Einaudi (1969).

Ogburn W. (1922), *Social Change. With Respect to Culture and Original Nature*, New York, Huebesch.

Parisi I. (2012), "L'immaginario", *Im@go. Rivista di studi sociali sull'immaginario*, I: 71-74.

Parsons T. (1951), *The Social System*, New York, Free Press of Glencoe; tr. it., *Il sistema sociale*, Milano, Comunità, 1965.

Popper K. R. (1978), *The Place of Mind in Nature*, San Francisco, Harper & Row; tr. it., *Il posto della mente nella natura*, in *Tre saggi sulla mente umana*, Roma, Armando, 1994.

Putnam H. (1990), *Realism with a Human Face*, Cambridge (Mass.), Harvard University Press; tr. it., *Realismo dal volto umano*, Bologna, il Mulino, 1995.

Pratkanis A. R., e Aronson E. (1992), *Age of Propaganda. The Everyday Use and Abuse of Persuasion*, New York, Henry Holt and Company; tr. it., *Psicologia delle comunicazioni di massa. Usi e abusi della persuasione*, Bologna, il Mulino, 1996.

Quine W. v. O. (1987)*, Quiddities*: An *Intermittently Philosophical Dictionary*, Belknap Press of Harvard University Press; tr. it., *Quidditates: quasi un dizionario filosofico*, Milano, Garzanti, 1991.

Rorty R. (1982), "Mind as Ineffable", in Elvee Richard (*ed. by*) (1982), *Mind in Nature*, San Francisco, Harper and Row, 60-95.

Rorty R. (1989), *Contingency, Irony, and Solidarity*, Cambridge, Cambridge University Press; tr. it*., La filosofia dopo la filosofia: contingenza, ironia e solidarietà*, Roma-Bari, 1989.

Ryle G. (1949), *The Concept of Mind*, London, Hutchinson & Co. Ltd; tr. it., *Il concetto di mente*, Roma-Bari, Laterza, 2007.

Sapir E. (1929), "The Unconscious Patterning of Behavior in Society", in Ethel S. Dummer (*ed. by*) (1929), *The Unconscious: A Symposium* New York, Knopf, 114-142.

ISSN: 2281-8138

Sartre J-P. (1936), *L'Imagination*, Paris, Puf; tr. it., *L'immaginazione – Idee per una teoria delle emozioni*, Milano, Bompiani, 1962, 5-151.

Sartre J-P. (1940), *L'imaginaire. Psychologie phénoménologique de l'imagination*, Paris, Gallimard; tr. it., *Immagine e coscienza. Psicologia fenomenologica dell'immaginazione*, Torino, Einaudi, 1960.

Sartre J-P. (1943), *L'Être et le Néant: essai d'ontologie phenomenologique*, Paris, Gallimard; tr. it., *L'essere e il nulla*, Milano, Il Saggiatore, 1965.

Sartre J-P. (1946), *L'Existentialisme est un humanisme*, Paris, Les Editions Nagel S.A.; tr. it., *L'esistenzialismo è un umanismo*, Milano, Mursia, 1970.

Searle J. R. (1992), *The Rediscovery of the Mind*, Cambridge-London, The MIT Press; tr. it., *La riscoperta della mente*, Torino, Bollati Boringhieri, 1994.

Searle J. R. (1998), *Mind, Language and Society. Philosophy in the Real World*, New York, Basic Books; tr. it., *Mente, linguaggio, società. La filosofia nel mondo reale*, Torino, Rosenberg & Sellier, 2009.

Shannon C. e Elwood W. W. (1949), *The Mathematical Theory of Communication*, Urbana (Illinois), University of Illinois Press; tr. it., *La teoria matematica delle comunicazioni*, Milano, Etas Libri, 1971.

Smelser N. J. (1962), *Theory of Collective Behavior*, New York: Free Press.

Sonzogni B. (2005), "Eterogeneità globale e modernizzazione. Attualità di Ogburn", *Sociologia e ricerca sociale*, 78: 77-116.

Sonzogni B. (2022), "Simulabilità computazionale degli stati mentali. Uomo e macchine alla luce di alcuni dualismi interpretativi", in Bonolis M. e Lombardo C. (a cura di) (2022), *Sociologia degli stati mentali (Teoria e ricerca)*, Milano, Angeli, 185-201.

Süskind P. (1985), *Das Parfum – Die Geschichte eines Mörders*, Zürich, Diogenes; tr. it., *Il profumo*, Milano, Longanesi, 1985.

Weber M. (1919), *Wissenschaft als Beruf*, Tübingen, Verlag J.C.B. Mohr; tr. it., *La scienza come professione*, in *Il lavoro intellettuale come professione*, Torino, Einaudi, 1948, 3-43.

Weber M. (1904), "Die «Objektivität» sozialwissenschaftlicher und sozialpolitischer Erkenntnis"; tr. it., *"L'«oggettività» conoscitiva della scienza sociale e della politica sociale"*, in Weber (1922a), *Gesammelte Aufsatze zur Wissenschaltslehre*, Tübingen,

N. 23 – Year XIII / July 2024 pp. 87-108
imagojournal.it
ISSN: 2281-8138

Verlag J.C.B. Mohr; tr. it., *Il metodo delle scienze storico-sociali*, Torino, Einaudi, 1958 (1967), 53-141.

Weber M. (1906), "Kritische Studien auf dem Gebiet der Kulturwissensschaftlichen Logik", in M. Weber (1922); tr. it., *Studi critici intorno alla logica delle scienze della cultura*, in *Il metodo delle scienze storico-sociali*, Torino, Einaudi, 1958 (1967), 143-237.

Weber M. (1922b), *Wirtschaft und Gesellschaft*, Tübingen, Verlag J.C.B. Mohr; tr. it., *Economia e società*, Milano, Comunità, 1961.

N. 23 – Year XIII / July 2024 pp. 87-108
imagojournal.it
ISSN: 2281-8138