

Chiara Lucifora,
Carmelo Mario Vicario

IL CERVELLO MORALE

Dalle Scienze Cognitive
all'Intelligenza Artificiale

Prefazione di
Giuseppe Craparo



per coltivare le conoscenze

FrancoAngeli semi

Informazioni per il lettore

Questo file PDF è una versione gratuita di sole 20 pagine ed è leggibile con **Adobe Acrobat Reader**



La versione completa dell'e-book (a pagamento) è leggibile **con Adobe Digital Editions**.

Per tutte le informazioni sulle condizioni dei nostri e-book (con quali dispositivi leggerli e quali funzioni sono consentite) consulta [cliccando qui](#) le nostre F.A.Q.

Semi

Per coltivare le conoscenze

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati
possono consultare il nostro sito Internet: www.francoangeli.it
e iscriversi nella homepage
al servizio "Informatemi" per ricevere via e-mail le
segnalazioni delle novità.

Chiara Lucifora,
Carmelo Mario Vicario

IL CERVELLO MORALE

Dalle Scienze Cognitive
all'Intelligenza Artificiale

Prefazione di
Giuseppe Craparo



Grafica della copertina: *Alessandro Petrini*

Copyright © 2023 by FrancoAngeli s.r.l., Milano, Italy

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito www.francoangeli.it.

Indice

Prefazione, di Giuseppe Craparo	pag. 7
Introduzione	» 13
1. Le scienze cognitive della morale	» 15
1. I classici dilemmi morali	» 24
2. Cognizione, emozioni e moralità	» 37
2.1. Cognizione e moralità	» 38
2.2. Emozioni e moralità	» 41
3. Stati motivazionali e comportamento morale	» 46
2. Neuroscienze dell'etica	» 49
1. Correlati neurali della moralità	» 58
2. Neurologia e Psichiatria del ragionamento morale	» 61
3. Neurochimica della moralità	» 65
3. La morale in gioco: due diversi sistemi intelligenti, uno stesso problema morale	» 69
1. <i>Autonomous Driving Systems</i> e dilemmi morali	» 69
2. È possibile stabilire una politica etica per gli AVs?	» 78
3. ChatGPT e moralità	» 82
3.1. Cosa è ChatGPT	» 82
Conclusione	» 93
Bibliografia	» 99

Prefazione

di Giuseppe Craparo¹

Nell'ultimo trentennio, le ricerche neuroscientifiche hanno fornito un contributo importantissimo in vari ambiti della conoscenza umana, ad esempio: filosofica, psicologica, artistica, sino ad arrivare, tra le altre discipline, alle scienze giuridiche. Cito di seguito alcuni dei contributi che, in virtù anche della mia formazione di psicologo clinico, ritengo particolarmente interessanti: gli studi sul rapporto tra cognizione ed emozione (si pensi, per esempio, a Panksepp, LeDoux, Damasio); le ricerche sui fenomeni mentali connessi “alle coscienze” (si considerino, a tal proposito, le scoperte di Edelman); la teoria polivagale di Porges, che ha permesso di comprendere il ruolo del SNA nella ricerca di protezione o di connessione/interazione nel rapporto con l'ambiente; gli studi di Rizzolatti, Gallese e colleghi sui “neuroni specchio” coinvolti nelle interazioni sociali, in particolar modo nella comprensione implicita degli scopi e delle emozioni altrui (di grande rilievo è il costrutto, proposto da Gallese, di “simulazione incarnata”); per arrivare agli studi sul cervello morale, argomento di questo libro. Trovo che, al di là degli specifici argomenti trattati, questi (ed altri) studi vadano tutti nella stessa direzione, ossia supe-

¹ Professore ordinario di Psicologia clinica, Università degli Studi di Enna Kore.

rare l'idea di una separazione del corpo dalla mente, integrando necessariamente e virtuosamente i due versanti. È proprio in virtù di questa integrazione, o superamento della dualità cartesiana, che si parla oggi di “psiche incarnata”, di *embodied cognition*, così come, in filosofia, si fa sempre più strada l'idea di una *embodied phylosophy* (anche in psicologia clinica la centralità riconosciuta al corpo e al suo stretto legame con il mondo psichico – conscio e inconscio – inducono a considerare la relazione terapeutica come una relazione incarnata).

La lettura del libro di Lucifora e Vicario sulle basi neurobiologiche della morale ha stimolato in me alcuni interrogativi, in particolar modo sulla coscienza morale (più avanti spiegherò perché parlo di coscienza morale e non semplicemente di morale) e nello specifico: in che rapporto è la morale con la coscienza? Quali sono le particolarità psicologiche della coscienza morale? Quanto la qualità delle relazioni primarie può incidere sulla compromissione o sullo sviluppo normale di una coscienza morale? Qual è il rapporto fra coscienza morale e funzionamento di personalità? Quali sono le cause di una compromissione della coscienza morale? In che maniera si può intervenire clinicamente nei casi di compromissione della coscienza morale? Quanto incide il contesto socio-culturale sullo sviluppo di una coscienza morale? La coscienza morale è ereditabile? Sono solo alcune domande che evidenziano la complessità dell'argomento. Obiettivo di questa mia prefazione non è certo rispondere a questi interrogativi. Mi limiterò ad alcune brevi considerazioni partendo però dalla coscienza per poi passare alla coscienza morale.

Prima di riferirci alla coscienza morale è d'uopo porsi la domanda: cosa è la coscienza? Alla maniera di Sant'Agostino potremmo sostenere che per coscienza intendiamo “qualcosa che pensiamo di conoscere fino a quando qualcuno non ci chiede di definirla”. Si tratta certamente di un concetto complesso, controverso, per certi versi anche misterioso di cui alcuni autori hanno sottolineato il rapporto con l'atto, tanto da parlare di atto cosciente e di atto di coscienza: in entrambe

queste declinazioni la coscienza è posta in relazione ad una azione o un atto, come ad esempio l'atto del parlare, del percepire, del sentire o di un comportamento, ecc.; altri hanno delineato il rapporto della coscienza con il linguaggio, nel senso che la coscienza è in stretto rapporto con la capacità dell'essere umano di saper parlare e di conseguenza con la sua natura di "essere parlante": la conseguenza di ciò è che la coscienza è un "fattore tipicamente umano"; c'è invece chi è andato alla radice della questione affermando che la coscienza non può prescindere dall'*essere* della coscienza.

Per quanto sia un argomento controverso, mi sembra che non ci siano dubbi sul fatto che tutti siamo dotati di una coscienza, in quanto dotati di un cervello e di una mente. A tal proposito, il neurobiologo Joseph LeDoux sottolinea la distinzione tra *coscienza individuale*, ovvero l'essere vigili, svegli e in grado di percepire e descrivere l'ambiente circostante, e *coscienza dello stato mentale*, ossia l'essere consapevoli di star facendo esperienza di qualcosa. Dal punto di vista delle Neuroscienze Cognitive, che considerano la mente cognitiva un sistema di elaborazione il cui processamento delle informazioni avviene per lo più inconsciamente, ciò di cui siamo coscienti è il risultato dell'elaborazione di uno stimolo/informazione. Non tutti i risultati di questo processamento implicito però entrano nel campo della coscienza, alcuni infatti si mantengono inconsci, o non consci: il termine non conscio (o inconscio cognitivo) "si riferisce ai processi che svolgono funzioni che possono produrre o non produrre contenuti consapevoli" (in *Ansia*, 2015 trad. it. 2016, p. 221).

I due neuroscienziati Damasio e Panksepp distinguono una forma primitiva di coscienza, di natura affettiva, dalla coscienza cognitiva. La prima coinvolge gli stati sottocorticali intesi come "stati procedurali (forse veramente inconsci) sensorio-percettivi e affettivi impliciti, organizzati a livello neuronale sottocorticale" (cit. in LeDoux, 2015, p. 256), mentre la coscienza cognitiva è in rapporto ai processi corticali.

Poco incline a ridurre la coscienza alle attività neurali, per

il filosofo Thomas Nagel la coscienza è una qualità soggettiva, un'esperienza fenomenologica e qualitativa descritta dall'autore come *qualia*. Secondo questo filosofo, infatti, quello che possiamo osservare nel cervello non è coscienza, non è una esperienza soggettiva (*qualia*) ma attività neurale. Insomma, non è il cervello ad essere cosciente ma la mente. È questa una considerazione che, dal mio punto di vista, potremmo allargare anche alla morale: motivo per cui preferisco parlare di coscienza morale, nel senso che la morale è uno stato mentale della coscienza, di cui è possibile, come fanno Lucifora e Vicario, studiare i correlati neurobiologici: in questo senso va inteso, almeno è come l'ho inteso io, il titolo di questo libro: *Cervello morale*. Titolo che tradurrei appunto come: *Correlati neurobiologici della coscienza morale*.

Su cosa sia la morale (argomento anch'esso molto vasto) in *Neurobiologia della morale*, Patricia Churchland sostiene che "l'*etica* o la *morale*" [è corretto utilizzare questi due concetti come sinonimi?] "è uno schema quadrimensionale per il comportamento sociale a quattro dimensioni plasmato da processi cerebrali interconnessi:

- 1) *prendersi cura* (basato sull'attaccamento a parenti e amici e sull'interesse per il loro benessere);
- 2) *riconoscimento degli stati psicologici altrui* (basato sui benefici di predire il comportamento degli altri);
- 3) *soluzione di problemi nel contesto sociale* (per esempio, come dovremmo distribuire beni che scarseggiano, risolvere dispute sulla proprietà terriera, punire i criminali);
- 4) *apprendimento delle pratiche sociali* (tramite rinforzo positivo e negativo, per imitazione, per tentativi ed errori, in virtù di tipi di condizionamento, per analogia)" (2011, trad. it. 2012, p. 22).

L'autrice delinea quindi quelle che considera le caratteristiche principali della moralità: una preoccupazione primaria alla vita dell'altro; la capacità di cogliere gli stati mentali propri e

altrui (capacità che gli psicologi chiamano mentalizzazione o funzione riflessiva); l'insieme di diritti e di doveri socialmente determinati che distinguono i comportamenti morali da quelli che non lo sono; l'apprendimento e l'interiorizzazione dei codici morali.

Nella mia pratica clinica, è capitato di confrontarmi con pazienti con una grave compromissione della coscienza morale. Le cause di queste compromissioni possono essere molteplici e in rapporto fra loro: relazioni primarie traumatiche (per lo più storie precoci e croniche di maltrattamenti fisici, trascuratezza emotiva e/o abusi sessuali), ambiente evolutivo particolarmente incline a uno stile di vita antisociale, povertà, discriminazione, presenza di disturbo psichiatrico ecc.

Secondo lo psicoanalista Otto Kernberg, la compromissione della coscienza morale coinvolge in particolar modo coloro che rientrano in un *continuum* narcisistico che va da forme di narcisismo meno grave alla personalità psicopatica passando per il cosiddetto narcisismo maligno. All'interno di questo continuum, la personalità psicopatica rappresenta per lo psicopatologo (clinico e forense) una sfida proprio in virtù della grave compromissione, se non addirittura in alcuni casi dell'assenza, della coscienza morale che permetta loro di apprendere dall'esperienza così da evitare di commettere nuovamente azioni devianti, lesive per sé e per gli altri. È quello che Robert Hare ha sottolineato, definendo gli psicopatici "predatori intraspecie che usano fascino, manipolazione, intimidazione e violazione per controllare il prossimo e soddisfare i propri egoistici bisogni; mancando di morale ed empatia, riescono freddamente a prendere e a fare ciò che vogliono, violando norme e divieti sociali senza il minimo senso di colpa o rimpianto" (Hare, 2009). Assenza di colpa e di rimorso esprimono la mancanza di una coscienza morale.

Proprio sul rapporto fra emozioni e morale (o coscienza morale), Lucifora e Vicario citano numerose ricerche da cui sembra emergere una "coevoluzione" fra disgusto e morale.

Un argomento interessante che mi ha ricordato l'esperimento di Paul Rozin e Carol Nemeroff (1994) in cui fu chiesto ai partecipanti la disponibilità a indossare un maglione: proposta a cui molti risposero positivamente, indossandolo. Una volta indossato, fu detto loro che quel maglione era appartenuto a Adolf Hitler. Dopo questa informazione, la maggioranza dei partecipanti reagì con delle smorfie di disgusto. Secondo Rozin e Nemeroff, tale reazione di disgusto era in rapporto al timore di poter essere contaminati da una forma di "hitlerità", come se il semplice indossare il maglione potesse, nella loro immaginazione, influenzare, avvelenare, i loro giudizi morali ed estetici.

Non vado oltre con le sollecitazioni avute dalla lettura di questo libro; lascio al lettore la curiosità di scoprire un testo che, nonostante affronti tematiche impegnative, ha il pregio di una scrittura chiara e comprensibile anche ai non addetti ai lavori.

Introduzione

Il crescente interesse nei confronti delle Neuroscienze da parte dei filosofi ha favorito l'emergere di una nuova disciplina nel settore della Filosofia Morale chiamata *Neuroetica*.

Il principale merito di questa disciplina consiste nell'aver manifestato apertura nei confronti dei metodi d'indagine e dei modelli esplicativi propri delle Neuroscienze Cognitive e Sociali al fine di incrementare l'attuale conoscenza dei processi e dei meccanismi che determinano il ragionamento e il comportamento morale.

In questo senso la Neuroetica si configura come una disciplina dall'approccio integrato che, pur adottando il metodo sperimentale nello studio del ragionamento morale e delle sue espressioni sul piano comportamentale, mantiene salda l'ossatura di uno spirito filosofico.

Più in generale, il presente volume intende fornire una panoramica sui principali indirizzi di ricerca nel campo del comportamento e ragionamento morale, sfruttando la cooperazione tra diverse discipline quali Filosofia, Psicologia, Neuroscienze e intelligenza artificiale.

In linea con le Scienze Cognitive, la prima parte del testo assume un tono filosofico offrendo una panoramica del concetto di moralità indagato attraverso le lenti della filosofia mo-

rale, della psicologia cognitiva e affettiva. A tal proposito, verranno presi in esame alcuni risultati sperimentali sull'influenza esercitata dalle funzioni esecutive e dalle emozioni sul comportamento e ragionamento morale. Inoltre, in questo contesto sarà affrontato il tema dell'origine della moralità dedicando spazio alla letteratura *neosentimentalista*.

Segue una seconda sezione dal taglio neuroscientifico con l'obiettivo di fornire una panoramica sui correlati neurobiologici della moralità attraverso una disamina degli studi condotti su individui sani e su pazienti affetti da disordini di natura neurologica e psichiatrica. In questo contesto viene anche preso in esame il ruolo predittivo della neurochimica nel comportamento morale e le applicazioni derivanti dall'uso delle tecniche di stimolazione cerebrale non invasiva.

Il testo termina con una accurata indagine relativa alla relazione tra etica e intelligenza artificiale prendendo come esempio due casi studio (Autonomous Driving Systems e ChatGPT3) che condividono lo stesso problema: "cosa vuol dire compiere una scelta morale?".

1

Le scienze cognitive della morale

La capacità morale umana, ossia la capacità di distinguere il bene dal male, è oggi oggetto di studio in diverse discipline, che muovono dalla Psicologia, alla Filosofia e alle Neuroscienze. Qui ne delineiamo le principali teorie.

Dal punto di vista psicologico, si inizia a parlare di moralità con gli studi di Jean Piaget (1977), il quale delinea un approccio costruttivista allo sviluppo morale, intendendola come il prodotto dell'interazione del bambino con l'ambiente fisico e sociale. In tal senso, Piaget individua due tipi di moralità nello sviluppo del bambino. Il primo viene definito come basato sulla volontà dell'adulto e dura fino al nono anno di età, qui le regole imposte (dai genitori) sono inviolabili, le azioni stesse vengono giudicate sulla base delle conseguenze positive o negative che esse comportano. Il secondo, invece, si manifesta intorno al dodicesimo anno di vita. Qui le regole non vengono più intese come imposte da un'autorità ma rappresentano un sistema che garantisce un mutuo rispetto tra pari. In tale ottica, il giudizio morale non si concentra più sulle conseguenze ma sulle intenzioni delle azioni (Siegal, 1980).

Sulla stessa linea, Lawrence Kohlberg (1978) codifica un modello evolutivo morale a tre livelli.

Il primo livello è definito pre-convenzionale e dura fino al

nono anno di età, qui nella prima fase il bambino obbedisce alle norme morali per timore della punizione; nella seconda fase l'obbedienza alle regole si basa su accordi reciproci.

Il secondo livello, definito convenzionale, denota un processo di interiorizzazione della morale, definito da due fasi. La prima riguarda la creazione di un'idea generale del "buon comportamento", mentre la seconda è relativa al riconoscimento di un'istituzione in grado di mantenere l'ordine sociale.

Il terzo livello definito post-convenzionale riguarda il riconoscimento delle leggi e delle istituzioni poste.

Su tale base, lo psicologo formula il noto Dilemma di Heinz, che può essere così descritto (Kohlberg, 1978):

“Una donna è sul letto di morte. Esiste un farmaco che i medici pensano possa salvarla. Si tratta di un derivante del radium che un farmacista della stessa città ha scoperto recentemente. Il farmaco è costoso da produrre ma il farmacista lo fa pagare dieci volte in più di quanto costa a lui produrlo: paga 200 dollari per il medicinale e ne fa pagare 2000 per una piccola dose. Il marito della donna, Heinz, va da tutti quelli che conosce a chiedere in prestito il denaro ma riesce ad avere solo 1000 dollari, ovvero la metà del costo. Spiega al farmacista che sua moglie sta morendo e gli chiede di vendergli il medicinale a un prezzo più basso o di farglielo pagare più avanti. Ma il farmacista risponde “no, ho scoperto il farmaco e con questo diventerò ricco”. Così Heinz, disperato, entra nel laboratorio dell'uomo per rubare il farmaco per sua moglie”.

Le risposte che vengono date al dilemma riguardano le sopracitate fasi dello sviluppo morale e possono essere così riassunte:

Livello pre-convenzionale (bambini):

- obbedienza: commettere il crimine poiché ruberebbe solo il farmaco e non altro; non commettere il crimine perché verrebbe arrestato
- accordi reciproci: commettere il crimine per essere felice di salvare la moglie; non commettere il crimine perché la prigione è un luogo orribile

Livello convenzionale (adolescenti):

- buon comportamento: commettere il crimine perché sarebbe visto come un bravo marito; non commettere il crimine perché rubare è sbagliato
- legalità: commettere il crimine per salvare la moglie ma pagarne le conseguenze in carcere; non commettere il crimine poiché la legge vieta di rubare

Livello post-convenzionale (adulti):

- contratto sociale: commettere il crimine poiché la vita è un bene principale; non commettere il crimine perché non è mai un diritto rubare
- giustizia universale: commettere il crimine perché salvare una vita umana è un dovere; non commettere il crimine perché altri potrebbero avere bisogno di quel farmaco

In tal senso, dunque, la moralità è intesa come un concetto che evolve nel tempo, in modo graduale.

Muovendosi dall'ambito psicologico all'ambito filosofico, nell'ambito della filosofia morale si è soliti distinguere la teoria consequenzialista relativa all'idea di John Stuart Mill (Hooker, 2000) basata sul raggiungimento del bene superiore, dalla teoria deontologica in riferimento all'imperativo categorico di Immanuel Kant, che garantisce i diritti e i doveri di tutti i cittadini (Hill, 2002), dalla teoria della virtù risalente ad Aristotele e basata sul raggiungimento della felicità attraverso l'esercizio virtuoso (Dudley, 1999).

L'approccio *consequenzialista*, rappresentato storicamente da autori come J. Bentham e J.S. Mill, sostiene che un'azione morale è quella che massimizza la maggiore felicità, o assenza di dolore, per il maggior numero di persone (Hooker, 2000; Darwall, 2003a; Greene, 2013).

L'etichetta *consequenzialismo* si riferisce al fatto che, secondo questo approccio, la moralità di alcune azioni, decisioni o comportamenti viene valutata sulla base delle conseguenze di tale decisione per le persone coinvolte: "dovremmo fare

tutto ciò che produrrà le migliori conseguenze generali per tutti gli interessati” (Greene, 2013, p. 107).

In tal senso, il processo di decisione morale può essere sostanzialmente ridotto ad una sorta di analisi costi-benefici, che mira alle conseguenze delle azioni ex ante, che potremmo definire come “Etica delle Conseguenze”.

Valutare le conseguenze di determinate azioni conduce inevitabilmente al principio di massimizzazione dell'utilità attesa per il maggior numero di parti interessate, da qui il concetto di *utilitarismo* teso alle sole conseguenze immediate, valutate come somma del dispiacere prodotto meno il piacere.

Il principio $U > 0$ (utilità maggiore di zero) è un concetto particolarmente utilizzato nell'ambito dell'economia, in cui è stata portata avanti l'idea di *homo oeconomicus*, ossia la possibilità di intendere l'essere umano come dotato di una razionalità olimpica, che gli permette di massimizzare la sua utilità in ogni circostanza (Gray, 1987). Tuttavia, nella vita quotidiana l'applicazione di tale principio è spesso contraddetta. Ciò è stato dimostrato attraverso l'uso del ben noto *Ultimatum Game* (Güth *et al.*, 1982), un gioco di contrattazione economica che consente di creare delle situazioni di iniquità allo scopo di studiare come le persone prendono decisioni quando è richiesta la divisione di risorse e come reagiscono all'offerta ricevuta (vedi Tabella 1 per maggiori dettagli).

Tale gioco può essere così descritto (Thaler, 2000):

“A questo gioco partecipano due soggetti A (proponente) e B (ricevente) con lo scopo di dividersi una somma di denaro.

A decide come dividere la somma;

B decide se accettare o rifiutare.

Se B accetta, la somma verrà divisa secondo quanto stabilito da A; se B rifiuta nessuno dei giocatori riceverà il premio in denaro”.

Sulla base del principio $U > 0$, la decisione migliore per A potrebbe essere così riassunta: 99% per sé e 1% per l'altro. In tal modo A avrà massimizzato il suo guadagno riducendo al minimo le perdite e B avrà a sua volta massimizzato il suo

guadagno secondo il principio $U > 0$, in quanto 1 è sempre maggiore di 0.

Tuttavia, come notato negli studi di Thaler (2000) la maggior parte dei riceventi rifiuta somme inferiori al 20%, manifestando importanti emozioni negative. Allo stesso modo la maggior parte dei proponenti non offre cifre eccessivamente basse per paura di un rifiuto da parte dei compagni.

Conferme di ciò arrivano anche dall'ambito evoluzionistico, Proctor e colleghi (2013) mostrano con un esperimento in laboratorio che sia i bambini che gli scimpanzé possiedono un primordiale senso di equità, reagendo in modo emozionale a ricompense che ritengono minori rispetto a quelle offerte.

	Somma N
Proponente (A)	Proposta divisione
Ricevente (B)	Accettare / Non accettare

Tabella 1 - *Ultimatum Game* (Güth et al., 1982)

Dalla parte opposta troviamo, invece, filosofi ed etici che tendono ad ignorare le conseguenze di azioni e decisioni considerandole irrilevanti per la moralità. Questo è l'approccio dei cosiddetti deontologici, il più famoso è Immanuel Kant, che ha proposto il concetto di *imperativo categorico* come unica guida razionale al processo decisionale morale.

“Agisci solo secondo quella massima, la quale puoi, allo stesso tempo, intendere come una legge universale” (Kant, 1785). Secondo le teorie deontologiche, la scelta moralmente corretta non riguarda le conseguenze delle azioni, ma il rispetto o meno dei diritti e doveri fondamentali degli individui (Hill, 2002; Darwall, 2003b). In questo caso, si distingue all'interno dell'etica delle regole, la cosiddetta etica della prescrizione divina in cui bisogna perseguire le norme emanate da Dio e l'etica kantiana in cui l'imperativo categorico richiede di agire sulla massima di azioni che potrebbero diventare leggi universali. Qui si privilegia un ragionamento *ex post*, basato sia