

NOTIZIE DI FILOSOFIA

EVENTI CULTURALI - LABORATORIO DI IDEE

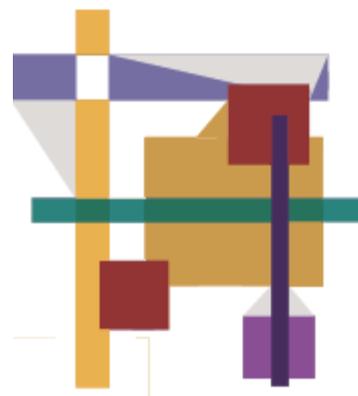
Humana.Mente

Il Pensario della Biblioteca Filosofica

NOTIZIARIO TRIMESTRALE — N° 5, APRILE 2008

www.humana-mente.it

redazione@humana-mente.it



Verso una neuro-filosofia?



Biblioteca Filosofica © 2007 - Humana.mente,
Periodico trimestrale di Filosofia, edito dalla Biblioteca Filosofica -
Sezione Fiorentina della Società Filosofica Italiana,
con sede in via del Parione 7, 50123 Firenze
(c/o la Facoltà di Scienze della Formazione dell'Università degli Studi di Firenze) -
Pubblicazione regolarmente iscritta al Registro Stampa Periodica del Tribunale di Firenze
con numero 5585 dal 18/6/2007.

**REDAZIONE - Via del Parione 7, Firenze, presso Biblioteca Filosofica - Facoltà di Scienze della
Formazione, Università degli Studi di Firenze**

- **Direttore editoriale: Prof. Alberto Peruzzi**
- **Direttore responsabile: Duccio Manetti.**

- **Supervisore scientifico: Prof. Marco Salucci**

Redattori e Area di Ricerca:

Scilla Bellucci - Filosofia del Linguaggio / Antropologia Filosofica
Laura Beritelli - Filosofia del Linguaggio / Ermeneutica Filosofica
Alberto Binazzi - Filosofia della Mente / Scienze Cognitive
Matteo Borri - Filosofia e Storia della Scienza
Giovanni Casini - Filosofia della Scienza / Logica
Chiara Erbosi - Bioetica / Filosofia della Mente
Riccardo Furi - Filosofia della Mente
Tommaso Geri - Filosofia del Linguaggio / Ermeneutica Filosofica
Matteo Leoni - Antropologia Filosofica
Stefano Liccioli - Filosofia Morale
Umberto Maionchi - Filosofia della Scienza / Logica / Filosofia del Linguaggio
Francesco Mariotti - Storia e Filosofia delle Neuroscienze / Filosofia della Mente
Giovanni Pancani - Filosofia Politica / Antropologia Filosofica
Daniele Romano - Filosofia e Storia della Scienza
Emilio Troia - Didattica della Filosofia
Fabio Vannini - Filosofia della Scienza
Silvano Zipoli - Filosofia e Storia della Scienza
Segretario di Redazione: Matteo Leoni

I contenuti di Humana.mente sono sottoposti a *refereeing*

**Il comitato scientifico della rivista è composto dai membri del Consiglio Direttivo
della Biblioteca Filosofica:**

Alberto Peruzzi (presidente)
Gaspare Polizzi (vice presidente)
Chiara Cantelli (segretaria tesoriera)
Fabrizio Desideri
Ubaldo Fadini
Rosa Martiniello
Marco Solinas

Progetto Grafico: Duccio Manetti
Sito web: www.humana-mente.it
La struttura del sito è realizzata da Emiliano Mazzetti e Simone Gallorini
Per informazioni e collaborazioni: redazione@humana-mente.it

INDICE

- Editoriale - pag. I
- Filosoficamente scorretto - pag. V
- Inaugurazione Associazione di Promozione sociale 'Studi Umanistici Rogersiani' – pag. VII
- Qual è il “laboratorio” per l’arte oggi? Relatori: Gianfranco Baruchello e Giuseppe Di Giacomo – pag. IX
- Pensare la musica, comporre la musica: un ritmo comune? Relatori: Sergio Givone, Daniele Lombardi – pag. XI

PAPERS

- Peruzzi, A., *Qualche riflessione sulla teoria dell'identità* - pag. 1
- Lanfredini, R., *La teoria dell'identità. Alcuni problemi epistemologici* – pag. 21
- Salucci, M., *La teoria dell'identità, precisazioni e risposte* – pag. 33
- Erbosi, C., - *La questione del libero arbitrio tra compatibilismo e incompatibilismo* - pag. 49
- Binazzi, L., – *Modelli mentali: un'analisi epistemologica* - pag. 65
- Furi, R., - *Evoluzione e libertà* - pag. 103
- Battilotti, G.,- *Logica, computazione e teorie quantistiche della mente* - pag. 119
- Beritelli, L., - *Alcune considerazioni sui limiti semantici della fisiologia* - pag. 133
- Borri, M., - *Neuroimaging: continuità e innovazione* - pag. 149

RECENSIONI

- *Neuroetica*, Laura Boella - pag. 165
- *Il volto e la parola*, di Felice Cimatti - pag. 169
- *Seconda natura. Scienza del cervello e conoscenza umana*, di Gerald Edelman - pag. 173
- *Menti morali*, di Marc D. Hauser - pag. 177
- *Mind Time*, di Benjamin Libet - pag. 181
- *Immagini della mente*, a cura di Lucignani G. e Pinotti A. - pag. 185
- *Psicologi per natura*, Cristina Meini - pag. 189
- *La mente etica*, di Michael Gazzaniga - pag. 195
- *La mente sociale*, di Cristina Meini e Massimo Marraffa - pag. 199
- *Neurofenomenologia*, di M. Cappuccio - pag. 207
- *Pensieri Materiali*, di Simone Gozzano - pag. 211
- *Simulating Minds*, di Alvin I. Goldman - pag. 219

RILETTURE

- *Mente e mondo fisico*, di Jaegwon Kim - pag. 223
- *La mente e il computer*, di P.N. Johnson Laird - pag. 229
- *Il cervello emotivo. Alle origini delle emozioni*, di Joseph LeDoux - pag. 235
- *La società della mente*, di Marvin Minsky - pag. 237
- *Empirismo e filosofia della mente*, di Wilfrid Sellars - pag. 241
- *Neurophilosophy*, di Patricia S. Churchland - pag. 251

INTERVISTE

- Lucio Luzzatto - pag. 267

Editoriale

VERSO UNA NEURO-FILOSOFIA?

Oggi come non mai la ricerca scientifica sembra spingere sempre più in là la frontiera delle ricerche sul cervello ed imporre una nuova dinamica agli studi sulla mente. Il rapporto tra la materia e il pensiero ha occupato il centro del dibattito filosofico per tutto il corso della storia.

Con la pubblicazione del quinto fascicolo *Humana.Mente* festeggia il suo primo anno di vita editoriale e lo fa dedicando a questo tema l'intero numero.

Gli stati mentali e le nostre sensazioni, che fino a ieri sembravano non poter costituire oggetto di ricerca scientifica, stanno entrando pian piano nel mirino della sperimentazione. L'uomo non può ancora affermare di essere arrivato a cogliere appieno la sua natura cosciente, ma negli ultimi anni le neuroscienze hanno compiuto passi da gigante nella direzione di una definizione molto precisa di alcuni 'pezzi' del nostro sistema cognitivo.

Il dibattito sul rapporto tra mente e cervello ha subito una svolta ad inizio Novecento quando una parte della filosofia analitica decise di puntare il proprio obiettivo non più soltanto sull'analisi del linguaggio, ma sull'analisi della mente, produttrice del linguaggio stesso. Da lì in poi sono nati fecondi filoni di ricerca e si sono sviluppate differenti concezioni e teorie del mentale. A questo proliferare di studi si sono aggiunte poi le neuroscienze, tutte quelle discipline, dalla linguistica cognitiva alla neuropsichiatria, dalla neurofisiologia alla neurobiologia, che hanno fatto dell'indagine sull'interazione tra sistema cognitivo e materia cerebrale l'oggetto della propria attenzione.

Anche se la spinta 'analitica' degli studi sul mentale sembra oggi essersi spenta, le neuroscienze rendono ancora vivo e fertile il terreno della ricerca in questo campo. Sul versante prettamente teorico tutte le posizioni si sono dovute confrontare con la teoria principe tra le concezioni della mente – la cosiddetta teoria dell'identità – che ha dominato la scena dagli anni cinquanta in poi del secolo scorso.

Chiunque si sia occupato di filosofia della mente ha dovuto infatti fare i conti con l'idea secondo la quale pensieri e stati mentali non sono altro che gli stati del nostro cervello; sono quindi i neuroni a 'causare' i nostri pensieri e le nostre sensazioni e di lì non si "scappava".

Di fatto la teoria dell'identità ha inaugurato e reso indipendente quella disciplina che oggi etichettiamo come filosofia della mente. Tra i pochi che in Italia hanno dedicato attenzione e sforzi a questa teoria vi è un ricercatore dell'Università di Firenze, Marco Salucci, che nel 2005 a questo tema ha dedicato una monografia (*La teoria dell'identità*, Le Monnier, Firenze). In occasione della presentazione del testo l'autore e due studiosi dell'Università di Firenze, che hanno entrambi dedicato parte dei propri studi alla ricerca sul mentale – Alberto Peruzzi e Roberta Lanfredini – si sono confrontati sulla validità o meno di questa teoria. In questo numero pubblichiamo gli estratti e il frutto di quella discussione.

Abbiamo voluto provare a rispondere ad alcune domande sulla discussione tra libero arbitrio e causalità mentale: che rapporto vi è tra stati mentali e stati cerebrali? Nonostante il valore impositivo degli *inputs* del cervello l'uomo può ancora dirsi libero? Quali sono le attuali ricerche in corso e di quali tecniche fanno uso? Come è stato affrontato il rapporto tra determinismo e libertà? Quali modelli mentali la psicologia cognitiva è stata in grado di elaborare?

Anni fa sembrava che la meccanica quantistica potesse offrire una speranza alla soluzione di questo mistero. Alcuni scienziati ed epistemologi, tra i quali spiccava Roger Penrose, tentarono la strada della ricerca di un legame tra i due ambiti più misteriosi e affascinanti della realtà naturale: il mondo del micro (regolato dalla quantistica) e la mente. Quegli studi si sono arenati sull'incapacità di mostrare un'interazione certa tra gli effetti quantistici e gli stati mentali coscienti. La meccanica quantistica ha comunque continuato il suo corso ed è oggi pilastro della fisica contemporanea. Una branca di questa disciplina, la logica quantistica, si è occupata tra l'altro di quale potesse essere l'approccio migliore al *mind/body problem*. Noi abbiamo chiesto ad una delle più brillanti ricercatrici in campo matematico di spiegarci quali soluzioni possano sussistere oggi.

Con le recensioni e le riletture abbiamo voluto prendere in esame la letteratura scientifica più recente e comprendere quali tra le ipotesi proposte negli ultimi anni sia quella più consistente e attualmente più libera da debolezze e fallacie. Il nerbo della discussione, volutamente riassunto con un interrogativo provocatorio è se si debba andare verso una filosofia completamente 'dipendente' e 'determinata' dalla ricerca scientifica, o se invece funzionalismo, teoria dell'identità e ogni altra teoria della mente, abbiano ancora qualcosa da dire e svolgano un ruolo teoretico imprescindibile anche per la scienza.

Il futuro va dunque verso una neuro-filosofia? Con questo non vogliamo invitare la categoria dei filosofi, a cui sentiamo sentitamente di appartenere a ricorrere ai servizi

della psichiatria... ma vogliamo provocare il dibattito chiedendoci se abbia un senso continuare a ignorare l'importanza delle scoperte sperimentali per la comprensione di un problema teorico come quello rappresentato dal rapporto mente-corpo, o se invece proprio queste scoperte possano offrire una *chance* alla filosofia per attualizzare molte delle domande su cui questa vastissima area del sapere umano si è sempre arrovellata, rivitalizzando lo status degli studi in corso.

Rubrica

FILOSOFICAMENTE SCORRETTO

Quarta puntata

CONSUMI E DEMOCRAZIA - UNA SOCIETÀ IN LIBERTÀ VIGILATA

Lo studio dei sistemi politici del mondo contemporaneo ha sempre trascurato un aspetto fondamentale della struttura profonda della ormai universale società dei consumi: il ruolo centrale del rapporto tra l'esercizio effettivo delle libertà individuali e la capacità/possibilità di esercitarle in un contesto di libertà pubblica formale. Intendo dire che l'esercizio effettivo e concretamente realizzabile delle libertà individuali formalmente garantite in un contesto "democratico" costituisce un aspetto marginale del dibattito culturale e politico attuale nelle società tecnologicamente avanzate (a parte qualche lodevole eccezione).

Questo punto mi sembra invece particolarmente importante dal momento che l'esercizio della democrazia non si fonda soltanto sul criterio di maggioranza elettiva e, conseguentemente, sul principio "una testa, un voto" ma sulla reale possibilità del singolo cittadino elettore di esprimere un suo personale "punto di vista", una "propria" opinione. Se viene a mancare questo indispensabile presupposto, cioè la concreta capacità di una scelta individuale basata su convinzioni "proprie", raggiunte attraverso una personale elaborazione delle "proprie" opinioni, tutta la struttura formale delle regole democratiche rischia di crollare. Viene ripetuto costantemente fin quasi alla noia che il fondamento della democrazia e del suo esercizio sta nelle "regole formali" e nella loro puntuale applicazione. Ma quasi nessuno si preoccupa di aggiungere che quelle regole "formali" hanno un "contenuto" profondo: riflettono e riproducono una rete interconnessa di norme etiche non scritte, di principi e comportamenti che appartengono al costume di una intera civiltà come eredità culturale della sua storia.

La sopravvivenza e lo sviluppo delle democrazie esistenti e di quelle potenziali dipende dunque essenzialmente dai meccanismi di formazione dell'individuo-cittadino e dal rispetto, quanto più possibile, della singolarità di ognuno, delle sue aspettative, delle sue credenze, delle sue inclinazioni, ecc. Nelle forme di governo delle democrazie occidentali, un ruolo centrale è giocato dall'opinione pubblica. Quanto più essa è differenziata e multiforme, tanto più è garantita la libera circolazione di idee, progetti, prospettive,

proposte. E quanto più i singoli cittadini sono abituati al libero confronto di convinzioni ed opinioni diverse ed originali, tanto più la vita democratica di quel paese si arricchisce ed è vitale.

Ma per poter raggiungere questo livello virtuoso di civiltà è appunto indispensabile la presenza di una forte penetrazione culturale tra i cittadini, di una maturità politica derivata da una naturalezza che si acquisisce con l'abitudine alla ricerca di strade nuove. Proporre nuove idee, sviluppare, come si usa dire oggi, la ricerca e l'innovazione anche e soprattutto a livello del singolo cittadino fa parte integrante della sua "educazione culturale" e quindi, aggiungiamo, della sua "educazione politica". Potremmo dunque dire con assoluta sicurezza, ispirandoci a Giulio Preti (un autore a noi molto caro), che col progredire della scienza e della tecnologia nella nostra civiltà, l'educazione politica e civile dei cittadini dipende in modo essenziale dal livello della loro cultura generale e di quella scientifica in particolare. E che la vita democratica di un paese si alimenta della creatività e dello spirito critico dei propri cittadini, della loro autentica libertà di pensiero in primo luogo.

Come ha detto qualcuno, "spesso non è la libertà che manca, mancano gli uomini liberi". E, aggiungerei, gli uomini spesso non nascono liberi, purtroppo, ma lo possono diventare.

Umberto Maionchi

Conferenze

INAUGURAZIONE ASSOCIAZIONE DI PROMOZIONE SOCIALE STUDI UMANISTICI ROGERSIANI

Il 29 marzo scorso, alle ore 17.00 presso il circolo culturale di viale Fratelli Rosselli 9, è stata inaugurata l'Associazione Culturale di Promozione Sociale (www.studirogersiani.it) con il primo di un ciclo di seminari che l'Associazione proporrà alla cittadinanza intera allo scopo di illustrare gli obiettivi dell'Associazione stessa e le sue attività. I seminari vogliono anche essere l'occasione per osservare l'applicazione del paradigma umanistico rogersiano nelle tematiche di volta in volta proposte.

L'associazione è costituita da un gruppo di psicologi e di psicoterapeuti che si identifica nel paradigma umanistico di Carl R. Rogers, il cui messaggio fondamentale è che le persone nascono con un bagaglio di risorse personali per migliorare in ogni caso le proprie condizioni esistenziali. Compito dello psicologo è quello di facilitare l'espressione di quel processo di attualizzazione che caratterizza tutti gli esseri viventi. In materia di malattia mentale, i rogersiani non credono nel determinismo: i disturbi non sono legati ad un'unica causa. Le persone costruiscono modalità di espressione del disagio, che tutti possono condividere, anche se lo esprimono con sintomi diversi. Nella visione rogersiana della natura umana, si tiene in una particolare considerazione la valenza emotiva. Per questa ragione viene identificato nell'empatia il principale strumento di cura. La teoria di Carl R. Rogers e la fenomenologia insegnano un profondo rispetto etico nei confronti delle persone, perché viene dato credito al loro modo di vedere il mondo. Attraverso il confronto e la relazione tra le persone è possibile aumentare il potenziale umano, per questo l'Associazione organizza gruppi di incontro. L'Associazione è convinta che il paradigma rogersiano possa essere applicato a qualsiasi settore umano (dalla psicologia clinica alla psicologia del lavoro, alla criminologia e psicologia forense). Verranno quindi organizzati corsi di formazione nell'applicazione del paradigma umanistico nei vari settori della psicologia.

Ha tenuto a battesimo l'Associazione il Prof. Alberto Peruzzi, ordinario di Filosofia teoretica presso l'Università di Firenze, il quale ha ricordato la figura e l'opera dello psicologo Carl R. Rogers, al lavoro del quale l'Associazione intende ispirarsi. Peruzzi ha ricordato l'originalità dell'approccio rogersiano sottolineandone i caratteri di diversità rispetto ad altre scuole psicologiche contemporanee. La persona è infatti il centro della

riflessione e della metodologia rogersiana, il paziente psicologico è comunque e sempre considerato non solo come malato ma come un valore depositario di risorse auto-terapeutiche che lo psicologo, mediante una relazione empatica, deve aiutare a crescere. Si tratta, ha affermato Peruzzi, di una forma moderna di aretè, di socratica cura del sé, tesa a valorizzare quanto di positivo c'è in ogni personalità per quanto disturbata possa essere.

La parola è poi passata al Presidente dell'Associazione, lo psicologo e criminologo Mario Ruocco, che, dopo aver illustrato i fini dell'associazione consistenti nel diffondere un pensiero umanistico tanto più necessario in un'epoca dominata dalla tecnica, ha esposto una relazione dal tema: *Dalla catastrofe massmediatica all'approccio centrato sulla persona*. Ruocco, ha voluto sottolineare che è necessario un confronto con la tecnica e non un mero rifiuto. Inaugurando così il primo seminario dell'Associazione ha caldeggiato il superamento di un atteggiamento pessimistico e rassegnato riguardo alla "catastrofe" che la tecnologia dei mass media ha prodotto nella società contemporanea. La relazione è stata seguita da un ampio dibattito fra il relatore e il pubblico presente.

Marco Salucci

Conferenze

QUAL È IL “LABORATORIO” PER L’ARTE OGGI?

Pensare il presente della musica e delle arti figurative 2008

Relatori: Gianfranco Baruchello e Giuseppe Di Giacomo

Venerdì 29 Febbraio l’artista Gianfranco Baruchello ed il filosofo Giuseppe Di Giacomo, professore ordinario di Estetica presso l’Università di Roma “La Sapienza”, si sono confrontati sul tema “Qual è il *laboratorio* per l’arte oggi?”. Baruchello, presentatosi come un vecchio signore che da molti anni fa il pittore, ha affermato che negli ultimi cinquant’anni non si è limitato ad usare il pennello, ma anche le mani ed il cervello per compiere un lavoro sperimentale, un’attività d’indagine. “La pittura negli ultimi quaranta anni – ha continuato – ha dimenticato l’oggetto, ha fatto un salto nel nulla: in questo niente ci si muove alla ventura, correndo anche il rischio di non essere compresi”. In questo senso il relatore ha citato la sua esperienza di lavoro e ricerca artistica a contatto con l’agricoltura. “Il pittore – ha concluso Baruchello – si posa come una farfalla sui vari aspetti della realtà quotidiana, sugli interstizi del quotidiano”. Ha preso dunque la parola Giuseppe Di Giacomo che ha domandato che cosa può succedere oggi in un laboratorio di un artista, proprio come quello di Baruchello. Citando Adorno ha affermato che la filosofia ha bisogno dell’arte per dire quello che la filosofia da sola non può dire e lo può mostrare l’arte tacendo, anche se nemmeno essa riesce ad esprimerlo fino in fondo. “Il filosofo non può dire ciò che l’arte dovrebbe fare oggi. Soltanto Adorno – ha detto – ha sostenuto che l’arte deve compiere certe cose oppure viene meno: ad esempio mantenere e lavorare sulla forma». Il relatore ha insistito proprio su questo punto sostenendo che l’arte non è più tale se rinunciamo alla forma, che è carica di contenuto: in essa i significati si sono stratificati. Paul Klee, artista e filosofo, ha teorizzato appunto un pensiero pittorico. In una conferenza del 1924 spiegò con un’immagine che cosa accade in un laboratorio di un artista: questi è come il tronco di un albero, le cui radici sono i suoi contatti con la realtà. L’opera d’arte è prodotta dalle radici, un pittore o uno scultore sono solo i mediatori. Come ha affermato lo stesso Klee: “L’arte non ripete le cose visibili, ma rende visibile”. Di Giacomo ha concluso il suo intervento soffermandosi su un ulteriore aspetto: non c’è arte senza elemento riflessivo. Citando ancora Adorno, egli ha mostrato come in un autore come Kafka e, soprattutto, Beckett, non c’è riflessione a livello di contenuto, ma di forma. In Beckett ad esempio che non riproduce il visibile, ma questo mondo disgregato, la forma

è montata insieme alla riflessione. In quest'ottica un'opera d'arte si dà cosa tra le cose, ma un aspetto la distingue dal resto: la forma appunto.

Stefano Liccioli

Conferenze

PENSARE LA MUSICA, COMPORRE LA MUSICA: UN RITMO COMUNE?

Pensare il presente della musica e delle arti figurative

Relatori: Sergio Givone, Daniele Lombardi.

Coordina: Eleonora Negri.

Per il ciclo *Pensare il presente della musica e delle arti figurative*, il 31 gennaio si è tenuto al Gabinetto Vieusseux un incontro sul tema *Pensare la musica, comporre la musica: un ritmo comune?*

Sergio Givone, ordinario di Estetica presso l'Università di Firenze, ha sottolineato la necessità di chiarire il significato dell'espressione "pensare la musica", che può riferirsi ad almeno due cose, molto diverse tra loro: può indicare una domanda relativa alla definizione della musica, volta a capire a quale genere artistico appartenga e dunque a una sua catalogazione. D'altro canto l'espressione in questione può anche rimandare a una domanda ben diversa: la musica è o non è necessaria alla filosofia? Bisogna cioè stabilire se la filosofia, quando incontra la musica, si interroga su se stessa, ovvero se, attraverso di essa, entra in contatto col suo stesso problema.

Givone ha poi evidenziato come la risposta data dal pensiero occidentale abbia sovente escluso la musica, pure riconosciuta come fonte di esperienze sublimi, dal novero delle cose essenziali alla filosofia, intesa altresì come sapere limpido, trasparente, «privo dell'intonazione che trasferisce il logos al di là di se stesso».

Una simile tendenza, presente sin dal pensiero antico, ha portato all'elaborazione di un progetto filosofico preciso, quello della purificazione linguistica; alla base di tale programma vi è la convinzione della possibilità di tradurre il linguaggio in un codice matematico esatto, purificato da ogni patos.

Givone ha fatto notare poi che proprio uno dei maggiori teorici di tale progetto, Wittgenstein, ha evidenziato come la sua realizzazione comporti la perdita dell'altro, di ciò che precede la parola, di ciò che non può essere detto. A questo punto la filosofia, attraverso la musica, è costretta a ripensare il proprio essere e il proprio statuto.

Dietro a questa tensione filosofica verso il non dicibile, vi è l'idea, coltivata da seguaci di Wittgenstein come S. Cavell, che la filosofia sia chiamata a proseguire il ritmo del mito, che debba essere, ha proseguito Givone, «necessariamente intonata».

Daniele Lombardi, musicista, ha articolato il suo intervento intorno alla nozione di 'musica', insistendo su come questa parola possa significare «un oceano di cose diverse», ragione per cui le critiche del pensiero occidentale alla musica intesa come negazione del logos sarebbero valide soltanto per una forma particolare di linguaggio musicale, relativo al linguaggio tonale quale si è sviluppato alla fine dell'ottocento.

Givone, riprendendo la parola, ha insistito sul rapporto di prossimità e lontananza che lega filosofia e musica: «costruire una teoria generale della musica, per quanto sia una strada alta, è impensabile, poiché essa è ormai troppe cose». Tuttavia, ha concluso, «la filosofia, se parliamo non di verità oggettive ma di senso, non può che essere musicale, intonata».

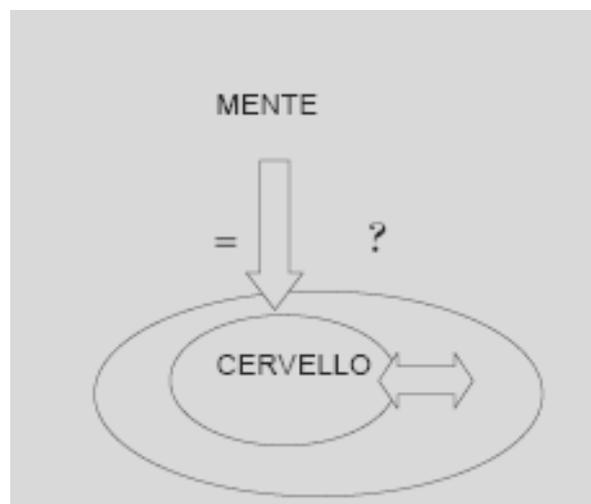
Matteo Leoni

QUALCHE RIFLESSIONE SULLA TEORIA DELL'IDENTITÀ

di Alberto Peruzzi

Estratto dalla presentazione dell'antologia curata da Marco Salucci *La teoria dell'identità: Alle origini della filosofia della mente* (Le Monnier, Firenze 2005)

(Liceo Gobetti, Bagno a Ripoli, 21 febbraio 2006)



1. Premessa occasionalista

Nel volume sono raccolti alcuni dei principali articoli che hanno segnato il percorso della teoria che afferma l'identità mente-corpo (nel modo che preciserò dopo). La filosofia della mente non sarebbe oggi quel che è senza il contributo della teoria dell'identità, quindi per capire l'oggi, si deve tener conto della teoria dell'identità. Capire però non equivale a risolvere i problemi e il fatto che senta il bisogno di ricordarlo vi dice subito che attribuisco una buona parte della responsabilità in questa lacuna alla teoria dell'identità. La silloge curata da Marco Salucci fornisce anche la prima traduzione italiana di molti dei testi raccolti e costituisce pertanto un punto di riferimento per chi intenda farsi una prima idea di quali siano gli argomenti addotti dai sostenitori della teoria dell'identità e quali siano le difficoltà cui si va incontro servendosi di tali argomenti. I meriti del volume non si fermano alla traduzione e alla

possibilità di trovare finalmente raccolti insieme numerosi contributi, sparsi in riviste e testi pubblicati in lingua inglese. Il volume infatti si completa con un prezioso contributo di Luca Malatesti, che ordina opportunamente i vari elementi emersi nel dibattito intorno a uno degli argomenti più discussi degli ultimi tempi (“l’argomento della conoscenza”), da alcuni considerato decisivo per l’accettazione o il rifiuto della teoria dell’identità. Ma soprattutto il volume si apre con un’ampia Introduzione di Salucci, ove i punti principali dei saggi raccolti sono sintetizzati, inquadrati criticamente e raccordati tra loro, arrivando a fare persino una tassonomia delle varie formulazioni, discutendo le obiezioni che la teoria dell’identità ha suscitato, aprendo l’orizzonte su una vasta letteratura critica, fornendo indicazioni utili su questioni tuttora dibattute. Pregevole è infine la chiarezza con cui Salucci ricostruisce la teoria dell’identità nel suo sviluppo a partire dai lavori originali di Place, Feigl e Smart nella seconda metà degli anni Cinquanta.

Per tutte queste ragioni il volume si raccomanda come testo di riferimento per corsi universitari sulla filosofia della mente, ma più in generale anche come guida per chiunque desideri seriamente avvicinarsi a uno dei temi più controversi nel dibattito recente in filosofia, non semplicemente perché per farlo occorre partire dalle fonti primarie – e di sicuro i testi qui raccolti sono ormai considerati dei “classici” –, ma soprattutto perché in essi vengono alla luce i problemi di fondo in maniera esemplare.

La presentazione del volume mi offre l’occasione per esprimere alcune riflessioni sulla teoria dell’identità. Seppure si tratti di riflessioni di carattere generale, che non entrano nei dettagli e che lasciano fuori molte questioni, alle quali lo stesso Salucci ha rivolto attenzione nelle pagine introduttive, spero che possano essere ugualmente utili.

2. Di che cosa si tratta?

I sostenitori della teoria dell’identità non solo affermano che “gli stati mentali sono stati fisici” ma argomentano anche che “gli stati mentali sono stati fisici” è un’ipotesi empirica, quindi scientificamente accertabile. La precisano dicendo che gli stati fisici in questione sono stati *cerebrali* e proponendo (o dando per implicito) il suo rafforzamento: gli stati fisici rilevanti sono *soltanto* quelli cerebrali.

Questo rafforzamento gioca un ruolo cruciale nel rispondere al quesito seguente: SE ci sono proprietà di stati mentali che non sono proprietà inerenti alla natura (fisica), ALLORA ne possiamo inferire che esistono *entità* che vanno al di là della scienza (naturale)? La risposta

sembra ovviamente positiva, in quanto ciò che diciamo esistere è caratterizzato dalle sue proprietà, siano esse grandezze misurabili o qualità, e a meno di dimenticare la critica di Locke alla nozione di sostanza, i tipi di queste proprietà (in fisica si fa non a caso un'analisi dimensionale) definiscono ciò cui ci si riferisce.

Poiché si vuole prospettare che lo studio della mente appartiene di diritto alla scienza naturale, è immediato il rifiuto di proprietà di stati mentali che non abbiano natura fisica.

Ma si noti che:

(a)

premessa (SE...) non è di per sé vincolata esclusivamente alle proprietà fisiche degli stati *cerebrali*;

(b)

i teorici dell'identità fanno propria la tesi che ogni proprietà naturale è definibile in termini fisici.

Ora, quali sono i confini della fisica? A tutt'oggi la biologia dei primati non è riducibile a fisica. La chimica fisica è ormai una solida disciplina, al pari della biochimica, ma tra le proprietà delle molecole e quelle degli organi e tra queste e quelle degli organismi ci sono ancora molte lacune da colmare, e questo vale specialmente per quanto riguarda i meccanismi di crescita e auto-organizzazione del cervello di un primate. Questi meccanismi difficilmente sono isolabili dalle interazioni con l'ambiente corporeo e dalle interazioni di questo con la fisica, la chimica e la biologia dell'ambiente (senza per questo svaporare in un regresso all'infinito). Però, qualcuno potrebbe insistere, "in linea di principio, nulla esclude che ...", e i dati finora ottenuti sono quanto mai incoraggianti al fine di provare che ..." - No, questo è l'avvio di una difesa troppo facile di qualunque tesi *empirica*. Per restare alla fisica: "in linea di principio", la teoria unificata delle interazioni fondamentali dovrebbe ugualmente essere garantita o fortemente avvalorata, ma nessuno ha ancora oggi scoperto come, e proprio per questo ci si lavora. Come si lavora cercando indizi sperimentali, si lavora anche a ridiscuterne i presupposti teorici.

D'altra parte, "in linea di fatto", la fisica è qualcosa di molto meno uniforme di quanto farebbe comodo pensare. La fisica è la base della cosmologia odierna, ma questa a sua volta pone problemi alla fisica. In particolare, la cosmologia odierna configura stati della natura che sono non meno impegnativi, dal punto di vista ontologico, di quanto lo sia ammettere stati mentali

non identificabili con stati cerebrali, solo che li configura in un quadro esplicativo unitario, mentre chi come il dualista ha voluto separare nella maniera più netta, e definitiva, stati mentali e stati cerebrali lascia a bocca asciutta in linea di principio e in linea di fatto: *hic sunt leones*. In base a quel poco detto fin qui si può già notare che, ad (a), negare la separazione di stati mentali e stati cerebrali non implica affermare la loro perfetta identità (e su questo tornerò fra poco); inoltre, ad (b), “Tutta la scienza si può ridurre a fisica?” è una domanda cui si risponde solo sfogliando la margherita, o se preferite: giocando a bocce ferme. Ai filosofi piace farlo, a me no. Evidentemente, o non sono un filosofo oppure la filosofia non comincia quando finisce il movimento delle bocce scientifiche. Ci sono tre considerazioni che al riguardo mi stanno a cuore più di tutte quelle che poi seguono. Quindi tanto vale che le esponga subito invece di lasciarle come dessert.

La PRIMA CONSIDERAZIONE è appunto che una riduzione sistematica di una teoria diciamo di macrolivello a un'altra di microlivello comporta sempre dei principi di correlazione che sono a loro volta una cosa molto, ma molto, delicata. Soprattutto, la riduzione (a parte casi banali) diventa possibile solo quando la teoria di microlivello è stata arricchita di tutti quei concetti che sono richiesti per una riduzione soddisfacente. In altre parole, la teoria di microlivello non è più quella cornice generale che era prima, perché vi sono stati predisposti elementi *mirati* alla riduzione. Non sto dicendo che la cornice generale è, come tale, indistinta; sto dicendo che tra tutti i possibili stati di aggregazione che consente, si devono specificare le condizioni che rendono possibile proprio quelli che ci interessa spiegare (perché sono quelli che di fatto si sono realizzati). Per fare un esempio, quando si dice che la nozione di numero reale si riduce a quello di numero naturale (arimetizzazione dell'analisi), lo si può dire solo perché si sfruttano le risorse della teoria degli insiemi, che non fanno parte dell'aritmetica elementare. Che cosa, nel caso di mente e corpo, ci sia a svolgere il ruolo della teoria degli insiemi non è affatto chiaro. In altre parole: ogni riduzione funziona come una pompa idraulica: non crea e non distrugge nulla, ma trasferisce. Il mediatore del trasferimento può essere un soggetto teorico altrettanto vasto e impegnativo del bacino sorgente e del bacino destinazione. Il vantaggio di ogni operazione del genere, se ben congegnata, è l'unità del sapere (che nel caso in questione si declina come chiusura causale del mondo: non ci sono cause di eventi naturali (fisici) che non siano eventi naturali (fisici), altrimenti sarebbe violato il principio di conservazione dell'energia – e questa violazione comporta, a catena, tutta una serie di

inesplicabilità, anche laddove proprio non avremmo voluto (nel caso che intendessimo negare la riduzione). Ma non si deve essere ipocriti: lo scarto tra la cornice generale (quella del microlivello e del megalivello) e le isole di struttura (che portano al mesolivello) così identificate e gerarchizzate resta; si tratta di descrivere al meglio questa gerarchia di scarti, piuttosto che farla collassare in un corto circuito immediato. Consentitemi una brevissima chiosa al richiamo che ho fatto al principio di conservazione: non è necessario interpretarlo, come non lo è per ogni altro principio teorico, in termini di causa-effetto. Il modello nomologico-deduttivo della spiegazione, pur con tutti i suoi difetti e i raffinamenti che richiede per estenderlo al caso delle spiegazioni probabilistiche, esclude il riferimento alla nozione di causa. Che in filosofia della mente si continui invece a fare questo riferimento dovrebbe quanto meno indurre gli interessati a motivare la conservazione del rapporto causa-effetto. Invece, mi spiace notare che la cosa è data immancabilmente per scontata: vergogna! Nel modello della mente come sistema dinamico – alcuni dicono con inutile leziosità “modello dinamicista” – si affronta la questione in maniera più coerente, e con ciò siamo alla seconda considerazione.

La SECONDA CONSIDERAZIONE è che nella mente non ci sono solo *stati* ma anche *transizioni-di-stato (processi)*. La mente è un sistema dinamico. Una teoria dell'identità che lasciasse fuori i processi sarebbe solo pubblicità. Quindi si dovrebbe prestare maggiore attenzione (per usare un eufemismo) alla questione dei processi. Di che natura sono? Quali strutture manifestano? Quali sono le loro differenziate condizioni di possibilità? Il che significa in primo luogo capire quali sono i processi mentali (a partire da quelli più elementari della percezione e della cognizione) e capire *quali* sono le transizioni di stato sul piano neurochimico per poter andare avanti nella discussione.

Se no si parla a vuoto. E con ciò siamo alla TERZA CONSIDERAZIONE: la filosofia della mente e in particolare la teoria dell'identità, con tutte le loro pur debordanti diatribe, sono cresciute in uno splendido isolamento dalla psicologia cognitiva. Cioè, si è continuato a parlare degli stati mentali a livello intuitivo, di modo che resta nebuloso *che cosa* dovrebbe essere identificato con stati cerebrali, anche qualora questi fossero identificabili *chiaramente e in ogni caso* – cosa che a sua volta non è, perché l'architettura del cervello è dannatamente complessa.

Per inciso, le ricerche dei neuroscienziati sono forse state motivate, stimolate, guidate dai lavori dei teorici dell'identità? Non mi pare. Alcuni dei problemi oggi più studiati riguardano la neuropsicologia, ma se ci siamo dimenticati la psicologia, non possiamo certo ricordarci della *neuropsicologia*.

3. Psicologia + filosofia = zero

L'operazione di insabbiamento della psicologia, e mi riferisco innanzitutto alla psicologia cognitiva, ha proceduto con grande astuzia, grazie alle risorse offerte dall'analisi del linguaggio. Proprio qui sta il peccato originale della teoria dell'identità, ma più latamente di quasi tutta la filosofia della mente, passata e presente, perché la filosofia della mente e nella fattispecie la teoria dell'identità così come si è configurata negli ultimi 60 anni, attraverso i lavori tradotti da Salucci, è un'articolazione della filosofia analitica; e la filosofia analitica, di ogni marca, è nata con l'esclusione dei processi cognitivi da ogni possibile discorso logico-filosofico sui pensieri linguisticamente espressi.

(Beninteso: c'è e c'è sempre stata una rigogliosa fioritura di letteratura filosofica sulla mente – sull'anima, sullo spirito, sull'ego, sulla coscienza o su quel non-so-che che ho dimenticato dalla lista – che ora vedeva come una minaccia barbarica l'indagine scientifica ora ne pronosticava le magnifiche sorti e progressive, ma non era "filosofia della mente" così come l'intendiamo oggi perché non si serviva in maniera essenziale dell'analisi del linguaggio.)

Non è un caso che quando si discute con gli psicologi, la loro reazione ad argomenti filosofici pro o contro l'identità è di uno scetticismo ironico, come a dire: voi filosofi, pur con tutto l'armamentario della logica, non avete la minima idea di quanto sia difficile studiare gli stati e i processi mentali e invece vi lanciate in programmi o contro-programmi ambiziosissimi, per realizzare una minima parte dei quali dovrete invece sporcarvi le mani con la percettologia, lo sviluppo infantile, la clinica, ecc. Il guaio è che poi gli psicologi non possono mica disinteressarsi delle basi della loro disciplina e così si avventurano in considerazioni delle quali non intravedono la natura filosofica (e i relativi trabocchetti) oppure l'hanno presente ma se la cavano troppo a buon mercato.

Su questo punto, la teoria dell'identità avrebbe potuto avere un ruolo di grande utilità, ma *non l'ha avuto*; perché appunto disdegnava il commercio di idee con gli psicologi cognitivi (e gli stimoli derivanti dai loro guai). D'accordo, non dovrebbe esserci neanche bisogno di dire che una "buona" filosofia della mente ha da tenere presenti le varie dimensioni del problema

mente/corpo. Invece c'è. Lascio a voi di completare l'entimema. Per non nascondermi dietro a un dito e non dare l'impressione di apparecchiare solo un bel discorso super-mega-epistemologico, aggiungo che le tre precedenti considerazioni si traducono in aspetti operativi della ricerca. È quel che ho cercato di far vedere, forse senza riuscirci come speravo, in *Mind and Causality* (John Benjamins, Amsterdam 2004), dove l'impianto di fondo, centrato sull'idea della mente come sistema dinamico interattivo, comporta in maniera essenziale i rapporti fra il cervello, il resto del corpo e i *pattern* di percezione-azione – quindi un modello naturalistico, ma non l'identità stati mentali = stati cerebrali. La mia idea, per quanto peregrina nel sostituirsi all'astuzia della ragione, era che anche i lavori raccolti nel volume che non sapevano di inserirsi in tale impianto in realtà lo suffragavano. Ma torniamo al punto su cui insistono i teorici dell'identità (TI d'ora in poi): non si può né teorizzare il dualismo né liquidare il problema mente-corpo con argomenti concettuali, a priori, logici, di analisi del linguaggio. Bene, ma gli argomenti con i quali i TI ci arrivano non sono argomenti empirici: sono argomenti non meno concettuali. A loro interessa provare che gli argomenti di tipo logico a sostegno dell'incoerenza dell'identità sono fallaci.

Ma come cercano di provarlo? Come si fa a dimostrare (p. 7) che “non ci sono motivi logici per rifiutare la tesi dell'identità”? Place si proponeva nel 1956 di mostrare che “la tesi secondo la quale la coscienza è un processo cerebrale non può essere respinta con argomentazioni logiche” (p. 13). Come poteva pensare di riuscirci e più in generale come ci si può riuscire? Due casi: ci limitiamo ai motivi logici *finora addotti*, ma in tal caso non si può escludere che un altro motivo logico, questa volta efficace, venga scoperto domani; oppure mettiamo in piedi un argomento che copra *tutti i possibili* motivi logici. I TI ragionano come se questi due casi fossero uno solo. La furbizia non soddisfa la ragione. In realtà i TI facevano e fanno ancora di meno: si preoccupano esclusivamente degli argomenti addotti contro l'incoerenza dell'identità mente corpo (o meglio, fra stati mentali e stati cerebrali). Sarebbe come se nella discussione sui fondamenti della matematica chi avesse trovato un sistema per escludere i paradossi finora noti affermasse che ha dimostrato che la matematica è coerente (o almeno la possibilità che così sia).

Non so di nessuno che abbia affermato una cosa del genere. Inoltre, i TI fanno un'aggiunta caratteristica: così come l'identità non è esclusa a priori, non è neppure provabile a priori. Se è vera può solo esserlo a posteriori. Ma loro adducono argomenti a posteriori (empirici, fattuali) a questo scopo? No. E questo costringe a fare qualche considerazione storica.

4. Neoempirismo senza comportamentismo?

Infatti, proprio qui si vede la filiazione dal neoempirismo più di quanto sembri a prima vista. Un olista come Quine non scinderebbe mai argomenti a priori pro o contro la tesi dell'identità e argomenti a posteriori, perché se sono argomenti che permettono di spiegare qualcosa appartengono all'ambito del sapere scientifico e *nella scienza* non c'è un confine netto tra i due tipi di argomenti, o se volete ci sono argomenti di entrambi i tipi.

Per fare un solo esempio: con quale tipo di argomento Galileo contesta l'idea che un corpo più pesante arrivi al suolo con maggior velocità di uno leggero, una volta lasciati cadere dalla stessa altezza? Con una prova empirica? No, perché la presenza dell'aria provoca un attrito diverso nei due casi. Lo fa con una riduzione all'assurdo, considerando due sfere di peso diverso vincolate fra loro da una catena. Questa separazione (e con essa la filiazione neoempiristica) si manifesta anche nel modo di concepire il ruolo degli argomenti filosofici. Con perfetta ortodossia, Feigl attribuiva alla filosofia il compito preliminare di una "chiarificazione logica ed epistemologica dei concetti" (p. 8) e gestiva poi la difficoltà di metter d'accordo senso comune e scienza con la dottrina dei due linguaggi più il criterio dell'economia ontologica (rasoio di Occam).

La chiarificazione cercata passava ovviamente per l'analisi linguistica. Dunque, l'analisi linguistica avrebbe dovuto aprire la strada a quella propriamente scientifica. L'ha aperta o la sta aprendo? In che misura? Non saprei dire (lo chiedo nuovamente a Salucci) ma, se così è, lo è esattamente nello stesso modo dei dualisti seri, cioè identificando specifici fenomeni e mostrando come si possano spiegare al meglio, o come non si possono spiegare così e così. Per coerenza, bisognerebbe anche domandarsi: un'analisi concettuale che prova i limiti dell'analisi concettuale non mette a repentaglio se stessa? L'analisi logica non può dimostrare a priori *nulla* circa il mondo? Ma allora l'argomento con cui si voleva legittimare l'indagine empirica sulla mente da un lato non poteva rifarsi ad alcun esito di questa (oltre all'analisi concettuale), dall'altro, lo stesso argomento, in virtù proprio di quanto afferma, non può dimostrare a priori nulla, in questo caso che gli argomenti logico-concettuali contro l'identità sono sbagliati.

(Altro quesito per Salucci che mi viene in mente a questo proposito: i teorici dell'identità, e più specificamente i suoi padri, si sono mai preoccupati di confrontarsi con il Teorema di Gödel?)

La questione però non finisce qui, perché nell'argomento dei TI e nell'analisi preparatoria (concettuale) c'era già in effetti un cavallo di Troia del più-che-concettuale. Cosa? La traccia del metodo comportamentistico. Però, le spiegazioni di tipo comportamentistico ovviamente non andavano bene, perché escludevano di poter guardare dentro alla scatola nera, cioè di ammettere come legittimi oggetti d'indagine gli stati interni della mente.

Di solito (e lo illustra anche una diffusa introduzione alla filosofia della mente come quella di Bechtel, pure tradotta in italiano da Salucci) si distingue un comportamentismo filosofico da uno scientifico, e più precisamente psicologico: il comportamentismo filosofico riguarda l'analisi del significato di enunciati su stati mentali e comportamenti, il comportamentismo scientifico, come teoria psicologica, riguarda il modo in cui uno specifico metodo e modello della psicologia si articola in termini di meccanismi stimolo-risposta e rinforzo/inibizione. Storicamente, per rispetto della comodità dei filosofi, è fin troppo giusto. Teoreticamente lo è ben poco, perché (1) senza il comportamentismo come dottrina empirica il comportamentismo filosofico non avrebbe avuto la forza che presumeva di avere, (2) il comportamentismo in psicologia ha già per conto suo dirette conseguenze circa l'analisi del significato, perché della psicologia fa parte anche lo studio del linguaggio in tutti i suoi aspetti, quindi anche la capacità semantica. Questo è chiaro a partire da Bloomfield – e siamo nel 1933! – e ancor più con Skinner (1957). Per i filosofi lo è soprattutto, anche se in una versione *sui generis*, allorché questo nesso si è esplicitato nei lavori di Quine dal 1952 al 1960.

Negli anni cinquanta, però, il comportamentismo si stava già aprendo alla considerazione del "comportamento" ... dei neuroni. E questo è un *trait d'union* che non bisognerebbe dimenticare. Per fare un solo esempio: è in un'ottica comportamentistica che si collocava anche la ricerca di Hebb, la cui "regola" per l'apprendimento da parte di una rete di neuroni (naturali o artificiali) è poi diventata famosa. Una precisazione è d'obbligo: ammettere stati interni che causano comportamenti implica il dualismo? No, ricorda giustamente Salucci, lo implica solo se gli stati interni non sono stati fisici. Ma, aggiungo, c'è una ulteriore lacuna: gli stati interni che sono causa di comportamenti potrebbero essere a loro volta causati da comportamenti. Chi nega l'implicazione deve dunque dire se ci sono o no stati interni (fisici) che non sono causati dall'esterno. Se dice che non ci sono, la differenza dal

comportamentismo si annulla. Se dice che ci sono, è tenuto a spiegare in che senso intende la *causalità* e quali sono le soglie tra bacini causali nel sistema accoppiato mente-mondo. Non mi pare che i TI ci abbiano pensato. Né hanno pensato a un altro guaio: la fisica del Novecento che dovrebbe servire a descrivere gli stati fisici cui si riducono gli stati mentali ha sbandierato più volte il superamento della stessa nozione di causalità come di un fossile. (Il suo accennato modello nomologico-deduttivo ne ha tratto le conseguenze.) Occorreva allora un chiarimento di come fosse legittimo, dato che era necessario per i TI, usarla ancora. In realtà, per la filosofia della mente, e proprio nel senso che interessava ai TI, ha fatto di più Chomsky (e la teoria della competenza linguistica come ancorata a un "organo biologico") di quanto potessero fare le aperture allo studio dei meccanismi stimolo-risposta-rinforzo/inibizione nel comportamento dei neuroni. Ciò per una ragione semplicissima: perché ha permesso, pur non intenzionalmente, di focalizzare il discorso sui danni cerebrali corrispondenti a deficit linguistici (ritardi ereditari e non acquisiti, afasie ecc.) – il primo a pensarci fu Lenneberg.

5. Questioni semantiche in generale

Sempre a proposito del comportamentismo, c'è infine da sfatare un equivoco. Sfatarlo ci permetterà di entrare nel merito delle questioni semantiche coinvolte nella teoria dell'identità. Come gli empiristi del Settecento ammettevano strategie intrinseche d'apprendimento (per associazione), anche i comportamentisti dovevano ammettere che ci sono non solo comportamenti effettivi, ma *disposizioni al comportamento*, e queste *disposizioni* sono interne. Il problema è che per loro non c'era altro. Anzi, l'identità di una disposizione era definita *estensionalmente* (in termini di cause-effetti, stimoli-risposte, comportamentali).

Feigl invece recuperava la distinzione fregeana fra senso e riferimento: gli enunciati del linguaggio mentalista si riferiscono agli stessi stati di cose degli enunciati del linguaggio fisicalista, ma in modo diverso.

(Due precisazioni per inciso: 1. la distinzione tra senso e riferimento non può essere confusa con quella tra due linguaggi che parlano di una realtà neutrale, per il semplice motivo che la dottrina dei due linguaggi può essere sottoscritta negando la distinzione di Frege, mentre la stessa distinzione può essere sfruttata, applicandola al metalinguaggio in cui si parla del linguaggio fisicalista e di quello mentalista, per negare che i due linguaggi siano o possano essere disgiunti; 2. quando Place considerava "palesamente falso" che gli enunciati sulla

coscienza sono enunciati sui processi cerebrali, se non stava esprimendosi in modo *palesamente fuorviante*, voleva sottolineare che il senso "inteso" degli uni non coincideva con quello degli altri, quindi non gli bastava la sola dimensione referenziale.)

A ciò si connette un'altra inferenza fallace, in linea con il rasoio di Occam: tutto ciò che si dice della mente e tutto ciò che si dice del cervello *ha come riferimento il cervello*, ma a rigore si sarebbe potuto dire esattamente il contrario (*ha come riferimento la mente*), adottando una linea che fu già di Leibniz. La negazione del dualismo è il monismo, ma non c'è un solo tipo di monismo, non importa qui se degli enti o della proprietà.

(Tipi di enti vs tipi di proprietà: una distinzione che, se non è soltanto un fossile empiristico ... e di quelli che piacciono tanto ai neomedievali oxbrigensi, esige una teoria che nessuno ha ancora elaborato a dovere. Chi vuol esser analitico, sia ... Non si deve, però, fermare all'abc dell'analisi logica del linguaggio.)

Ecco un altro dei guai di un ragionamento puramente concettuale che vuol essere sottile, ma non lo è quanto dovrebbe. Lo rilevo con un altro ragionamento concettuale, sì, ma siccome non presumo una linea di demarcazione fra due tipi di analisi, non ho doveri di coerenza da rispettare.

E l'appello alla distinzione fra senso e riferimento? Un altro presupposto lasciato a sguazzare nel vuoto, mentre andrebbe giustificato con argomenti solidi. Vi prego di notare che all'epoca (fine anni cinquanta) questi argomenti erano latitanti, anzi più che dubbi alla luce delle critiche di Quine. Non solo: lo stesso presupposto, per coerenza, avrebbe *dovuto essere*, oltre che giustificato concettualmente o empiricamente, essere anche spiegato in termini della teoria dell'identità. Le critiche di Quine erano davvero superabili con l'avvento della semantica dei mondi possibili? La risposta è controversa, ma anche se non lo fosse, è dubbio che l'allargamento modale servirebbe ai TI. Insomma, se la teoria dell'identità vuol essere il perno di una buona filosofia della mente, (i) ha bisogno di una buona semantica e in più (ii) è anche tenuta a mostrare che il 'buono' di questa semantica non solo va d'accordo con l'impostazione filosofica adottata ma ne scaturisce. Non mi pare che i due desiderata (i) e (ii) siano mai stati presi seriamente in esame. Quando lo sono stati (l'uno separatamente dall'altro), per esempio con Kripke e Lewis, i risultati non puntavano nell'esatta direzione del recupero della distinzione fregeana. Sembra proprio che io stia sparando a zero sulla teoria dell'identità, ma in realtà sto solo cercando di aiutare i suoi sostenitori a liberarsi dal peso di assunti non necessari e a ridurre le aspettative, affinché la discussione si sposti su quel che

conta. Come segno di benevolenza, farò una *parentesina* che vuole sgombrare il terreno da un'accusa troppo facilmente devastante che è stata o è possibile che sia rivolta alla teoria.

6. Parentesina

All'interno dell'epistemologia post-neoempiristica, tutte le teorie del riferimento, inteso come qualcosa di determinabile solo relativamente a una cornice teorica, sono debtrici di una dimostrazione che il recupero della distinzione fra senso e riferimento è fattibile e solido. Che io sappia, una simile dimostrazione non è mai stata fornita. E la ragione è dovuta a una fallacia tipicamente commessa. Ecco la fallacia: si ragiona come se il riferimento (r) fosse così carico di senso (s) da dire che $s(a) \neq s(b)$ implica $r(a) \neq r(b)$. Classicamente questo equivale a $r(a) = r(b)$ implica $s(a) = s(b)$. E poiché si continua a sottoscrivere il principio fregeano secondo cui $s(a) = s(b)$ implica $r(a) = r(b)$, il risultato – se sottoscritto in nome di un'ontologia essenzialmente carica-di-teoria o di oggetti carichi-di-noemi non fa, qui, alcuna differenza - è indiscernibile dal collasso della distinzione tra s e r teorizzato dai referenzialisti (causali o meno). Tale collasso è la bancarotta per chiunque creda che la teoria del significato dei post-neoempiristi è più fine di quella neoempiristica o che la teoria del significato di Husserl è "più fine" di quella di Frege, perché "più fine" infatti implica conservare la distinzione s/r e appunto raffinarla, non annullarla. Quindi chi ragiona così non può ragionare così. Ma non può neppure salvare la distinzione banalizzandola. Infatti, la differenza rispetto ai referenzialisti sta solo nel fatto che il collasso è rovesciato: ora, tutto è determinato dal senso. Siccome ci si guarda dallo specificare un criterio per l'identità di senso (sinonimia), il senso è inteso nella maniera più fine possibile, dunque produce classi di equivalenza che sono singoletti e di conseguenza ciò si applica al riferimento. Ma allora la tesi del riferimento carico-di-senso diventa banalmente vera, contro ciò che si voleva suggerire – e se non è banale, è indicibile, perché l'equazione $s = r$, se non banalmente vera, è falsa. In ambedue i casi la tesi è praticamente inutilizzabile.

E lo è perché il collasso porta a una nuova monadologia, che per coerenza non finisce dove si vorrebbe, ovvero con i confini tra una teoria e l'altra, ma permea ogni enunciato di *una stessa teoria*; pertanto la relazione di uguaglianza non ha più ragion d'essere. Purtroppo, nella conoscenza umana ha rilievo, e ancor prima legittimità non banale, la scoperta di uguaglianze. La scienza, modernamente intesa, e non dunque come una immensa tassonomia di qualità accessibili direttamente ai nostri sensi, è quel che è anche grazie alle equazioni che contiene. Una filosofia della scienza che elimina le equazioni o le banalizza o non ne sa ricostruire le

condizioni di rilevanza conoscitiva si lascia sfuggire qualcosa di importante. I TI non ringrazieranno certo per il fatto che ora, in considerazione del modo-in-cui-pensano-la-realtà, la tesi “stati mentali = stati cerebrali” può essere considerata empiricamente corretta perché è l'empiria-così-come-si-dischiude-ai-loro-occhi quella che conta; diranno piuttosto che è una vittoria destituita di valore. Né saranno i loro rivali a ringraziare perché, per la stessa ragione, la loro fatica argomentativa poteva ridursi al semplice riconoscimento di un diverso modo di parlare e pensare.

Chiusa parentesina.

C'è qualcosa in quest'ultime considerazioni che può servire in relazione al dibattito connesso con l'argomento di Jackson? Non ci ho pensato. Passo ancora una volta la palla a Salucci. Però non posso esimermi da una minima chiosa sui “qualia”, sempre per compensare l'impressione di star sparando a zero contro la teoria dell'identità. La respingo, sì, ma non per le ragioni dei dualisti, e in filosofia, così come nella scienza, le tesi contano se contano le ragioni. Il problema è quello di fornire, in conformità con la teoria dell'identità, una spiegazione delle qualità con cui appaiono alla nostra propria coscienza i nostri propri stati.

(Chiedo di passaggio a Salucci: una disposizione al comportamento non potrebbe avere le sue brave qualità d'accompagnamento, come facenti semplicemente parte del comportamento cerebrale in termini di catene causali fra neuroni?)

Per onestà, devo confessare che la questione delle qualità esperite in prima persona (*che cosa si prova a...*) non mi ha mai solleticato. Le discussioni in merito dubito siano mai arrivate a un punto minimamente solido e dubito pure che possano arrivarci fintanto che, come al solito, si travasano in funzione teorica o metateorica nozioni del senso comune che non ci si preoccupa di precisare né concettualmente né empiricamente.

(*) “Sì – è la tipica replica – ma dopotutto non c'è bisogno di questa precisazione, perché se c'è una cosa certa, è il modo in cui un *mio* stato interiore appare *a me*”.

Chiamo tale replica (e ogni altra consimile) “la barba di Cartesio”. È la barba che cresce con la presunzione di aver trovato nell'intimità cosciente dell'autoriferimento la chiave d'accesso alla suprema certezza, cui si accompagna la presunta certezza (!) che la chiave è persa non appena si esce da questa intimità. La cosa è tanto più curiosa, o sospetta, ai nostri giorni

rispetto a quando la barba cominciò a crescere, perché se lo scetticismo circa ogni oggettività empirica c'era già allora, almeno Cartesio faceva della certezza del cogito un fondamento per costruirci sopra qualcosa, mentre oggi chi si fa crescere la sua barba (per esempio, puntando sui *qualia*) non se ne serve per costruire qualcosa, bensì mette su un argomento che poi, se applicato con coerenza, si ripercuote ben al di là dell'ambito inteso, con effetti devastanti sulle spiegazioni scientifiche. Come dato ricavabile dalla letteratura, si potrebbe anzi dire che chi oggi si fa crescere la barba di Cartesio ha a cuore soltanto di negare la fondabilità o la fondatezza oggettiva di qualunque conoscenza non riguardi il suo mondo privato così come appare in presa diretta al *de cuius*.

(Per inciso, siccome non si fanno processi alle intenzioni, mi attengo alla problematicità della replica (*) su sinteticamente riassunta, che, a rigore, non è chiaro neppure se è veicolabile a parole, cioè, non è chiaro se quel che dice può mai corrispondere a quel che si vorrebbe che dicesse. E come non si fanno processi, non si legittima una posizione teoretica con i suoi intenti. Ma ... se può servire, tengo a precisare che sono lungi dal voler proporre una qualche forma di riduzionismo o di formalismo neutrale. Tant'è vero che considero i concetti di spiegazione e di dimostrazione come del tutto vuoti indipendentemente dalla consapevolezza di qualità cognitive. La scienza moderna nasce con – tra le altre cose – la trattazione quantitativa delle qualità, ma prosegue anche con una trattazione matematica delle strutture qualitative, che non è necessariamente quantitativa, nel senso delle quantità numeriche solitamente prese a modello. Per fare un solo esempio, il concetto di gruppo di simmetria è *qualitativo*.)

Qui potrebbe invece tornare utile, contro l'ulteriore crescita dalle barba, il ben noto argomento di Wittgenstein in *Della certezza*. Ma siccome non c'è tempo per entrare nel merito, mi limito a notare che la teoria dell'identità mostra la sua utilità almeno come barbiere. È proprio il *modo* in cui i nostri stati ci appaiono, prima della relativa certezza, che richiede di essere specificato empiricamente – non all'ingrosso però, protetti dalla nostra normalità; bensì attraverso lo studio dei danni cerebrali in relazione a deficit cognitivi. I TI hanno espresso un programma che al riguardo è deficitario, ma sempre meglio che *hic sunt leones*.

Ecco perché in passato ho più volte insistito sull'importanza dei dati clinici relativi ai deficit di propriocezione e le sindromi che portano alcuni soggetti a trattare i loro pensieri come pensieri di un altro. Ci sono deficit cognitivi (in quale dominio) che non hanno alcun riscontro

cerebrale (in quale area)? Questa dovrebbe essere la domanda da porsi. Invece si continua a discettare in termini vaghi. Onde non apparire dogmatico e d'altra parte non farla troppo lunga, mi limito a rimandare alle osservazioni di Dennett e Hofstadter sul pur gustoso saggio di Nagel intitolato "Che cosa si prova a essere un pipistrello?" Vorrei invece tornare al punto da cui ho cominciato questa pur ondivaga serie di riflessioni. Place, ci ricorda Salucci a p. 13, non cerca "di dimostrare la verità dell'ipotesi dell'identità, ma cerca di mostrare che (a) non è necessariamente falsa e che, dunque, (b) è un'ipotesi plausibile, la verifica della quale è (c) affidata all'indagine empirica." Ma se questa è la strategia, è sicuramente adolescenziale: dall'impossibilità di dimostrare l'assurdità di qualcosa non scende la sua potenziale verità. Faccio un esempio: nessuna affermazione esistenziale, tipo "esistono solo 75 cose", "Nella galassia di Andromeda c'è un pianeta con una forma di parmigiano reggiano" ecc. non può essere confutata a priori, ma da questo non scende che a crederci ci guadagniamo qualcosa ai fini di una qualsiasi spiegazione.

Se questo appunto è marginale, non lo è un elemento ricorrente negli argomenti dei TI, cioè, l'asserita *contingenza* di ogni identità: l'acqua è H₂O, il lampo è un movimento di cariche elettriche e *dulcis in fundo* la mente è un insieme di stati fisici. Anche facendo finta che nel dire questo si tenesse conto non solo degli stati ma pure dei processi, l'asserita contingenza provoca a mio modesto parere più problemi di quanti si potesse sperare di risolvere: contingenza fino a che punto? Fino alla più arbitraria *casualità*? Ma quali proprietà, che non siano legate a convenzioni sociali o limitate a specifici eventi singoli (facendo come al solito astrazione da tanti fattori) sono coestensive e allo stesso tempo puramente contingenti? Se fosse così, quale vantaggio ci sarebbe a dire che le proprietà mentali coincidono con proprietà cerebrali? (Altrettanto contingentemente, potrebbero esserci stati mentali in una slavina sul Monte Rosa.)

7. Questioni semantiche specifiche

Infatti, nell'ottica della contingenza, già che ci siamo, potremmo anche scendere di un gradino e chiederci perché l'attivazione di una certa colonna di neuroni nella corteccia visiva corrisponde alla vista di una certa linea orizzontale o verticale. Anche questo è un fatto *contingente*? Forse i TI stanno trasformandosi in funzionalisti senza saperlo? Inoltre, come l'identità fra stati mentali e cerebrali, perché non è contingente anche il fatto che il calore di un gas è l'energia cinetica media delle sue molecole? Qualunque senso io riesca a dare alla

identità “per composizione” come diversa da quella “per definizione”, è un senso che non avalla direttamente la tesi che l’identità per composizione è contingente. La avalla indirettamente? Ma allora i TI dovrebbero dirci *come* e non mi pare che ce lo dicano. (Chiedo l’ennesima conferma a Salucci.)

Tre chiarimenti sono opportuni in proposito.

PRIMO CHIARIMENTO: confesso di non aver mai capito cosa intenda affermare chi dice che “non è contraddittorio che l’acqua non sia H₂O”, ovvero: l’acqua potrebbe non essere H₂O. (Prego il gentile pubblico di non pensare che sto dimenticando il fatto che l’acqua che beviamo contiene anche molte sostanze disciolte, come riportato sulle etichette delle bottiglie di acqua minerale.) Proviamo ad aiutare chi lo dice.

- Sta facendo un’osservazione sull’arbitrarietà del segno? Banale.
- Sta facendo una considerazione sul rapporto tra linguaggio ordinario (L₀) e linguaggio di una teoria chimica (L_T)? Allora, per favore, metta le virgolette differenziate con indici (L₀) e (L_T) tutte le volte che parla di un concetto linguisticamente espresso; anzi, già che c’è, lo faccia per tutte le parole che sta usando. Impraticabile.
- Sta esprimendo una tesi sulla chimica di questo e di altri universi possibili? Ci dica allora a quale teoria chimico-cosmologica è dovuta questa ipotetica possibilità e in base a quali criteri raccordare l’uso di “acqua-qui-nel-mio-mondo” fatto da soggetti che appartengono a mondi diversi in cui una volta l’acqua è XYZ invece che H₂O, ma i soggetti che sono la nostra controparte hanno tutti gli stessi stati mentali che abbiamo noi quando pensiamo all’acqua, e un’altra volta l’acqua è ancora H₂O, ma i soggetti che sono la nostra controparte non hanno tutti gli stessi stati mentali che abbiamo noi quando pensiamo all’acqua. Credenziali scientifiche (circa la *biologia* umana) smarrite.

Ogni dottrina filosofica che divida tutto quanto in argomenti logici su contraddittorità o non-contraddittorità, e argomenti empirici, si scorda che chi, come il buon Carnap, praticava l’analisi concettuale, quando contrapponeva all’esperienza e alle verità osservativo-sperimentali un tipo peculiare di verità, non intendeva solo le verità logiche ma anche quelle

concernenti i significati dei termini. Ci si dica allora se la contraddittorietà a favore o contro la quale si ragiona è davvero quella esclusivamente logica o no, perché, se non lo è, quali sarebbero mai le verità analitiche circa la mente o il cervello?

SECONDO CHIARIMENTO: chi si sia interessato di filosofia del linguaggio potrebbe pensare che io dico questo perché voglio sostenere il carattere necessario di ogni equazione nomologica, nella linea della teoria kripkiana del riferimento rigido, seguita (una volta tanto) coerentemente fino in fondo. Ma non è così: voglio piuttosto contestare la sensatezza di parlare di necessità e di contingenza in astratto, cioè su un presunto piano puramente logico. In breve: o la necessità e la contingenza sono nozioni relative ai nessi inferenziali e non ai contenuti degli enunciati, oppure sono cariche di contenuto epistemico (possibilmente non micromonadologico). Lascio a voi la scelta. Ho un'idea al riguardo, ma porta a una concezione, un po' stramba se volete, della competenza logica e ad una teoria di come logica ed esperienza cinestesica si interfacciano – "stramba" perché costituisce una duplice anomalia rispetto al modo "standard" di impostare l'analisi delle nozioni modali e di valutare il loro impatto sull'epistemologia ... e questo modo entra pesantemente nella discussione sulla teoria dell'identità. Quale che sia la vostra scelta, in un caso e nell'altro né gli argomenti che sono bersaglio dei TI né gli argomenti dei TI risultano attendibili. Il guaio della filosofia, anche in questo caso, è che si perde un sacco di tempo a mostrare che certi argomenti non funzionano e, invece di uscire dalla scacchiera truccata, si continua a farle la guardia, così non resta tempo per cominciare a fare la minima ricerca che potrebbe precisare i termini delle questioni affrontate e conseguire qualche conoscenza degna di una pur minima considerazione. Se quando parliamo di mente, conoscenza, significato, pensiamo di esprimere pensieri che siano o possano diventare conoscenze (se no, che cosa sono: chiacchiere da salotto, pettegolezzi, umori?), non ci dobbiamo fermare sulla soglia. Ma se andiamo avanti, usciamo dalla filosofia come consuetamente intesa e insegnata.

TERZO CHIARIMENTO: sono fra coloro che non sottoscrivono l'identità fra stati mentali e stati cerebrali perché sottoscrivo la tesi della multirealizzabilità, cioè, penso che si possano riprodurre almeno alcuni stati mentali in sistemi artificiali, cyborg o futuri OGM che non hanno da cima a fondo la stessa architettura cerebrale di *homo sapiens sapiens* (ma per quel che può contare, di un macaco o di un cane). Con ciò tengo fisse le proprietà cerebrali e

definisco funzionalmente le proprietà mentali. Si potrebbe fare l'inverso? In tal caso dovremmo pensare a sistemi con la (nostra, diciamo) architettura cerebrale definita funzionalmente e ancorare strettamente la nozione fenomenologica di stato mentale. Potrebbe anche darsi ..., ma se la variazione nell'implementare stati mentali è già problematica, questa lo è ancora di più. A scanso di equivoci, è un'ipotesi ben diversa dal funzionalismo, perché (1) non sottoscrivo in alcun modo la completa separabilità di software e hardware e (2) mi preoccupo di mettere le carte in tavola via via che scollo uno stato mentale da uno stato cerebrale, ma non l'insieme degli stati mentali dall'insieme degli stati cerebrali in un colpo solo – quindi la multirealizzabilità è vincolata. Sono i vincoli (dinamici) che una scienza naturale della mente ci aspettiamo metta appunto in luce. Lo so che quest'ultimo discorso esige un chiarimento su come si intendono le "proprietà causali" degli stati in esame, ma per pigrizia (e per non annoiarvi) lascio fuori la problematica relativa, che pure è presente nell'antologia curata da Salucci, per esempio nel saggio di David Lewis.

8. Identità in che senso?

I tre chiarimenti precedenti ci portano direttamente ad affrontare il problema: in che modo va intesa l'identità stati mentali = stati cerebrali?

Sull'identità in un linguaggio modale, come quello in cui presumiamo di esprimerci, ci sarebbe tantissimo da dire. Ma qui sì che dovremmo entrare in dettagli di carattere logico-matematico. Restando sulle generali, il principio di Leibniz (identità degli indiscernibili / indiscernibilità degli identici) è un principio più che controverso. Viene usato in entrambe le forme soprattutto in versione contrapposta: se a e b sono discernibili mediante almeno una proprietà allora $a \neq b$, e se $a \neq b$ allora sono discernibili. Ma, a parte la questione di cosa si possa intendere con "tutte" quando parliamo di "tutte le proprietà", si noti che se in più vogliamo distinguere senso e riferimento, siamo nei guai grossi. Nel caso che l'essere identificata come la stella del mattino sia una proprietà di Venere e sia una proprietà diversa dall'essere

identificata come stella della sera, allora non si può più dire che *la stella del mattino = la stella sera*. Inoltre è un principio comunemente accettato della logica modale che (*) se si è dimostrato $a = b$ allora si può dimostrare che necessariamente $a = b$. E ci sono logiche in cui

vale anche che (**) se si è dimostrato $a \neq b$ allora si dimostra che *necessariamente* $a \neq b$, ma non è garantito che (**) scenda da (*).

Il guaio qui è: cosa vuol ancora dire “necessariamente”? Al riguardo non si può barare pur di scrivere articoli filosofici. Le lingue naturali in cui si fa filosofia (anche della mente) sono semanticamente chiuse ma sono anche sintatticamente chiuse. Perciò le teorie che hanno l'una o l'altra proprietà di chiusura hanno bisogno di essere trattate con doppi guanti bianchi per evitare paradossi. Ma i filosofi parlano simultaneamente di tutti i concetti che vogliono, sfruttando insieme le risorse del linguaggio naturale e quelle di qualsiasi teoria faccia loro comodo. Questo è barare.

Place ci ricorda le diverse accezioni di “è”, ma con ciò non dice nulla che non si sapesse già ($=, \subseteq$ e \in vanno distinti). L'identità che ci interessa, per trattare il caso in cui ci siano a e b tali che $a = b$ ma sia possibile che $a \neq b$, non è un'identità diversa da quella intuitiva. Ciò che conta è invece distinguere tra il fatto che $a = b$ sia dimostrabile in una teoria e sia invece semplicemente vero in uno dei suoi modelli. In un articolo del 1989 (“The theory of descriptions revisited”) ho classificato la fenomenologia di casi del genere e l'ho fatto in termini di *quali* proprietà strutturali sono preservate o no da *quali* trasformazioni dell'universo di discorso.

Rileggendo l'Introduzione di Salucci, mi sono ricordato che una delle motivazioni era proprio quella di uscire dal ginepraio cui la contrapposizione fra teorie descrittiviste del riferimento (Frege-Russell-Strawson) e teorie del riferimento rigido (Kripke) ci condanna. Infatti, la considerazione di come variano i modelli di una teoria, in modo da precisare entro quali margini $a = b$ sia *contingente*, è proficua anche per capire cosa veramente s'intenda con *lo stato mentale m = lo stato neurale n*. Perché? Perché siamo obbligati a esplicitare tutti i presupposti taciti del nostro uso di una qualsiasi proprietà caratterizzante univocamente qualcosa.

Con questo fastidioso autoriferimento mi fermo. Ci sarebbero diversi altri argomenti che Salucci affronta nella sua Introduzione e in un modo che raramente mi è capitato di trovare altrettanto efficace, per come sono riassunti, organizzati e reciprocamente soppesati, fornendo così un prezioso filo d'Arianna nel labirinto dell'odierna filosofia della mente. Tra essi: il mulino di Leibniz, il ruolo dei contesti “intensionali”, il problema “correlazione nomologica o identità”, la sopravvenienza. Sarà per un'altra volta.

Alberto Peruzzi

LA TEORIA DELL'IDENTITÀ: ALCUNI PROBLEMI EPISTEMOLOGICI. DISCUSSIONE CON MARCO SALUCCI

di Roberta Lanfredini

1. Premessa: la teoria dell'identità fra anti-dualismo e anti-eliminativismo

La teoria dell'identità si oppone sia al dualismo cartesiano sia al comportamentismo. Il che significa che i teorici dell'identità sono al tempo stesso anti-eliminativisti (gli stati mentali esistono, non sono illusioni) e anti-dualisti (gli stati mentali non sono ontologicamente distinti da stati fisici).

La teoria dell'identità non nega quindi gli stati mentali, ma fonda la propria specificità sul tentativo - quasi un gioco di equilibrio - di affermare l'esistenza di stati mentali da un lato e l'identità fra stati mentali e stati cerebrali dall'altro.

Non esiste un qualche attrito, è lecito chiedersi, nel voler sostenere entrambe le cose? La risposta del teorico dell'identità è negativa, sempre che interpretiamo fisico e mentale in senso né *sostanziale* né *concettuale*.

- **Non sostanziale:** la distinzione fra stati mentali e stati fisici non implica una diversità ontologica fra due sostanze (mente o anima e corpo o cervello). La teoria dell'identità non esprime un monismo che si contrappone al dualismo cartesiano, ma dichiara la legittimità della sola distinzione fra *stati* mentali e *stati* cerebrali.
- **Non concettuale:** l'ipotesi dell'identità è un'ipotesi empirica. Il problema mente-corpo non può avere una risoluzione *a priori*. Ciò che ai teorici dell'identità interessa non è quindi sostenere l'identità fra stati mentali e stati cerebrali con argomenti logici o a priori, bensì sostenere tale identità utilizzando indagini empirico-scientifiche.

Come deve essere interpretata, quindi, la dichiarazione di esistenza di stati mentali? Secondo il teorico dell'identità in questo modo: esistono *enunciati* su stati mentali e tali enunciati non possono essere tradotti in *enunciati* su stati cerebrali. La riprova sta nel fatto che possiamo descrivere le nostre sensazioni senza sapere che esistono stati cerebrali: «un contadino analfabeta potrebbe benissimo riuscire a parlare delle sue sensazioni senza sapere nulla dei suoi processi cerebrali, proprio come può parlare del

lampo senza sapere nulla di elettricità».¹ Analogamente è possibile affermare senza cadere in contraddizioni che qualcuno ha uno stato mentale e simultaneamente negare che abbia uno stato cerebrale: d'altro canto, il teorico dell'identità sostiene che l'indipendenza *logica* fra i due tipi di enunciati non dice nulla sulla indipendenza *ontologica* degli oggetti a cui gli enunciati si riferiscono. Il fatto che esistano modalità di accesso, caratterizzazioni, processi differenti, non esclude che tali modalità di accesso, caratterizzazioni o processi si riferiscano in realtà *allo stesso oggetto*. Al contrario, la specificità dell'ipotesi dell'identità sta proprio nel dichiarare che stati mentali e stati cerebrali sono punti di vista distinti che convergono sullo stesso oggetto e non punti di vista distinti che convergono su oggetti distinti.

Sembrerebbe esserci un residuo intrattabile di *concetti* che si raggruppano intorno a quelli di coscienza, esperienza, sensazione e immagine mentale, per i quali è inevitabile chiamare in causa qualche genere di *processo interno*. È possibile, naturalmente, che alla fine si possa trovare un resoconto comportamentistico soddisfacente di tale residuo. Per i nostri scopi presenti, tuttavia, assumerò che questo non possa essere fatto e che gli enunciati sulla sofferenza e sul dolore, su come le cose ci appaiono, su come le udiamo e le sentiamo, su come le sogniamo o ce le raffiguriamo nell'occhio della mente, siano enunciati che si riferiscono a eventi o processi che sono in qualche senso privati o interni all'individuo di cui sono predicati. La questione che desidero sollevare è: nel fare una simile assunzione ci impegnamo inevitabilmente a sostenere una posizione dualista, per la quale sensazioni e immagini mentali formano una categoria di processi separata che sta oltre e al di là dei processi fisici e fisiologici con i quali si sa che sono correlate? Mostrerò che accettare la nozione di processo interiore non implica il dualismo e che la tesi secondo la quale la coscienza è un processo cerebrale non può essere respinta con argomentazioni logiche.²

Così facendo, i teorici dell'identità si fanno promotori di un *anti-riduzionismo epistemologico* ma di un chiaro *riduzionismo ontologico* (sia nel senso dell'anti-dualismo delle sostanze sia nel senso dell'anti-dualismo delle proprietà).

Rimane a questo punto da chiarire qual è il *tipo* di riduzione ontologica che viene qui proposto. Questione che ne richiama immediatamente un'altra, cioè qual è il senso di identità che viene adottato quando si dichiara che gli stati mentali *sono* stati fisici.

¹ J.J.C. Smart, *Sensations and brain processes*, in "Philosophical Review", 68, 1959; trad. it. *Sensazioni e processi cerebrali*, in M. Salucci (a cura di) *La teoria dell'identità*, cit. p. 79.

² U.T. Place, *Is consciousness a brain process?* In "British Journal of Psychology", 47, 1956; trad. it. *La coscienza è un processo cerebrale?*, in M. Salucci (a cura di), *La teoria dell'identità*, cit. p. 62.

2. Quale identità?

La teoria dell'identità dà a questo proposito le seguenti indicazioni. Il senso del «sono» nella affermazione «gli stati mentali sono stati cerebrali» deve essere interpretato non come *definizione* (es.: «il rosso è un colore» oppure «il quadrato è un rettangolo equiangolo»): la coscienza non è infatti uno stato cerebrale per definizione e l'identità non è necessaria né a priori; una mente disincarnata è perfettamente concepibile. Il senso di quel «sono» è piuttosto da intendersi o nel senso della *composizione* (es. «una nuvola è una massa di goccioline d'acqua e di altre particelle in sospensione»; oppure «il lampo è un movimento di cariche elettriche» o ancora: «acqua è H₂O») o del senso della *denotazione* (es. il classico esempio fregeiano: «la stella della sera è la stella del mattino»). In entrambi i casi, l'identità vale fra caratterizzazioni alternative dello stato di cose a cui le espressioni si riferiscono. E in entrambi i casi l'indipendenza logica non corrisponde a una indipendenza ontologica.

La teoria dell'identità, sia nel caso venga concepita come identità per composizione sia nel caso in cui venga concepita come identità per denotazione, mostra tuttavia alcune difficoltà epistemologiche che i teorici dell'identità sembrano non prendere adeguatamente in considerazione. Prenderemo qui in esame tre ordini di difficoltà:

1. difficoltà riguardanti la relazione fra osservazione e teoria;
2. difficoltà riguardanti la cosiddetta fallacia fenomenologica;
3. difficoltà concernenti l'applicazione del concetto di denotazione al caso della relazione fra stati mentali e stati cerebrali.

3. Osservazione e teoria

Il primo problema consiste nello stabilire quando due osservazioni possono essere considerate osservazioni dello stesso evento. I casi di identità per composizione segnalati dal teorico dell'identità risultano infatti essere molto differenti. Proviamo a prenderli in considerazione uno alla volta.

- Una nuvola è una massa di particelle in sospensione.

In questo caso l'identità può essere stabilita da un'ordinaria procedura di osservazione: l'osservazione della nuvola è dello *stesso tipo* di osservazione della massa di particelle in

sospensione. Si tratta, quindi, dello stesso atto di osservazione che si sposta da un macrolivello a un microlivello senza soluzione di continuità.

- Il lampo è un movimento di scariche elettriche.

In questo caso, per quanto si possa osservare da vicino il lampo non si osserveranno mai scariche elettriche. L'identità sussiste quindi, in questo caso, fra due entità che possono essere osservate con soluzione di continuità. Uno dei due termini (scariche elettriche) sta inoltre per una ipotesi teorica in grado di spiegare il tipo di stimolazione visiva (il lampo) a partire da un movimento di scariche elettriche.

- Lo stato mentale è lo stato cerebrale.

In questo caso la situazione è ancora diversa. Lo stato cerebrale non è infatti *direttamente* osservabile mentre lo stato mentale non è, in senso proprio, osservabile. Stando a quanto sostenuto da filosofi come Kant e Husserl, lo stato di coscienza non è oggetto di conoscenza. Per Kant una rappresentazione acquisisce statuto oggettivo, quindi conoscitivo, solo se è rappresentazione di qualcosa nello spazio. È quindi possibile avere una rappresentazione oggettiva, e di conseguenza conoscitiva, di tavoli o sedie, ma non dei propri desideri, timori, percezioni. Si tratta di una posizione radicalmente anti-cartesiana: per Cartesio io conosco la mia coscienza più direttamente e esattamente di quanto conosca qualsiasi altra cosa e forse (se non fosse per l'intervento di Dio) non potrei conoscere nessun'altra cosa. Per Kant, al contrario, si può dire con ragionevolezza che conosco ogni altra cosa più propriamente di quanto conosca me stesso: non si può essere passivi spettatori delle proprie operazioni interne.

Husserl, dal canto suo, al fine di distinguere fra l'essere intenzionalmente rivolti a oggetti «esterni» e l'essere intenzionalmente rivolti a oggetti «interni», introduce la differenza fra conoscenza e riflessione. Gli stati di coscienza possono essere resi oggetti di atti riflessivi ma non di atti conoscitivi né introspettivi. L'aver un accesso privato e diretto ai nostri stati mentali non significa, quindi, rendere gli stati mentali oggetto di osservazione. In conclusione, fra il rendere oggetto gli stati mentali e il rendere oggetto gli stati cerebrali vi è uno scarto: non si tratta di un'identità stabilita mediante un'ordinaria procedura di osservazione (se gli stati mentali non sono oggetto di osservazione, è impossibile verificare l'identità psicofisica mediante osservazione); né si tratta di una

identità stabilita mediante il ricorso a entità teoriche (gli stati cerebrali non sono entità teoriche ma entità indirettamente osservabili).

L'identità fra stati mentali e stati cerebrali non potrà quindi essere stabilita mediante un confronto fra ordinari processi di osservazione e procedure speciali di osservazione (come nel caso 1). Lo stesso Place giunge a riconoscerlo: «nel caso dei processi cerebrali e della coscienza non c'è continuità fra i due sistemi di osservazioni», il che significa che «un'indagine introspettiva ravvicinata non rivelerà mai il passaggio degli impulsi nervosi attraverso migliaia di sinapsi nel modo in cui invece l'osservazione ravvicinata di una nuvola rivelerà una massa di piccole particelle in sospensione»³.

D'altro canto, nemmeno il ricorso a un qualche tipo di teoria (come nel caso 2) ci permetterà di stipulare tale identità e questo per due ordini di motivi: gli stati cerebrali non sono entità teoriche e gli stati mentali non sono entità osservative.⁴ Come è possibile, quindi, parlare di *equivalenza empirica* se «le operazioni richieste per verificare gli enunciati sulla coscienza e gli enunciati sui processi cerebrali sono fundamentalmente diverse»?⁵

L'unica cosa che sembra possa essere detta è che fra eventi cerebrali e eventi mentali vi è una correlazione sistematica e costante, condizione che è tuttavia del tutto insufficiente per sancire l'identità.

Che cosa è, dunque, che ci conduce a dire che due insiemi di osservazioni sono osservazioni di uno stesso evento? Non può essere solamente il fatto che due insiemi di osservazioni sono sistematicamente correlati in modo tale che ogni qual volta c'è un lampo c'è anche un movimento di cariche elettriche. Ci sono innumerevoli casi di simili correlazioni in cui non abbiamo nessuna tentazione di dire che i due insiemi di osservazioni sono osservazioni dello stesso evento. C'è una sistematica correlazione, per esempio, fra il movimento delle maree e le fasi della Luna, ma questo non ci induce a dire che le registrazioni dei livelli delle maree sono registrazioni delle fasi lunari e viceversa. Parliamo piuttosto di una connessione causale fra due processi o eventi indipendenti.⁶

³ Ibid. p. 67.

⁴ Ciò rende inappropriata, in riferimento all'identità fra stati mentali e stati cerebrali, la considerazione di Place: «noi trattiamo due insiemi di osservazioni come osservazioni dello stesso evento in quei casi in cui l'insieme delle osservazioni scientifiche e tecniche stabilite dal contesto di un appropriato corpo di teorie scientifiche ci fornisce una spiegazione immediata delle osservazioni che facciamo comunemente. Così noi concludiamo che il lampo non è niente di più di un movimento di cariche elettriche poiché sappiamo che un movimento di cariche elettriche attraverso l'atmosfera, come quello che si verifica quando vediamo un lampo, determina quel tipo di stimolazione visiva che induce un osservatore a riferire di aver visto un lampo di luce» (p. 68).

⁵ Ibid. p. 67.

⁶ Ibid. p. 68.

La costante e sistematica correlazione è infatti una condizione minimale accettata anche dal dualista senza troppe difficoltà.

4. La «fallacia fenomenologica»

Il secondo ordine di problemi concerne la cosiddetta *fallacia fenomenologica*.

Analizzare accuratamente in che cosa consiste questa fallacia chiarisce bene la posizione dei teorici dell'identità e il loro collocarsi in uno spazio estraneo sia al dualismo sia all'eliminativismo. Prendiamo in considerazione la descrizione che ne fa Place:

Se assumiamo, per esempio, che quando un soggetto riferisce di una post-immagine verde egli stia dicendo che esiste un oggetto che è letteralmente verde dentro di lui, è chiaro che abbiamo a che fare con un'entità per la quale non c'è posto nel mondo fisico. Nel caso della post-immagine verde non c'è nessun oggetto verde nell'ambiente interno che corrisponda alla descrizione che il soggetto ne dà. Neppure c'è qualcosa di verde nel suo cervello; certamente non c'è niente che potrebbe essere emerso quando egli riferisce di aver visto la post-immagine verde. I processi cerebrali non sono il genere di cose alle quali possono essere propriamente applicati i concetti di colori.

La fallacia fenomenologica sulla quale si fonda l'argomento che sto discutendo deriva dall'errata assunzione secondo la quale poiché la nostra capacità di descrivere le cose nel nostro ambiente dipende dalla nostra consapevolezza di esse, le nostre descrizioni delle cose siano in primo luogo descrizioni della nostra esperienza cosciente e solo secondariamente, indirettamente e inferenzialmente, descrizioni degli oggetti e degli eventi che si trovano nel nostro ambiente. Si assume tacitamente quanto segue: poiché riconosciamo le cose del nostro ambiente dal loro aspetto, dal loro odore, dal gusto o dal suono, cominciamo con il descrivere le loro proprietà fenomeniche, cioè le proprietà del loro aspetto, dei suoni, degli odori, dei sapori, ecc. che esse producono in noi e inferiamo poi le loro proprietà reali dalle loro proprietà fenomeniche. Di fatto, accade l'inverso. Noi cominciamo dall'apprendere e riconoscere le proprietà reali degli oggetti del nostro ambiente. Impariamo a riconoscerle, ovviamente, da come ci appaiono, dai loro suoni, odori e sapori; ma questo non significa che noi dobbiamo imparare a descrivere l'aspetto, il suono, l'odore e il sapore delle cose prima che possiamo descrivere le cose stesse. In realtà è solo dopo che abbiamo imparato a descrivere gli oggetti nel loro ambiente che impariamo a descrivere la nostra consapevolezza di essi. Noi descriviamo la nostra esperienza conscia nei termini di mitologiche "proprietà fenomeniche" che si suppone ineriscano a mitologici "oggetti" in un mitologico "campo fenomenico", ma attraverso il riferimento alle effettive proprietà fisiche degli oggetti, degli eventi e dei processi fisici concreti che normalmente, benché forse non nella fattispecie in esame, fanno sorgere quel genere di esperienza conscia che cerchiamo di descrivere. In altre parole, quando descriviamo una post-immagine come verde, non stiamo dicendo che c'è *qualcosa*, la post-

immagine, che è verde; stiamo dicendo di avere quel genere di esperienza che normalmente abbiamo quando – e che abbiamo imparato a descrivere come quando – stiamo osservando una chiazza di luce verde.⁷

Prendiamo quindi le mosse dal seguente enunciato: «*Mi sembra ci sia una chiazza di luce di forma circolare e di colore giallo-arancione sul muro*».

Due sono, per il teorico dell'identità, le tesi da evitare. La prima è negare che esistano genuini enunciati qualitativi, cioè enunciati qualitativi dotati di senso (ho una post immagine giallo-arancio) e contemporaneamente affermare la sola esistenza di disposizioni a comportarsi in un certo modo, in questo caso la disposizione a dire che vi è una chiazza di luce di forma circolare e di colore giallo-arancio sul muro. La seconda è affermare che esistano enunciati qualitativi che si riferiscono genuinamente a oggetti propri, non oggetti fisici quindi ma fenomenici. Tali oggetti, insieme alle loro proprietà sono infatti, per il teorico dell'identità, "pendagli" o "ciondoli" nomologici (*nomological danglers*), correlazioni nomologiche sospese e lasciate inspiegate dal punto di vista fisico.⁸

La fallacia fenomenologica consisterebbe quindi nell'operare un passaggio ontologico indebito dalla nozione di stato mentale come *modalità di accesso fenomenica* alla nozione di stato mentale come *oggetto fenomenico*.

Se così interpretata, la fallacia fenomenologica introduce come nuovo e essenziale elemento la nozione di causa e, insieme ad essa, l'idea di un mondo trascendente nel senso metafisico del termine. Il fatto che si parli di *post-immagine* e non, semplicemente, di dati fenomenici direttamente esperiti, ne costituisce un chiaro indizio.

Per prevenire osservazioni irrilevanti, vorrei chiarire che con "lampo" intendo l'oggetto fisico pubblicamente osservabile, non un dato sensibile visivo. Dico che l'oggetto fisico

⁷ Ibid. pp. 70-71.

⁸ È da segnalare, a questo proposito, come un argomento praticamente identico alla fallacia fenomenologica sia rintracciabile in Berkeley per confortare la tesi opposta, cioè la tesi dell'*immaterialismo*. «può forse essere obiettato che, se l'estensione e la figura esistono solo nella mente, segue che la mente è estesa e ha una figura; poiché l'estensione è un modo, o un attributo, che, parlando nei termini degli scolastici, è predicato del soggetto nel quale essa esiste. Rispondo che queste qualità sono nella mente in quanto sono percepite da quest'ultima, cioè non nella forma del *modo* o dell'*attributo*, ma solamente in quello dell'*idea*; e non segue che l'anima o la mente sia estesa, perché l'estensione esiste in essa e basta, più di quanto non segua che essa sia rossa o blu, perché questi colori sono riconosciuti, sotto tutti i rispetti, esistere in essa e in nessun altro luogo. Riguardo a ciò che i filosofi dicono del soggetto e dei modi, questo sembra davvero infondato e inintelligibile. Per esempio, diranno che nella proposizione "un dado è duro, esteso e quadrato" la parola *dado* denota un soggetto o sostanza, distinta dalla durezza, dall'estensione e dalla figura che sono predicate di essa, e nella quale esse esistono. Questo non posso comprenderlo: a me sembra che un dado non sia distinto in niente da quelle cose che sono chiamate i suoi modi o accidenti. E dire che un dado è duro, esteso e quadrato, non è attribuire quelle qualità a un soggetto distinto da esse e che le sostiene, ma solamente una spiegazione del significato della parola *dado*». (*Trattato sui principi della conoscenza umana*, Milano, Bompiani, 2004, p. 333-5).

pubblicamente osservabile è di fatto una scarica elettrica, non un suo semplice correlato. Il dato sensibile o, piuttosto, l'aver il dato sensibile, la "visione" del lampo, può benissimo, secondo me, essere correlato alla scarica elettrica. Secondo il mio modo di vedere è uno stato cerebrale causato dal lampo. Ma non dovremo confondere la sensazione del lampo con il lampo più di quanto non confondiamo la sensazione di un tavolo con il tavolo.⁹

Se guardiamo la cosa dal punto di vista fenomenologico, due sono i risultati teorici che emergono dalle precedenti considerazioni, il primo dei quali verrebbe immediatamente sottoscritto dal fenomenologo. Si tratta dell'idea che esista una rigorosa distinzione fra *esperienza* di un oggetto e *oggetto* esperito. È l'esperienza a essere (eventualmente) identica a stati cerebrali. L'esperienza percettiva di un disco giallo-arancio non è giallo-arancio, esattamente come non lo è lo stato cerebrale. D'altro canto, che lo stato cerebrale non sia giallo arancio non esclude il fatto che esso sia identico all'esperienza percettiva di un disco giallo-arancio.

Esperienza percettiva e stato cerebrale, si potrebbe tradurre nella terminologia fenomenologica, sono entità *noetiche*; gli oggetti dell'esperienza percettiva e dello stato cerebrale sono entità *noematiche*.

Il secondo risultato è che per il teorico dell'identità le *proprietà* fenomeniche semplicemente non esistono. È questa, fra l'altro, la tesi che qualifica la fallacia fenomenologica: la fallacia consiste infatti nell'assumere che parlare di esperienze percettive significhi *anche* parlare di oggetti fenomenici verso i quali le esperienze percettive sarebbero dirette.

Per il teorico dell'identità, insomma, se è vero che esistono *processi* psichici irriducibili, è altrettanto vero che non esistono *proprietà* psichiche irriducibili.

Quando descriviamo una esperienza visiva di un disco come giallo-arancio, non stiamo dicendo che c'è *qualcosa* che è *effettivamente* giallo-arancio.

L'esperienza percettiva del vedere un'arancia viene infatti descritta dicendo che è simile all'esperienza che abbiamo quando vediamo *realmente* una arancia. È esattamente in questo che consiste il resoconto topico-neutrale delle esperienze qualitative: «quando qualcuno dice: "vedo una post-immagine giallo-arancio" sta dicendo qualcosa come "accade qualcosa che è simile a ciò che accade quando ho gli occhi aperti, sono desto e c'è un'arancia ben illuminata di fronte a me, cioè quando vedo realmente un'arancia"»; e

⁹ J. J. Smart, *Sensazioni e processi cerebrali*, cit. pp. 78-79.

ancora: «dire che qualcosa mi sembra verde è semplicemente dire che la mia esperienza è simile all'esperienza che ho quando vedo qualcosa che è realmente verde».¹⁰

La teoria dell'identità presuppone così un'ipotesi radicalmente realista: gli *oggetti* fenomenici e le loro proprietà semplicemente non esistono; esistono solo gli oggetti esterni che causano determinate esperienze le quali sono, a loro volta, processi cerebrali. Il significato di *essere realmente verde*, ad esempio, è fisico; il linguaggio delle proprietà private non denota autonomamente.

In questo senso, sembra non esservi differenza fra la teoria dell'identità e l'ipotesi sostenuta da un eliminativista come Churchland. Com'è noto, Paul Churchland ritiene che l'anti-riduzionismo in filosofia della mente (e, in particolare, il cosiddetto argomento della conoscenza sostenuto da Jackson e Nagel) cada in una sorta di *fallacia intensionalista*. La fallacia consiste nel derivare la natura di ciò che è conosciuto dal modo in cui è conosciuto. La *modalità* tramite la quale un oggetto è conosciuto non corrisponde necessariamente a una *proprietà* (nel senso ontologico) dell'oggetto. In altri termini, il fatto che le proprietà fenomeniche o qualitative siano conosciute soggettivamente non ha nulla a che vedere con le proprietà effettive delle sensazioni stesse. Quando Mary vede per la prima volta i colori, come nel famoso esempio di Jackson, ella non conosce un *oggetto* diverso ma ha una *conoscenza* diversa dello stesso oggetto.

Qui però, *l'oggetto* del conoscere è esattamente lo stesso, da entrambi i punti di vista, quello soggettivo e quello oggettivo, ed è qualcosa di essenzialmente fisico: la configurazione del vostro corpo e delle vostre membra. Non c'è qui nulla di soprannaturale, nulla che vada al di là dei limiti delle scienze fisiche (...) L'esistenza di un accesso epistemologico personale, di prima persona, a un qualche fenomeno non implica che il fenomeno in questione sia di natura non fisica.¹¹

Se adottassimo un punto di vista fenomenologico, sembra tuttavia non esistere un vero e proprio argomento che sia in grado di giustificare la radicale negazione del dualismo delle proprietà sostenuta dal teorico dell'identità. Una volta appurato che esistano modalità conoscitive alternative, cosa ci garantisce che esse convergano su un unico oggetto e non comportino, viceversa, una pluralità di oggetti di riferimento?

O, quantomeno, cosa ci garantisce che gli oggetti a cui ci riferiamo abbiano non solo proprietà fisiche ma anche proprietà fenomeniche?

¹⁰ Ibid. p. 86.

¹¹ P. M. Churchland, *The engine of reason, the seat of the soul*, Cambridge, The Mit Press, 1995; trad it. *Il motore della ragione, la sede dell'anima*, Milano, Il Saggiatore, 1998, p. 216.

5. Identità come denotazione

Il terzo ordine di problemi riguarda l'intendere l'identità fra stati mentali e stati cerebrali non come identità per *composizione* bensì come identità per *denotazione*.

Si tratta di un argomento che ripropone, e in modo ancora più radicale, la conflittualità fra livello epistemologico-soggettivo e livello ontologico-oggettuale.

L'analogia con il famoso esempio di Frege viene utilizzata dal teorico dell'identità per consolidare le due tesi che caratterizzano la sua posizione teorica: la non eliminazione degli stati mentali da un lato e la negazione del dualismo delle proprietà dall'altro.

Così come «vedo la stella della sera» non significa «vedo la stella del mattino», analogamente «processo cerebrale» non ha lo stesso significato di «avere un'esperienza visiva». Ma lo stato mentale è (contingentemente) uno stato cerebrale, analogamente alla stella della sera la quale è (contingentemente) la stella del mattino, cioè Venere. Il significato di un'espressione non è ciò che l'espressione denomina. Se l'identità fra stati mentali e stati cerebrali è contingente e non analitica allora deve essere possibile fornire definizioni estensive dei due lati dell'identità logicamente indipendenti. Nell'identità «Espero è Fosforo» infatti identifichiamo Espero come la prima stella che compare alla sera e Fosforo come l'ultima stella che scompare al mattino. In entrambi i casi ci stiamo riferendo allo stesso oggetto, cioè il pianeta Venere.

Qualcosa di simile dovrebbe verificarsi nel caso dell'identità psico-fisica. L'espressione «stato cerebrale» sta a significare «un processo fisico così e così che avviene in quell'oggetto molle e biancastro che si trova nel cranio». Analogamente l'espressione «stato mentale» sta a significare «un processo psichico in cui accade, ad esempio, di vedere una macchia giallo-arancione sul muro». Cosa dire, tuttavia, della denotazione? A che cosa si riferiscono le due espressioni?

Se affermiamo che le espressioni denotano proprietà fenomeniche, siamo allora costretti a riconoscere che gli stati mentali si riferiscono a proprietà loro proprie, cosa che il teorico dell'identità vuole evitare. Se d'altro canto sosteniamo che le espressioni denotano *proprietà* cerebrali, allora violiamo la condizione dell'indipendenza logica del mentale, appiattendo così il riduzionismo ontologico nell'eliminativismo e facendo collassare quel gioco di equilibrio fra anti-eliminativismo e anti-dualismo che caratterizza la teoria dell'identità.

6. Conclusione: «il» problema della filosofia della mente

L'analisi dell'identità psico-fisica mediante l'utilizzo di concetti come quello di composizione e di denotazione fa emergere un problema ancora più generale. Stati mentali e stati cerebrali sono infatti *noeticamente* identici, mentre lampo e scarica elettrica, acqua e H₂O, la stella della sera e la stella del mattino sono oggetti che eventualmente risultano *noematicamente* identici. I due livelli sembrano essere sistematicamente confusi dai teorici dell'identità. Il fatto è che gli stati cerebrali, benché osservabili, descrivibili, in una parola rappresentabili, non sono tuttavia *solo oggetti* di rappresentazione ma anche *la fonte* della rappresentazione stessa. La stessa cosa si verifica nel caso degli stati mentali: essi sono in qualche misura *oggetti* di ulteriori stati mentali e allo stesso tempo la loro origine. Detto in termini generali, lo statuto degli stati mentali e degli stati cerebrali è costitutivamente ambiguo.

È questa ambiguità che rende la domanda «gli stati mentali sono *effettivamente* stati cerebrali?» dal punto di vista epistemologico difficilmente affrontabile. Tale domanda sottintende infatti il passo da un livello che possiamo definire *ontico* (un livello pur sempre legato al nostro accesso cognitivo e intenzionale) a un livello che possiamo definire *ontologico*. Se affrontiamo la questione ontologica in termini di riduzione ci imbattiamo in una inevitabile chiusura cognitiva: per poter effettivamente dichiarare l'identità fra stati mentali e stati cerebrali occorrerebbe infatti che il cervello fosse in grado di esercitare esattamente ciò che non è in grado di esercitare; la possibilità, cioè, di auto-rappresentarsi come oggetto.

Credo quindi che a quella domanda né il riduzionismo né l'anti-riduzionismo, proprio in virtù del loro comune fondamento epistemologico, possano dare in linea di principio una risposta.

Roberta Lanfredini

LA TEORIA DELL'IDENTITÀ, PRECISAZIONI E RISPOSTE

di Marco Salucci

Prima di cominciare vorrei esprimere un doveroso ma sincero ringraziamento per l'attenzione con cui Lanfredini e Peruzzi hanno esaminato il tema oggetto del mio lavoro *La teoria dell'identità. Alle origini della filosofia della mente* (Firenze, Le Monnier 2005). Il livello di analisi, la puntualità e l'ampiezza con cui discutono dei vari argomenti sono rilevanti, tanto rilevanti da eccedere l'occasione che li ha sollecitati – una presentazione pubblica del libro - e da configurare, ritengo, un significativo momento del dibattito filosofico sulla mente nella letteratura specialistica in lingua italiana. Se dovessi mettere insieme un altro *reader* sulla filosofia della mente aggiungerei senz'altro i loro interventi, che hanno decisamente il carattere di contributi originali.

Lanfredini e Peruzzi non hanno in realtà dibattuto di una singola teoria, ma più spesso di un intero settore disciplinare – la filosofia della mente – e non di rado anche di questioni pertinenti a settori limitrofi – epistemologia e filosofia del linguaggio. Affrontare puntualmente le loro osservazioni è dunque in questa sede un compito impossibile. E ciò non solo per ovvi motivi di spazio ma anche perché, onestamente, devo ammettere che non ho una risposta per tutti i problemi che sono stati sollevati. La mia simpatia nei confronti della teoria dell'identità non significa adesione totale, molti problemi segnalati sono problemi reali.

L'ipotesi dell'identità psicofisica è stata proposta come un programma di ricerca, non come una teoria già articolata; è dunque fisiologico che alcuni problemi siano ancora aperti. Tali mi sembrano, per esempio, quelli segnalati da Peruzzi là dove si domanda a proposito dei processi mentali:

di che natura sono? Quali strutture manifestano? Quali sono le loro differenziate condizioni di possibilità? Il che significa in primo luogo capire quali sono i processi mentali (a partire da quelli più elementari della percezione e della cognizione) e capire quali sono le transizioni di stato sul piano neurochimico.

E ancora:

Il vantaggio di ogni operazione del genere [la riduzione] se ben congegnata, è l'unità del sapere (che nel caso in questione si declina come chiusura causale del mondo: non ci sono cause di eventi naturali (fisici) che non siano eventi naturali (fisici), altrimenti sarebbe violato il principio di conservazione dell'energia – e questa violazione comporta, a catena, tutta una

serie di inesplicabilità, anche laddove proprio non avremmo voluto (nel caso che intendessimo negare la riduzione). Ma non si deve essere ipocriti: lo scarto tra la cornice generale (quella del microlivello e del megalivello) e le isole di struttura (che portano al mesolivello) così identificate e gerarchizzate resta; si tratta di descrivere al meglio questa gerarchia di scarti, piuttosto che farla collassare in un corto circuito immediato.

Interpreto tali affermazioni nel modo più favorevole: come inviti alla prudenza. Sostenere l'identità psicofisica necessita di conoscenze ulteriori, anche empiriche: ma ciò mi sembra rientrare nello spirito della teoria dell'identità la quale non ha mai sostenuto che gli stati mentali fossero riconducibili a stati cerebrali ora e qui, allo stato attuale (si tratti del 1960 o del 2008) delle conoscenze, ma solo che è sensato ipotizzare che potrebbero esserlo un giorno. È il sapere qui e ora (si tratti del 1600 o del 2008) che gli stati mentali *non* sono stati fisici che è inaccettabile, come sostenuto dai dualisti del passato e dai funzionalisti del presente.

Vorrei insistere un momento sulla questione. Osserva ancora Peruzzi:

Place si proponeva nel 1956 di mostrare che 'la tesi secondo la quale la coscienza è un processo cerebrale non può essere respinta con argomentazioni logiche'. Come poteva pensare di riuscirci e più in generale come ci si può riuscire? Due casi: ci limitiamo ai motivi logici finora adottati, ma in tal caso non si può escludere che un altro motivo logico, questa volta efficace, venga scoperto domani.

Per capire l'affermazione di Place occorre tenere presente che quando scriveva *c'erano* motivi di natura logica. Gli autori di questi argomenti (Smart, in particolare, li enumera in modo meticoloso nel suo saggio tradotto nel libro che qui discutiamo) avevano nomi e cognomi, alcuni noti altri oggi dimenticati, fra quelli noti e non dimenticati: Platone e Cartesio. Riconosco però che Place avrebbe fatto bene ad aggiungere alla frase le parole "argomentazioni logiche finora note". La postilla "argomentazioni logiche finora note" sembra anzi oggi essere più necessaria che mai dal momento che la storia del dibattito sull'ipotesi identitista successiva alla prima formulazione è in realtà la storia di reiterati tentativi di mostrare a priori la sua insostenibilità – da quello di Kripke all'argomento della conoscenza - . Ma è anche la reiterazione dei tentativi di chiudere i conti con il problema mente corpo a tavolino che ha mantenuto aperto uno spazio per l'intervento di coloro che avrebbero desiderato qualche volta abbandonare il tavolino per il laboratorio.

Resta vero – come osserva ancora Peruzzi - che i sostenitori dell'identità psicofisica non hanno prodotto evidenze empiriche a sostegno del loro punto di vista, ma ciò dipende

dallo stato delle neuroscienze dell'epoca, non da quello della filosofia. Dunque il dibattito sull'identità psicofisica non avrebbe potuto essere diverso: non c'è stata e non c'è nessuna prova empirica della falsità dell'ipotesi identitista: se un dibattito c'è stato è stato necessariamente su argomenti a priori. O *questo* dibattito, o *questa* storia della teoria dell'identità, o nessun dibattito, nessuna storia. Si potrebbe affermare, anzi, che mentre è notevolmente cresciuta la ricerca di argomentazioni a priori contro l'identità, la teoria stessa pare essere rimasta sostanzialmente nella forma originaria, salvo le precisazioni necessarie per difenderla dalle obiezioni (questo spiega il sostanziale silenzio più che trentennale di Place il quale, dopo il suo intervento del 1956, ha ripreso la parola nel 1988)¹.

Il carattere che ho appena descritto della storia della teoria dell'identità è interpretato da Peruzzi come l'aver vissuto "in uno splendido isolamento dalla psicologia cognitiva". Sarei incline a temperare tale giudizio poiché scorgo il peso dell'eredità dell'ipotesi dell'identità psicofisica nel passaggio dal cognitivismo classico al nuovo cognitivismo esattamente nel punto in cui quest'ultimo si mostra molto più sensibile del primo ai risultati delle ricerche neurobiologiche. Il funzionalismo stesso si è in qualche autore coniugato con il riduzionismo per produrre forme più mature e articolate di entrambi (per restare in Italia vorrei ricordare il lavoro di S. Gozzano). D'altra parte, anche il cognitivismo ha i suoi limiti² e ripensare la tesi dell'identità ha aiutato e aiuta a identificarli. L'accusa di "vivere in uno splendido isolamento" (particolarmente grave per una teoria che affidi alla scienza il proprio destino) vale per i fondatori della teoria, quando la psicologia cognitiva ancora non c'era o quando, pur essendoci, era utilizzata dal funzionalismo in senso antiidentitista. Il contesto storico filosofico in cui è nata la teoria dell'identità è quello comportamentistico che precede - comunque come quadro concettuale e problematico anche se non sempre cronologicamente - il cognitivismo. Oggi la situazione mi pare mutata: né il fisicalismo né il cognitivismo si presentano più nelle forme "pure" del passato. Il sottotitolo del volume che discutiamo è *Alle origini della filosofia della mente*: che il lettore sappia che leggendo Place e Smart non ha a che fare con l'ultima moda in fatto di filosofia della mente.

In ciò che segue continuerò a raccogliere qualche nota sparsa e ondivaga su qualcuno degli argomenti sollevati. Molte questioni importanti resteranno dunque tagliate fuori. Oltre a quelli che ho detto sopra, c'è però un altro motivo – più serio, credo – per cui non

¹ U.T. Place, *Thirty Years on--Is Consciousness still a Brain Process?*, in "Australasian Journal of Philosophy", 66, 1988, pp. 208-219.

² Limiti che, fra l'altro, Peruzzi stesso ha contribuito a individuare.

discuterò di tutte le questioni che i miei interlocutori pongono: esso deriva dal punto di vista da cui considero complessivamente il problema mente-corpo. Non mi pare che questo sia stato tenuto in debito conto da Lanfredini e da Peruzzi (in realtà Peruzzi ne fa cenno, ma per scartarlo subito). Il risultato è che a volte *non* ci incontriamo su un terreno di discussione comune.

Vorrei dunque preliminarmente enunciare la mia posizione, perché credo che soltanto avendola presente sia possibile per il lettore orientarsi nella discussione, e anche capire perché ho scelto di lasciare fuori alcune questioni. Dunque: ritengo che uno stato mentale sia uno stato cerebrale vissuto in prima persona (tale affermazione mi porta in diretta collisione con Peruzzi, in particolare là dove manifesta la sua insofferenza per il dibattito sui *qualia* e sull'argomento della conoscenza). In altre parole: alcuni stati cerebrali sono particolari stati fisici che hanno la proprietà di apparire in un modo caratteristico a coloro, e solo a coloro, che li possiedono; tale modo caratteristico è ciò che definisce uno stato mentale (qualitativo). Ora, *l'apparire* in un certo modo fa parte dell'arredo ontologico del mondo – esattamente come *l'essere* in un certo modo - mentre *l'essere* apparenza *di* qualcosa condiziona l'indipendenza ontologica dell'apparenza stessa³. Con tale mossa credo sia possibile salvare (possibilità che, come vedremo subito, è decisamente contestata da Lanfredini) sia il realismo (gli stati mentali esistono realmente) che il riduzionismo ontologico (gli stati mentali non escono dall'ordine naturale delle cose). Le apparenze non sono necessariamente inganni; il fatto che il mare sia calmo in profondità non significa che non sia agitato in superficie (profondità e superficie hanno anche la loro brava efficacia causale ben diversa per una nave e un sottomarino). Apparenza e realtà possono essere due livelli o due modi di essere di uno stesso oggetto (il tedesco distingue *Scheinung* da *Erscheinung*)

Mi sembra che la strategia che ho appena delineato si possa considerare una forma di teoria dell'identità, o, comunque, che ne erediti lo spirito, se non la lettera originaria. Se la teoria dell'identità così come si è configurata nel passato non è in qualche modo l'antenata della posizione che ho appena espresso, allora non sostengo la teoria dell'identità. Ma se, invece, lo è allora anche la mia è una teoria dell'identità. Devo dire che i miei interlocutori intervengono sulla teoria dell'identità nella sua forma classica, quella che le è stata data dai suoi padri fondatori. Per tale aspetto la discussione si

³ Ho più volte espresso e motivato tale posizione - si veda per es.: *L'argomento della conoscenza*, in R. Lanfredini (a cura di), *Forma e contenuto*, Milano, LED 2002, pp. 43-44 - e non lo farò di nuovo: i miei interlocutori hanno il diritto di esprimere le loro opinioni tanto a lungo quanto desiderano poiché intervengono per la prima volta, per me è doverosa la brevità: la mia occasione l'ho già avuta.

configura come una discussione di tipo storico. Discussione più che legittima dal momento che il libro di cui parliamo ha l'intento di ricostruire la storia della teoria e di fornire al lettore i testi fondamentali. Discussione legittima, sì, ma meno interessante di quella che vedesse invece i partecipanti coinvolti nel difendere o nell'attaccare posizioni teoretiche: come storico non avrei nessun obbligo di aderire alle posizioni di cui ricostruisco la storia né, tanto meno, di difenderle. In realtà, però (fortunatamente), la discussione con i miei interlocutori è teoretica. Ed è perciò che lamento il fatto che ci si è concentrati troppo sulla lettera della teoria e poco sullo spirito.

Lo scopo del libro – come si sono chiaramente resi conto Lanfredini e Peruzzi - non è infatti soltanto storico: non mi ha interessato ricostruire filologicamente il pensiero di un autore ma saggiare la tenuta della sua posizione nel contesto del dibattito contemporaneo, anche se tale tenuta dovesse dipendere da qualche modifica (che non sia snaturalizzazione, evidentemente) delle sue tesi. Ma allora, invece che analizzare nei dettagli e rilevare le carenze di una teoria che, si rammenti, è ormai vecchia di mezzo secolo il problema (teorico) principale è: è possibile individuare un nocciolo caratteristico della teoria dell'identità che sopravvive ai molti elementi caduchi e ormai datati e che può ancora oggi costituire l'ispirazione di un programma di ricerca? La mia risposta è un, cauto, sì: qualunque ricerca biologicamente e naturalisticamente orientata, che tenga conto cioè degli elementi contingenti di cui e di come è fatto il mondo, noi stessi e il nostro sistema nervoso centrale è, a mio parere, erede della teoria dell'identità. Come spesso accade, ciò che questo significa appare più chiaramente sullo sfondo di ciò che esclude; e ciò che esclude sono posizioni come il dualismo ontologico e certo funzionalismo. Attualmente il volto del funzionalismo è cambiato radicalmente rispetto a quello della "prima" rivoluzione cognitiva: è degno di nota il fatto che alla nascita del nuovo cognitivismo non è stata estranea una linea di ricerca che ha ereditato alcune istanze della teoria dell'identità come, appunto, il richiamo alla costruzione di modelli della mente biologicamente plausibili.

Il lettore comprenderà con quale soddisfazione non possa esimersi dal segnalare, a questo punto, il recente libro di S. Gozzano⁴. Gozzano ha il merito di richiamare l'attenzione su quello che è il punto centrale a cui sembra arrivata la discussione oggi sul problema mente-corpo, quello dell'efficacia causale. Egli tenta una riedizione e una difesa della teoria dell'identità ("identificare il mentale con il fisico sembra essere l'unica opzione ancora aperta" p. 104). Benché su qualche particolare non sia d'accordo con lui, ci sono

⁴ S. Gozzano, *Pensieri materiali*, Torino, UTET 2007, recensito più avanti in questa stessa rivista.

nel suo libro molte cose che avrei voluto scrivere io; ma ce ne sono anche altre (suvvia, non siamo troppo autolesionisti!) che io stesso ho già avanzato in passato.⁵ Una in particolare, quella su cui tutto il libro di Gozzano è costruito: se vogliamo mantenere il principio dell'efficacia causale degli eventi mentali *dobbiamo* sostenere una qualche forma di teoria dell'identità. Questo mi sembra *il* (non *un*, ma *il*) punto cruciale del problema mente-corpo oggi⁶; un punto al quale i miei interlocutori non hanno dedicato, ritengo, attenzione sufficiente (un accenno, tuttavia, si trova in Peruzzi). Si può affrontare il problema mente corpo da molti punti di vista, ma il *mio* punto di vista è quello dell'efficacia causale. Peggio: ritengo che ogni altro punto di vista eluda la questione fondamentale. Se una qualche forma di teoria dell'identità è falsa, allora dobbiamo elaborare un nuovo modello di causalità – non solo mentale ma *anche* fisica –, il che, non c'è bisogno di dirlo, è impresa ardua e tale da coinvolgere non solo la filosofia della mente: l'intera scienza naturale ne sarebbe interessata. Qualcuno, e non farò nomi, sembra oggi disposto a rinunciare *in toto* all'immagine che le scienze naturali ci danno del mondo con la sola motivazione che la relazione mente-corpo è inspiegabile. Ma tale relazione è inspiegabile solo se la mente viene posta al di fuori del mondo fisico.

Ritorniamo per un momento alla questione dalla quale ho cominciato, quando ho richiamato la strategia che a mio parere permetterebbe di salvare il realismo e l'efficacia causale nei riguardi degli stati mentali. Secondo alcuni la principale difficoltà della teoria dell'identità starebbe proprio nell'impossibilità di mantenere la distanza dall'eliminativismo (bisogna ricordare però che per altri l'eliminativismo è desiderabile). L'esponente più autorevole di tale rilievo è, come è noto, J. Searle. Anche Lanfredini però avanza un'obiezione simile.

Scrivi infatti la mia interlocutrice:

la teoria dell'identità non nega gli stati mentali, ma fonda la propria specificità sul tentativo - quasi un gioco di equilibrio - di affermare l'esistenza di stati mentali da un lato e l'identità fra stati mentali e stati cerebrali dall'altro. Non esiste un qualche attrito, è lecito chiedersi, nel voler sostenere entrambe le cose?

⁵ In particolare sulla necessità di confrontarsi con il problema dell'efficacia causale degli stati mentali e della chiusura causale del mondo fisico cfr. *La teoria dell'identità e la sua giustificazione*, 2003, p. 75. Ma la causalità era una preoccupazione anche degli stessi fondatori della teoria dell'identità: cfr. lo stesso lavoro oggetto della presente discussione alle p. 31 e 32.

⁶ Come ieri: si tratta della forma che il problema seicentesco dell'interazione ha assunto in età contemporanea.

Se esiste attrito, ritengo, allora esiste in tutti i casi in cui si afferma che qualcosa è qualcos'altro, ed è quanto accade in moltissime strategie cognitive e conoscitive, a meno di non condannare la conoscenza a un elenco di enunciati irrelati (su questo mi pare che Peruzzi concordi con me quando parla di "condanna al monadismo"). La preoccupazione di Lanfredini è comprensibile, ma occorre tener presente che non sempre usiamo le affermazioni di identità in senso eliminativo ("quell'oasi che hai visto era soltanto un miraggio", il che significa "l'oasi non esiste") ma anche per definire la natura di qualcosa. Quando diciamo "quel tavolo è un aggregato di molecole" non vogliamo dire che il tavolo non esiste. Nella scienza è comune l'uso delle affermazioni di identità, molti esempi sono stati portati dagli autori di cui discutiamo ("il lampo è una scarica elettrica", "i geni sono molecole di DNA", "il calore è l'energia cinetica media" ecc.). Si può discutere quali siano le identificazioni eliminative e quali non lo siano, resta il fatto che non tutte le affermazioni d'identità sono eliminative. Coloro che hanno sostenuto la teoria dell'identità fra stati mentali e stati cerebrali hanno ritenuto di aver a che fare con il secondo tipo di identità, anche se possono essersi sbagliati nel discutere i casi particolari. Lanfredini trova, infatti, gli esempi di identità adottati in letteratura inadeguati, ma cosa dimostra con ciò? Solo che gli esempi adottati sono inadeguati! Ma questo non deve meravigliare poiché se non sappiamo come specificare l'identità fra stati mentali e stati cerebrali non possiamo neppure trovare esempi o metafore adeguate. Ci manca una più precisa definizione dei termini che devono essere identificati: per il buon motivo che la ricerca su entrambi i termini è ancora agli inizi. Una teoria del rapporto mente corpo in realtà non esiste! La teoria dell'identità è infatti soltanto un'ipotesi di ricerca. Su questo i suoi sostenitori sono stati chiarissimi.

Lanfredini insiste sulla difficoltà che la teoria dell'identità psicofisica ha nel mantenersi in equilibrio fra eliminativismo e dualismo:

se affermiamo [che] le espressioni [mentalistiche] denotano proprietà fenomeniche, siamo allora costretti a riconoscere che gli stati mentali si riferiscono a proprietà loro proprie, cosa che il teorico dell'identità vuole evitare. Se d'altro canto sosteniamo che le espressioni denotano *proprietà* cerebrali, allora violiamo la condizione dell'indipendenza logica del mentale, appiattendoci così il riduzionismo ontologico nell'eliminativismo e facendo collassare quel gioco di equilibrio fra anti-eliminativismo e anti-dualismo che caratterizza la teoria dell'identità.

In primo luogo devo ricordare che il problema dell'indipendenza delle definizioni dei due lati dell'identità è stato da tempo affrontato e risolto con la nozione di ruolo causale:

nozione sulla quale c'è un consenso più ampio di quello fornito dai seguaci della teoria dell'identità dal momento che è ampiamente utilizzato anche dai funzionalisti. Pertanto l'indipendenza logica del mentale non è minacciata. In secondo luogo riformulerei la frase "gli stati mentali si riferiscono a proprietà loro proprie", la quale già suggerisce che esistano stati mentali indipendenti, in "gli (alcuni) stati cerebrali hanno una proprietà fenomenica, cioè, quella di apparire in un certo modo a chi li possiede". Non vedo come questa proprietà degli stati mentali possa essere eliminata affermando che è una proprietà degli stati cerebrali e, neppure, come possa introdurre proprietà non fisiche: perché il manifestarsi di proprietà fisiche non può essere una proprietà fisica? Ho detto che la mia simpatia per la teoria dell'identità non mi rende cieco nei confronti dei suoi problemi, ecco adesso una prova di tale affermazione (apparirà però subito chiaro come non sia impossibile utilizzarla per affermare che *non* sostengo una teoria dell'identità). C'è una tesi di Lanfredini che sottoscrivo in pieno⁷

l'unica cosa che sembra possa essere detta è che fra eventi cerebrali e eventi mentali vi è una correlazione sistematica e costante, condizione che è tuttavia del tutto insufficiente per sancire l'identità. [...] La costante e sistematica correlazione è infatti una condizione minimale accettata anche dal dualista senza troppe difficoltà.

Questo è un punto che mi preme particolarmente: il modo con cui i teorici dell'identità hanno pensato di poter giustificare la loro posizione è stato troppo sbrigativo.

Scrive Peruzzi:

la prima considerazione è appunto che una riduzione [...] diventa possibile solo quando la teoria di microlivello è stata arricchita di tutti quei concetti che sono richiesti per una riduzione soddisfacente [...] Che cosa, nel caso di mente e corpo, ci sia a svolgere il ruolo della teoria degli insiemi non è affatto chiaro.

I difensori della teoria dell'identità si sono solitamente appellati alla giustificazione teorica o impegnati in un dibattito su questioni delicatissime come la riduzione o l'esistenza di leggi-ponte. Sono pessimista sulla praticabilità di tali strategie, riguardo al problema mente-corpo. Pertanto non affronterò per nulla il problema della riduzione. E la ragione è la seguente. Uno degli esiti più rilevanti che a parer mio ha avuto l'attuale dibattito sull'argomento della conoscenza sta nell'aver dato la giusta importanza a ciò che, d'altra

⁷ E non sarebbe possibile altrimenti avendola a lungo già argomentata in *La teoria dell'identità e la sua giustificazione*, op. cit.

parte, anche Place sapeva già: per quanto si osservino stati mentali non si possono in linea di principio osservare stati cerebrali e *viceversa*. Ciò che ci è dato osservare sono allora solo correlazioni fra le due serie di eventi. Il problema è se le correlazioni devono bastarci per sostenere una *qualunque* teoria relativa ai rapporti psicofisici (sia dell'identità o no). Lanfredini pensa di no e ricorda, giustamente, che i teorici dell'identità sono d'accordo con lei, e non con me, poiché le correlazioni sono compatibili con qualunque teoria della mente.

A mia volta, però, concordo con Lanfredini non solo sul fatto che non esistano prove diverse della correlazione, ma aggiungo che non *possono* neppure darsi a causa di un insuperabile scotoma, un limite cognitivo imposto dal punto di vista in prima persona⁸. Tuttavia mi sembra che si possa passare dalla correlazione all'identità rimanendo nell'ambito di una procedura non inconsueta nella normale pratica scientifica. Un esempio di giustificazione teorica di un'identità è quella fra massa gravitazionale e massa inerziale. Si può assumere l'equivalenza tra le due forme di massa nell'ambito della teoria della relatività, ma di tale equivalenza non c'è evidenza sperimentale⁹.

E ancora: dall'osservazione che due moli di idrogeno e una mole di ossigeno hanno le stesse proprietà dell'acqua possiamo inferire solo che le proprietà delle due moli di idrogeno e di ossigeno sono correlate a quelle dell'acqua, oppure anche che due moli di idrogeno e una di ossigeno sono la stessa cosa dell'acqua? Le ragioni che militano a favore del passaggio dalla constatazione di correlazioni sistematiche e costanti all'identità possono essere molteplici, le ho già esposte (anche) nel libro di cui discutiamo¹⁰ (Lanfredini non le menziona, non ritenendole evidentemente meritevoli di considerazione), ne ricorderò dunque soltanto alcune. Al contrario di altre, la teoria dell'identità si integra con l'intero *corpus* delle attuali conoscenze scientifiche; l'identità è la sola in grado di spiegare la causazione mentale; l'identità *spiega perché esistono correlazioni costanti* fra stati mentali e stati cerebrali; l'identità si suppone ontologicamente ed esplicativamente più economica e più semplice di altre; l'identità spiega perché i progetti di ricerca delle neuroscienze hanno successo.

Ma Lanfredini incalza:

d'altro canto, nemmeno il ricorso a un qualche tipo di teoria [come sostengono alcuni, per es. Place] ci permetterà di stipulare tale identità e questo per due ordini di motivi: gli stati cerebrali non sono entità teoriche e gli stati mentali non sono entità osservative.

⁸ Cfr. M. Salucci, *La teoria dell'identità e la sua giustificazione*, op. cit.

⁹ Cfr. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, MIT- J. Wiley and S. Boston- New York 1972, pp. 67-68.

¹⁰ Alle pp. 49-51.

Se qui si vuole utilizzare un argomento epistemologico allora ci si potrebbe spingere fino in fondo e sostenere che la difficoltà di distinguere entità teoriche da entità osservative mostra il contrario di ciò che si vorrebbe: gli stati cerebrali sono entità teoriche (purché non si consideri stato cerebrale solo il livello superficiale dell'attività biochimica ed elettrica – ma anche in tal caso...). Tuttavia, prima di avventurarmi su tale insidioso terreno sottoscriverei il punto di vista di J. Fodor secondo il quale la distinzione fra teorico e osservativo non ha valore ontologico – ovvero è una distinzione ontologicamente neutrale – ma piuttosto fa riferimento al tipo e alla catena inferenziale con i quali consideriamo qualcosa teorico e qualcos'altro osservativo¹¹. Infine, per concludere sulla questione particolare, non posso lasciarmi sfuggire l'occasione di mettere l'un contro l'altro i miei interlocutori. Nel paragrafo intitolato *Parentesina*, Peruzzi avanza una risposta robusta – alla quale rimando - a qualunque argomento di stile postneoeempirista secondo il quale il riferimento sarebbe determinabile solo all'interno di una cornice teorica.

In generale le osservazioni di Peruzzi hanno un respiro molto ampio: esse coinvolgono più di una volta e in misura massiccia questioni fondamentali della scienza, della logica e della spiegazione scientifica. Mi pare che la sola strategia praticabile per discutere il suo intervento sia quello di fare qualche considerazione di carattere complessivo, invece che affrontare le singole questioni.

L'ipotesi sostenuta dalla teoria dell'identità è: si può ipotizzare *oggi* che le neuroscienze in *futuro* possano mostrare come gli stati cerebrali danno origine a stati mentali? Come ho già cercato di mostrare, aderire a tale ipotesi è aderire ad una posizione che è, almeno, erede della teoria dell'identità. Il richiamo alla scienza futura ci conduce a un'obiezione che mi sembra stia dietro anche a molte altre questioni sollevate da Peruzzi. Mi pare che egli sostenga una qualche versione della tesi che Crane e Mellor hanno avanzato una decina di anni fa: il fisicalismo o è falso, se si appoggia alla scienza attuale, o è vuoto, se si appella alla scienza futura. A parer mio, la prima opzione ha di mira un falso bersaglio perché nessun fisicalista ha mai sostenuto che la scienza *attuale* sia in grado di spiegare i fenomeni mentali (Armstrong è molto esplicito al riguardo). Per quanto concerne la seconda opzione devo confessare che non riesco a comprenderla bene, perché non riesco a vedere che differenza ci sia tra affermare che “la scienza futura troverà la cura per il cancro” e “la scienza futura troverà la spiegazione dell'origine della coscienza”. O sono

¹¹ J. Fodor, *Materialism*, in D. Rosenthal, *Materialism and the Mind-Body Problem*, Indianapolis, Hackett 2000, p. 132.

entrambe accettabili (cioè sensate, non vuote, indipendentemente dal fatto che si riveleranno vere o false) o sono entrambe rigettabili. Se sono rigettabili allora lo era anche la frase “la medicina troverà una cura per la tubercolosi” detta da Aristotele, mentre invece Aristotele avrebbe avuto ragione a pronunciarla. Se si intende sostenere che è la natura stessa del mentale a fare la differenza, allora siamo chiaramente di fronte a una petizione di principio: si assume già senza averlo dimostrato che il mentale non possa essere oggetto di un'indagine scientifica.

Scrivo Peruzzi:

per restare alla fisica: ‘in linea di principio’, la teoria unificata delle interazioni fondamentali dovrebbe ugualmente essere garantita o fortemente avvalorata, ma nessuno ha ancora oggi scoperto come, e proprio per questo ci si lavora. Come si lavora cercando indizi sperimentali, si lavora anche a ridiscuterne i presupposti teorici. ‘In linea di fatto’, la fisica è qualcosa di molto meno uniforme di quanto farebbe comodo pensare. La fisica è la base della cosmologia odierna, ma questa a sua volta pone problemi alla fisica. In particolare, la cosmologia odierna configura stati della natura che sono non meno impegnativi, dal punto di vista ontologico, di quanto lo sia ammettere stati mentali non identificabili con stati cerebrali, solo che li configura in un quadro esplicativo unitario, mentre chi come il dualista ha voluto separare nella maniera più netta.

Ho citato questo passo perché mi pare significativo non solo di una tesi particolare ma del tono generale dell'intervento di Peruzzi. In molti luoghi del suo intervento mi sembra di intravedere un rimando implicito a una “morale”, la stessa enunciata da Amleto quando dichiara che esistono più cose fra cielo e terra di quanto possa immaginare la nostra filosofia. L'esistenza di livelli di complessità ancora da indagare, i risultati delle ricerche logiche, psicologiche e fisiche contemporanee denunciano in modo chiaro la semplicità, spesso la rozzezza, dei teorici dell'identità. Tanto è più difficile dare, per esempio, una definizione di oggetto fisico tanto più è malamente fondato il fisicalismo e qualunque teoria del rapporto mente-corpo che vi si appoggi. Se questo è effettivamente il pensiero di Peruzzi, lo condivido. Tuttavia, poiché il fisicalismo è un'opzione filosofica e non scientifica, esso non è tenuto a dare una definizione di oggetto fisico, la quale spetta alla scienza. Ora, il fisicalismo dei teorici dell'identità consiste appunto nell'assumere consapevolezza di tale limitazione della filosofia. La storia della scienza, dal meccanicismo seicentesco alla relatività e alla meccanica quantistica contemporanee, dimostra come il concetto di oggetto fisico è mutato ad opera della stessa scienza. Tutto ciò è noto, e Peruzzi non ha certo bisogno di apprenderlo da me. Ma vorrei sottolineare che l'idea che

la fisica non è data una volta per tutte è un punto di vista che è condiviso dai sostenitori dell'identità psicofisica. Armstrong, per esempio, ne parla esplicitamente. La teoria dell'identità è stata l'unica (comunque la prima) teoria della mente a voler passare la parola proprio alla scienza in tema di mente.

Fra le molte possibili concezioni del fisicalismo, quella che più si adatta alla teoria dell'identità è quella espressa dal motto para-protagoreo di W. Sellars "la scienza è la misura di tutte le cose, di quelle che sono in quanto sono, e di quelle che non sono in quanto non sono". Aggiungo: *compresa la possibilità di ridurre tutta la scienza a fisica* (questione sulla quale Peruzzi insiste in maniera critica). Un atteggiamento fisicalista deve almeno sostenere (mi limito a elencare i punti che possono avere rilevanza nella filosofia della mente): 1) fisico è ciò che ci dice sia la scienza fisica; 2) l'indagine empirica è necessaria per accertare come è fatto il mondo fisico; 3) se la mente è riducibile a stati fisici può essere deciso solo dall'indagine empirica; 4) nel caso sia possibile la riduzione spetta ancora all'indagine empirica stabilire di che tipo di riduzione si tratti (anche la scelta fra identità, emergenza, sopravvenienza o qualunque altro tipo di relazione dovrebbe essere affidata alla ricerca empirica). Queste quattro affermazioni costituiscono il nocciolo della teoria dell'identità. Inoltre danno significato all'affermazione – ampiamente criticata da Peruzzi – dei fondatori della teoria secondo la quale il compito del filosofo della mente si limita alla denuncia dell'infondatezza di argomenti che si pretendono conclusivi a priori.

Il punto 1 mi sembra importante in quanto stabilisce una differenza fondamentale fra il materialismo dei secoli passati (per esempio quello di Hobbes) e quello contemporaneo: il materialismo contemporaneo non ha una metafisica propria, poiché affida alla scienza il compito di pronunciarsi su quali sono i costituenti della realtà, compresa, ripeto, la sostenibilità dell'idea che i costituenti siano in ultima analisi fisici. Se una teoria fisica domani esigerà l'esistenza di puffi blu dentro i quark allora dovremo accettare l'esistenza di puffi blu dentro i quark. In tal caso non avremo abbandonato il fisicalismo ma avremo un fisicalismo dei puffi blu. Il punto 2 elimina qualunque ricerca effettuata a priori: non si possono ottenere informazioni sulla realtà senza il concorso dell'esperienza (secondo una delle tesi più caratterizzanti dei neopositivisti "il sintetico a priori non esiste"). Peruzzi ha ragione a ricordare quanto oggi – dopo Quine – sia insostenibile una netta distinzione fra a priori e a posteriori e di insistere sulla necessità di valutare le conseguenze che una simile revisione avrebbe su una qualunque forma di teoria dell'identità (argomento degno di una ricerca più approfondita e, di fatto, a quanto mi consta, fin'ora inesplorato). Ma siamo

certi che tale revisione avrebbe come conseguenza la ridefinizione del ruolo che la ricerca empirica (neurobiologica) potrebbe avere per la comprensione del rapporto mente-corpo? Il punto 3 è una banale conseguenza di come i padri fondatori Place e Smart hanno concepito la teoria dell'identità fra stati cerebrali e stati mentali: un'ipotesi empirica. Se, in generale, è legittimo fare ipotesi empiriche, allora lo è anche nel caso della natura del mentale. Questo non è banale: quando l'identità fu proposta secoli di dualismo avevano preteso di mostrare il contrario, cioè che la questione del mentale non è una faccenda passibile (anche) di indagine empirica. Perciò la teoria dell'identità può essere vera o falsa ma sicuramente non è né banale né vuota. E non lo è anche perché esistono teorie alternative, teorie che occhieggiano negli interventi di Lanfredini e Peruzzi¹². Vediamo la prima.

Scrive Lanfredini:

una volta appurato che esistano modalità conoscitive alternative, cosa ci garantisce che esse convergano su un unico oggetto e non comportino, viceversa, una pluralità di oggetti di riferimento?

Domanda legittima, ed è la stessa dalla quale partirebbe un dualista. Io vedo la questione dal punto di vista opposto: cosa garantisce che l'esistenza di più modalità conoscitive implichi la pluralità di oggetti? Scoprire identità, scoprire cioè che ciò che si credeva essere una pluralità di oggetti diversi è invece un unico oggetto, è uno dei compiti della conoscenza ("l'acqua è H₂O", "i geni sono molecole DNA", "Espero è Fosforo" ecc.), un compito spesso più impegnativo di quello contrario. È una ricca storia quella del pensiero che si è impegnato nel mostrare (almeno da F. Bacon) quanto sia ingannevole fidarsi, per esempio, della ricchezza del linguaggio per trarre conclusioni sulla ricchezza ontologica. È curiosa, pertanto, l'inversione di prospettiva che Lanfredini opera rispetto a una delle preoccupazioni principali della filosofia del linguaggio almeno della prima metà del secolo passato: dati due termini, cosa garantisce che esistano due riferimenti?¹³

La seconda teoria che occhieggia negli interventi dei miei interlocutori è l'idealismo, prospettiva che compare quando essi ricordano che, a rigore, il monismo, risultato dell'identità, può avere segno opposto: può essere materialista o idealista. Porre una questione del genere conferma quanto ho detto a proposito dell'ampiezza delle questioni che sono state poste. Scegliere fra idealismo e materialismo comporta scelte metafisiche,

¹² Che "occhieggino" non significa che siano sostenute dai miei interlocutori. Sono presenti però come termini di confronto.

¹³ L'obiezione più forte al dualismo è tuttavia, come ricordato, quella della chiusura causale del mondo fisico.

che vanno dunque al di là della filosofia della mente. Così come prendere posizione circa la pluralità del riferimento dei termini comporta scelte di filosofia del linguaggio. Così come prendere posizione circa il problema se gli enunciati sugli stati cerebrali coinvolgano termini teorici comporta scelte in epistemologia. Così come ... È vero, *tout se tiens* ma si rischia di approdare a una forma di olismo paralizzante. Due parole sull'idealismo, come sul dualismo, vorrei tuttavia dirle.

Perché, dunque, i teorici dell'identità hanno tanto polemizzato con i dualisti mentre non hanno speso quasi nessuna parola verso gli idealisti? Probabilmente hanno agito ragioni storiche: non bisogna trascurare il fatto che tutti gli idealisti classici tedeschi – per non parlare dei nostrani - hanno considerato la conoscenza empirica e la scienza naturale come forme inferiori di conoscenza rispetto alla filosofia. Veramente troppo per una teoria della mente che intende ispirarsi alla scienza!

La distinzione fra le due possibili forme di monismo, il materialismo e l'idealismo, può essere colta, per esempio, attraverso la questione della naturalizzazione del mentale. Infatti una volta identificati gli stati fisici con quelli mentali resta ancora da prendere una decisione circa il fatto se i fenomeni mentali siano fisici o se i fenomeni fisici siano mentali. Questa decisione non ha importanza (sarebbe, questa sì, triviale) se consistesse solo, per così dire, nel cambiare di segno ai fenomeni ("persino ai fisici, nei tempi recenti, la materia è divenuta, tra le loro mani, più sottile: essi sono giunti a materie imponderabili, come calore, luce, ecc."¹⁴): in realtà che i fenomeni fisici siano "attratti" da quelli mentali o viceversa comporta ridefinire radicalmente il compito, gli scopi e il valore della scienza (e conseguentemente anche della filosofia). Naturalizzare il mentale non è la stessa cosa che "mentalizzare" la natura ("ma alla materia vitale [...] manca [...] ogni altro essere determinato, onde possa essere collocata fra le cose materiali"¹⁵). Se il mentale è un fenomeno naturale, e nessuno ha ancora dimostrato che non lo sia, allora la scienza è l'ovvia candidata a spiegarlo: questo penso sia inaccettabile per qualsiasi idealista.

Specificare che non tutti gli stati cerebrali sono stati mentali è reso necessario dal fatto che vi sono processi cerebrali che non sono mentali, come quelli che governano certi processi fisiologici. Questo è banale, ma una conseguenza non banale è che l'identità psicofisica non è reversibile. Oltre a denunciare la particolarità dell'identità psicofisica rispetto ad altri tipi di identità, il fatto che l'identità non sia reversibile ha sia un significato ontologico, secondo il quale gli stati mentali dipendono da stati fisici (come per esempio sostenuto anche da alcune teorie della sopravvenienza o dell'emergenza del mentale sul

¹⁴ G.W.F. Hegel, *Enciclopedia delle scienze filosofiche*, Anmerkungen al par. 389.

¹⁵ Ibid.

fisico), sia un significato epistemologico. L'indagine dello stato fisico ha una priorità esplicativa su quella dello stato mentale e può gettare luce su alcuni tratti di quest'ultimo. Mentre una spiegazione dell'origine di stati mentali da stati fisici può almeno essere un'ipotesi per la ricerca empirica, quella contraria non si vede neppure come possa essere impostata. La priorità fondativa del fisico sul mentale sta alla base delle correnti teorie della sopravvenienza del mentale.

Concluderò queste brevi note dando prova di credere in quanto ho dichiarato all'inizio e cioè che molte questioni sollevate dai miei interlocutori sono reali. Ne ho ricordate tre che mi sembrano cruciali e dunque bisognose di una indagine più approfondita.

1. La questione del "solo correlazioni" sollevata da Lanfredini. Ritengo che la discussione sull'argomento della conoscenza sia stata così ampia per un motivo diverso da quello che i suoi animatori hanno creduto che fosse: il motivo vero è che l'argomento della conoscenza ci mette di fronte a un limite cognitivo. Come ho già detto c'è un limite cognitivo che impedisce di trovare prove empiriche a sostegno della teoria dell'identità. Questo, tuttavia, non è un argomento a priori contro la teoria: non è *contro* perché ritengo ci siano altre vie a favore (le correlazioni, appunto), non è *a priori* perché il limite cognitivo è un dato di fatto, radicato nella nostra costituzione psicofisica.

2. La questione, sollevata sia da Lanfredini che da Peruzzi, dell'arcaicità dei *paraphernalia* filosofici utilizzati dalla teoria dell'identità. Alcuni elementi dell'approccio neoempirista e analitico oggi non sono più sostenibili, fra questi la distinzione fra argomenti a priori e argomenti a posteriori, fra dati osservativi e dati teorici.

3. La necessità, sollecitata da Peruzzi, di rivedere la nozione di causalità. Poiché questo punto mi sembra, come ho già detto, quello più rilevante è anche quello a cui sono più sensibile. Non credo però che elaborare un nuovo modello di causalità sia un compito che possa e debba essere affrontato dalla filosofia della mente. Una revisione della nozione di causalità che fosse promossa dai filosofi della mente mi sembrerebbe *ad hoc*. C'è già chi, infatti, sembra disposto ad abbandonare il principio di chiusura causale per salvare una nozione – pregiudiziale e forse ancora più arcaica – di causazione fra mentale e fisico. L'abbandono del principio di chiusura causale avrebbe però un prezzo altissimo, dal momento che costituisce un pilastro delle scienze fisiche. In tal caso mi sembrerebbe corretto consultare le dirette interessate.

La questione della causalità mi dà lo spunto per ricordare una teoria che mi sembra la rivale più temibile dell'identità. Credo, ma potrei sbagliare, che Peruzzi pensi a questa quando, più volte, direttamente o indirettamente, accenna alla complessità del mondo

fisico. L'ipotesi circa il mentale *prima facie* più congrua con tale caratteristica del reale è quella dell'emergentismo o della sopravvenienza. Tale ipotesi, tuttavia, si scontra con il problema – per il quale ancora non c'è soluzione - della doppia causazione¹⁶. Il che, comunque, non fa che confermare la necessità di una riflessione sulla causalità.

Marco Salucci

¹⁶ O meglio, una soluzione è stata proposta: quella di non considerarlo più un problema! Sulla questione cfr. J. Kim, *Mind in a Physical World*, The MIT Press, Cambridge, 1998.

LA QUESTIONE DEL LIBERO ARBITRIO TRA COMPATIBILISMO E INCOMPATIBILISMO

di Chiara Erbosi

Il problema del libero arbitrio è da sempre stato oggetto di un fecondo dibattito filosofico che è arrivato fino all'epoca contemporanea. La problematica che ruota intorno alla domanda se gli uomini possono dirsi liberi, è estremamente complessa. Da un lato se si segue l'intuizione del senso comune, appare ovvio che l'uomo è in grado di controllare le scelte e le azioni che compie, che di esse è responsabile e che è dunque libero. Ma, approfondendo la questione, questa intuizione non appare più così scontata e anzi molti sono i problemi che sorgono intorno alla possibilità di affermare la libertà umana. Più si indaga su questa complessa problematica e più si infittiscono le difficoltà che riguardano la questione del libero arbitrio. In epoca moderna, per esempio, per molti, l'idea della libertà sembrava assolutamente inconciliabile con il cieco determinismo che si attribuiva alle leggi di natura. Gli uomini possono, e, in caso di risposta affermativa, come, sfuggire a tali leggi e dirsi liberi? In alternativa: è concepibile la libertà in un universo deterministico? Davanti a tali quesiti, l'intuizione secondo la quale l'uomo è libero non appare più così ovvia e scontata. Così se da una parte ci pare ovvia l'idea della nostra libertà, dall'altra le leggi del mondo naturale sembrano non lasciare alcuno spazio al libero arbitrio. Ma questo è solo un esempio della complessità del dibattito che in epoca contemporanea si è ampliato con numerose e variegate posizioni.

Esistono molti tipi di libertà: libertà politica, libertà religiosa, libertà di stampa. In tutti questi casi ci si riferisce alla libertà "da qualcosa": «libertà, ad esempio, dai vincoli che lo stato o le istituzioni politiche, religiose o economiche pongono agli individui; oppure libertà da costrizioni esterne»¹. Non si può infatti dire che un individuo agisca in maniera libera se costretto o impedito nelle sue azioni. In questo caso la libertà assume una connotazione negativa, dato che viene intesa come assenza di ostacoli o impedimenti².

¹ M. De Caro, *Il libero arbitrio. Un' introduzione*, Laterza, Bari, 2004, p. 5.

² Anche A. Schopenhauer dà una definizione della libertà in senso negativo sostenendo che essa è la mancanza di ostacoli e impedimenti. Nel suo volume *La libertà del volere umano* egli afferma che in senso fisico «uomini e animali si dicono liberi quando le loro azioni non sono ostacolati né da legami, né dal carcere, né da una paralisi insomma da nessun impedimento fisico, materiale, ma si svolgono conformemente alla loro volontà». A. Schopenhauer, *La libertà del volere umano*, Laterza, Bari, 1997, p. 44. Da sottolineare però che la libertà da ostacoli ed impedimenti fisici è in minima parte presa in considerazione da Schopenhauer perché oggetto del suo saggio è la libertà del volere. E infatti poche pagine più avanti chiarisce che cosa è oggetto della sua indagine e scrive: «In base al concetto empirico di libertà diciamo: "io sono libero se posso fare ciò che voglio" e con questo "ciò che voglio" la libertà è già definita. Ma dato che cerchiamo la libertà del volere, la domanda andrebbe formulata così: "puoi anche volere ciò che vuoi?"». Ivi, p. 46.

Negli ambiti poco sopra menzionati «il grado di libertà di cui un agente gode è inversamente proporzionale ai vincoli e alle costrizioni cui è sottoposto: così, per fare un esempio, la libertà politica aumenta con il diminuire dei vincoli che lo stato pone all'esercizio dei diritti dei cittadini»³. Quando si tratta però di libero arbitrio non è il grado di libertà che interessa la discussione quanto la sua stessa possibilità, cioè la capacità dell'uomo di autodeterminarsi e la responsabilità che l'essere umano deve portare per le proprie scelte e le proprie azioni. Non è un caso che i nessi che legano la libertà ai concetti di autodeterminazione e responsabilità siano fondamentali per il dibattito filosofico. Il dibattito sulla libertà assume dunque un'importanza fondamentale perché ha rilevanti implicazioni per varie discipline come l'etica, la filosofia del diritto, la filosofia politica: «A giudizio quasi universale, infatti, la libertà è da ritenersi *conditio sine qua non* di alcuni concetti fondamentali di queste discipline, come quelli di responsabilità e di autodeterminazione. In effetti, di norma gli agenti vengono considerati responsabili delle loro azioni (dal punto di vista legale, morale e politico) solo se essi le hanno compiute liberamente»⁴.

Nonostante dunque per molti autori, le disquisizioni filosofiche sul libero arbitrio siano una sorta di scandalo per quanto poco si sono rivelate proficue e conclusive⁵, il problema è di primaria importanza per i risvolti e le implicazioni che in tale dibattito sono coinvolte. Ora, due sembrano essere le condizioni della libertà. Perché si possa dire che un individuo ha compiuto un'azione libera,

1. è necessario che davanti all'agente si presentino vari corsi di azione alternativi
2. e che egli possa compiere scelte e azioni diverse da quelle che di fatto compirà.

La prima condizione della libertà è dunque la possibilità di agire diversamente rispetto al corso di azione scelto. Ma non è sufficiente che ad un soggetto si presenti una varietà di corsi di azioni possibili; è necessario che la scelta tra le varie possibilità alternative discenda dalla sua volontà autodeterminata e che non sia il frutto di una successione di eventi casuali. L'altra condizione essenziale per la libertà è dunque che l'individuo abbia il controllo sulle azioni che compie. Le scelte fatte non devono infatti risultare come frutto del caso o di fattori indipendenti dalla volontà della persona bensì devono essere «l'effetto di una catena causale in cui i desideri le credenze e le intenzioni dell'agente

³ M. De Caro, *Il libero arbitrio. Un' introduzione*, Laterza, Bari, 2004, p. 5-6.

⁴ Ivi, p. 7.

⁵ In questo senso si esprimono filosofi come John R. Searle che apre il suo libro *Libertà e neurobiologia. Riflessioni sul libero arbitrio, il linguaggio e il potere politico*, con queste parole: «la persistenza del problema del libero arbitrio in filosofia mi sembra costituisca una sorta di scandalo. Dopo secoli di riflessione sul libero arbitrio non mi pare siano stati fatti molti progressi ». John R. Searle, *Libertà e neurobiologia. Riflessioni sul libero arbitrio, il linguaggio e il potere politico*, Bruno Mondadori, Milano, 2005, p. 3.

giocano un ruolo determinante»⁶. In altri termini è necessario che gli individui abbiano il controllo sulle azioni che compiono e dunque che si possano auto-determinare è la seconda condizione della libertà.

Diverse sono le tesi che vengono elaborate sul libero arbitrio. Sicuramente una prima grande divisione deve essere fatta tra due gruppi di teorie: il compatibilismo e l'incompatibilismo. Il compatibilismo sostiene una definizione di libertà compatibile con il determinismo mentre l'incompatibilismo, presentandosi come la negazione del compatibilismo, nega la conciliabilità della libertà con il determinismo causale⁷. Sull'incompatibilismo si contrappongono poi diverse scuole di pensiero: ci sono gli incompatibilisti libertari i quali affermano «che la libertà è compatibile con l'indeterminismo causale, che nel mondo umano è presente una significativa quantità di indeterminismo e che gli essere umani godono della libertà»⁸. Vi sono poi gli incompatibilisti antilibertari che negano la libertà umana dal momento che sostengono che il mondo sia deterministico e che vi sia incompatibilità tra libertà e determinismo. Altri invece, gli scettici, credono che la libertà umana sia inconciliabile tanto con il determinismo quanto con l'indeterminismo e per tanto è impossibile.

Secondo il libertarismo la libertà è possibile solo ed esclusivamente in un mondo indeterministico. Le teorie libertarie si suddividono in tre gruppi:

Il primo gruppo sostiene una concezione che possiamo chiamare indeterminismo radicale, che si propone di spiegare la libertà congiungendo l'indeterminismo con una concezione non causale dell'azione. Il secondo gruppo difende invece l'indeterminismo causale, che coniuga l'indeterminismo con una concezione causale dell'azione. Il terzo gruppo infine, propugna una concezione comunemente detta *agent causation*, secondo la quale la libertà dipende da speciali poteri causali degli agenti incompatibili con il determinismo⁹.

Qualunque tipo di teoria si sostenga, il libertarismo incorre in enormi difficoltà difficilmente risolvibili e quindi sembra essere inadeguato a risolvere il problema della libertà.

La tesi fondamentale del libertarismo è che la libertà è possibile soltanto in un contesto indeterministico e che l'assenza di ogni tipo di determinazione è condizione fondamentale

⁶ M. De Caro, *Il libero arbitrio. Un' introduzione*, Laterza, Bari, 2004, p. 10.

⁷ Con determinismo si intende quella dottrina filosofica secondo la quale tutto ciò che accade è determinato in maniera casualmente necessaria da una catena ininterrotta di eventi avvenuti in precedenza unitamente alle leggi di natura valide in questo nostro mondo. L'indeterminismo sostiene invece la tesi opposta.

⁸ Ivi, p. 25.

⁹ Ivi, p. 27-28.

della libertà. In un mondo indeterministico infatti, niente è già stabilito e dunque si ha lo spazio perché gli agenti scelgano tra diversi corsi di azione tutti ugualmente possibili e si autodeterminino in base alle scelte compiute e da compiere. Solo in tal caso la decisione su quale corso di azione prendere tra i tanti possibili dipende solo ed esclusivamente dall'individuo. I libertari incentrano le loro teorie sulla difesa del poter fare altrimenti, sull'idea che l'uomo in quanto libero, possa intraprendere sempre una scelta o una azione diversa da quella che di fatto compirà. Ma se apparentemente il libertarismo sembra dare una risposta più che esaustiva al problema della libertà molte sono anche le difficoltà alle quali va incontro tra cui fondamentale è l'incapacità di dare conto di come gli agenti possono controllare le azioni che compiono. Il problema nasce dal fatto che in un contesto indeterministico la libertà sembra avvicinarsi molto all'idea di casualità. Una delle condizioni della libertà è che l'agente possa scegliere tra corsi di azioni alternativi e quindi il contesto indeterministico sembra perfetto perché questa condizione si realizzi dal momento che secondo questa teoria niente può prestabilire ciò che di fatto accadrà. In un mondo indeterministico le azioni tra le quali un individuo deve scegliere sono tutte egualmente probabili e lo spazio di possibilità è assolutamente aperto. Ma se è così, se davvero niente può determinare quale corso di azione si compirà, allora nemmeno l'agente può controllare le proprie scelte. Ma allora ciò che accadrà sembra più il frutto del caso piuttosto che il prodotto di una scelta libera e razionale. Dunque gli agenti non possono rendere conto delle scelte compiute perché queste sfuggono al loro controllo ed essi non sono in grado di determinare ciò che accade. Le scelte e le azioni fatte non sono il frutto di una deliberazione o di una volontà libera ma accadono per caso, per mera accidentalità. Un'azione però per poter essere definita libera non può essere casuale ma deve poter essere controllata dall'agente perché a lui deve essere imputabile e ciò può accadere solo se l'azione compiuta è frutto di una scelta libera e razionale e non del caso. Dunque il libertarismo, ben lontano dal dare conto della libertà, sembra renderla impossibile perché rende impossibile da parte degli agenti il controllo sulle azioni, vanificando anche così l'idea di responsabilità così strettamente legata a quella di libertà. Un agente infatti può essere responsabile di una azione che è stata il frutto di una deliberazione razionale e non di una azione che è accaduta per caso. Né l'indeterminismo causale né l'indeterminismo radicale riescono ad eludere tale problematica che resta dunque irrisolta. La terza forma di libertarismo, quella dell'*agent causation*, non nega un nesso causale tra la decisione dell'azione e l'azione che viene compiuta. Nega però che tale rapporto sia analizzabile nei termini del determinismo della fisica sognata da Laplace

perché attribuisce una causalità agli stati mentali, i quali non sono né determinati secondo modalità di causazione studiate dalla fisica, né causano secondo tali modalità.

L'indeterminismo radicale fonda le proprie argomentazioni su una concezione non causale dell'azione. Secondo tale teoria ci sono dei nessi indeterministici ma non causali: le azioni cioè sono eventi senza cause. Se gli eventi non hanno cause, non possono essere necessitati: fino al momento in cui una azione viene compiuta è in potere dell'agente fare una scelta diversa impostando un corso di azione alternativo. Come detto prima, questa teoria non è esente dal problema del controllo in quanto non riesce a spiegare come l'agente sia in grado di controllare le proprie azioni. Se un evento infatti è indeterminato e niente può controllarne l'accadere, allora nemmeno l'agente può dare conto delle azioni compiute e delle scelte fatte. Se infatti un evento non ha causa esso è indipendente dall'agente e da tutti gli eventi che lo riguardano («sia quelli che occorrono a lui quando agisce sia quelli che nel tempo hanno forgiato il carattere e il modo di comportarsi»¹⁰). Un tale evento dunque accade per caso e a parità di condizioni poteva anche non accadere:

La nozione di controllo, al contrario, richiede che l'agente scelga di compiere una specifica azione in forza dei suoi desideri, delle sue intenzioni, delle sue credenze; in una parola, richiede che egli scelga ed agisca nel modo in cui sceglie ed agisce perché (un perché che ha evidente carattere causale) quell'agente è la persona che è. Le azioni, considerate alla luce di questa concezione, si presentano come eventi casuali, incontrollabili: la rinuncia al nesso causale tra agente e azione sembra dunque compromettere non solo la possibilità dell'azione libera ma dell'azione tout court¹¹.

La seconda concezione libertaria è l'indeterminismo causale. Tale teoria si fonda sulla nozione di causazione indeterministica. Secondo tale tesi, lungo la catena causale che porta alla realizzazione dell'azione, interviene, tra il processo che porta alla deliberazione e il compimento dell'azione, un elemento indeterministico. In questo modo, secondo l'indeterminismo causale, è garantita la prima condizione della libertà, cioè la possibilità di agire diversamente: infatti, dal momento che le azioni non sono il frutto di una causazione deterministica potrebbero anche non accadere o verificarsi in maniera diversa da come di fatto accadono pur in presenza dello stesso passato e delle stesse leggi di natura. In questo modo De Caro spiega il fulcro del ragionamento dell'indeterminismo causale:

Per potere affermare che un agente causa una certa azione è sufficiente che egli intenzionalmente si adoperi in maniera adeguata per compiere quell'azione, anche se il

¹⁰ Ivi, p. 43.

¹¹ Ivi, p. 44.

risultato dei suoi sforzi in tal senso non è determinato o inevitabile. In una parola, l'idea è quella di una forma di causalità indeterministica che non necessiti l'effetto ma semplicemente aumenti la probabilità che esso accada: in questo modo l'effetto non è determinato dalla propria causa¹².

Quando dunque un agente dovrà compiere una scelta, avrà davanti a sé diverse motivazioni e diverse ragioni da valutare per decidere quale corso di azione intraprendere. Le valutazioni che il soggetto compirà non predeterminano il corso dell'azione che l'agente sceglierà ma sarà solo alla comparazione tra le diverse ragioni che le varie valutazioni assumeranno un peso relativo facendo propendere l'agente verso una scelta piuttosto che verso un'altra. Dunque le ragioni spingono l'agente a compiere certe scelte ma non le necessitano. Secondo la tesi dell'indeterminismo causale, i soggetti scelgono tra vari corsi di azioni alternativi, dopo aver privilegiato, senza essere però determinati in questa scelta, un insieme di motivazioni piuttosto che un altro:

Quando un'agente deve optare tra due corsi di azione alternativi, egli sceglie comparando le ragioni che militano rispettivamente in favore dell'uno e dell'altro; tuttavia, ed è questo il punto fondamentale, prima di questa valutazione le diverse ragioni (desideri, credenze, intenzioni) non hanno un peso oggettivo, ovvero non determinano già, intrinsecamente, quale sarà l'opzione che l'agente sceglierà. Piuttosto, è proprio al momento della pesatura, della comparazione delle diverse ragioni che l'agente svolge al momento della decisione, che tali ragioni assumono il peso relativo che fa propendere l'agente in favore di una azione o dell'altra¹³.

Il processo che porta alla scelta di una azione invece di un'altra è contemporaneamente indeterministico e causale. È indeterministico perché «il peso relativo dei diversi insiemi di ragioni per agire non è predeterminato»¹⁴. E' solo dell'agente il potere su un corso di azioni o un altro perché essi non sono già predeterminati. È causale invece perché la valutazione che viene compiuta in favore di una scelta piuttosto che di un'altra ha come effetto di produrre l'azione scelta. L'indeterminismo causale cerca di garantire la prima condizione della libertà, cioè la possibilità di agire altrimenti, tutelando lo stretto rapporto che intercorre e che lega casualmente «le ragioni, le valutazioni e le deliberazioni degli agenti alle azioni che ne discendono»¹⁵. Quindi, se è vero che gli indeterministi causali ritengono che il processo che porta al compimento dell'azione non sia deterministico, è

¹² Ivi, p. 45.

¹³ M. De Caro (a cura di), *La logica della libertà*, Meltemi, Roma, 2002, p. 23.

¹⁴ Ivi, p. 47.

¹⁵ *Ibidem*.

vero anche che non tutto è lasciato al caos e alla accidentalità, come accadeva nell'indeterminismo radicale. Nonostante questo, però, neanche l'indeterminismo causale riesce a dare spiegazione di come gli agenti controllino i corsi di azioni che scelgono. Quando un atto è indeterminato, niente lo determina nemmeno l'agente. Se anche l'agente causasse in maniera indeterministica le proprie azioni, non ne seguirebbe per questo che le controlla, cioè che sia capace di determinare quale corso di azione si svilupperà:

In sostanza, in tale scenario il contributo causale indeterministico dell'agente al compimento dell'azione non pare sufficiente a impedire che anche contro questa concezione venga elevata, sia pure riformulata la classica accusa antilibertaria: quella secondo la quale l'indeterminismo impedisce il controllo dell'agente sulle proprie azioni¹⁶.

Sia l'indeterminismo causale che l'indeterminismo radicale non riescono dunque a sfuggire alla critica di non riuscire a dare spiegazione del controllo che un agente deve avere sulle proprie azioni se queste devono dirsi libere. Secondo alcuni autori, a questa critica non è soggetta la terza ipotesi di libertarismo, l'*agent causation*. Secondo questa teoria all'agente viene riconosciuta la capacità di autodeterminare la propria volontà dando così inizio a nuove catene causali. Ciò permetterebbe di spiegare come gli agenti controllino le proprie azioni. Questa concezione si basa su due tesi:

- la prima è che la libertà non riguarda solo le azioni ma soprattutto la volontà del soggetto;
- la seconda è che la libertà ha bisogno di una particolare causalità che non è la causalità degli eventi descritti dalle scienze della natura (cioè intesi come distinti dalle azioni).

¹⁶ Ivi, p. 48. Sempre De Caro, nel suo libro *La logica della libertà*, si esprime così su questo argomento: «Ammettiamo dunque che, nel valutare le ragioni che militano in favore di due azioni alternative, un'agente ricorra effettivamente all'operazione indeterministica di pesatura [...]. Immaginiamo poi che, in virtù di ciò, egli decida di leggere un libro: dato il carattere indeterministico del processo, tuttavia, l'agente avrebbe potuto pesare diversamente le proprie ragioni e optare, di conseguenza, per una diversa alternativa: per esempio, avrebbe potuto scegliere di scrivere una lettera. A questo punto dobbiamo chiederci che cosa ha fatto sì che l'agente scegliesse di fatto la prima opzione (si noti che, esattamente nella stessa situazione, egli avrebbe potuto scegliere diversamente: se così non fosse, infatti, la sua scelta sarebbe stata determinata). Di conseguenza, dato che, in quello specifico momento, dagli stati interni dell'agente, dalle sue credenze, dai suoi desideri, sarebbero potute discendere scelte ed azioni diverse, non si può dire che tali stati abbiano fatto la differenza, ovvero che sia dipeso da essi che l'agente abbia scelto di leggere il libro invece di scrivere la lettera. Infatti, se un atto è indeterminato, per definizione nulla lo determina: *a fortiori* nemmeno l'agente con i suoi stati interni. Dunque, se anche fosse vero che l'agente causa indeterministicamente le proprie azioni, non ne seguirebbe che egli le controlla, ovvero che è in grado di determinare quale tra i possibili futuri si attualizza. In tal modo nemmeno l'indeterminismo causale sembra in grado di dare conto di un requisito fondamentale della libertà: quello dell'autocontrollo». M. De Caro (a cura di), *La logica della libertà*, Meltemi, Roma, 2002, p. 24-25.

Quando infatti entrano in gioco le decisioni e le scelte di soggetti razionali non si può ricorrere alla causazione che utilizziamo per gli eventi perché gli agenti possono dare origine a nuove catene causali e «possono fare ciò perché svolgono il loro ruolo causale in quanto sono sostanze»¹⁷. Mentre gli eventi sono causati, le sostanze non possono esserlo, bensì possono dare corso a nuove catene di eventi. Gli agenti, determinando liberamente la propria volontà, scelgono quali corsi di azioni compiere. La libertà degli agenti si rivolge principalmente allo loro stessa volontà determinandola e solo in conseguenza di ciò riguarda le azioni scelte dai soggetti. Tale teoria sembra dunque dare garanzia di entrambe le condizioni della libertà: dal momento che gli agenti determinano la propria volontà e in secondo luogo le proprie azioni, tale teoria è in grado di garantire un relativo controllo intenzionale da parte degli agenti sui corsi d'azione scelti; negando poi il determinismo lascia spazio anche alla possibilità di agire altrimenti. Da questa teoria sia l'autodeterminazione che la capacità di agire diversamente escono salve e quindi la libertà sembra spiegata. Secondo i critici di tale teoria però, l'idea di una causalità da parte degli agenti è qualcosa che manca di qualsiasi sostegno empirico, una misteriosa eccezione all'unica forma di causazione giustificabile che è quella tra eventi. Secondo i sostenitori dell'*agent causation* quando si analizzano le azioni indagando nelle cause si arriva in un punto in cui non si può più dire che un evento ne ha causato un altro, ma solo che un agente ha iniziato una nuova catena causale. Non esiste solo la causazione deterministica tra eventi ma anche una causazione degli agenti che lascia spazio alla libertà. Come detto sopra però, questi poteri causali negli agenti appaiono ai critici dell'*agent causation* come «un'eccezione all'ordine naturale, che è invece scandito dalle leggi della causazione tra eventi»¹⁸. Anche se dunque, tra le teorie del libertarismo, quella dell'*agent causation* sembra la più idonea a dare spiegazione e giustificazione della libertà, anch'essa incorre in difficoltà che dai suoi sostenitori non sono ancora state superate¹⁹. Le altre due tesi del libertarismo, quella dell'indeterminismo radicale e quella dell'indeterminismo causale, sono oggetto di critiche ancora più severe perché incapaci di spiegare come gli agenti possano controllare le azioni che compiono. Dunque nessuna delle teorie libertarie sembra dare una giustificazione convincente ed esaustiva della libertà.

Ma anche il compatibilismo non è esposto a critiche minori. Il compatibilismo fonda le proprie teorie su tesi esattamente opposte alle tesi libertarie e cioè sull'idea che

¹⁷ Ivi, p. 50.

¹⁸ Ivi, p. 54.

¹⁹ Si potrebbe però argomentare contro i critici dell'*agent causation* che, se tale tesi sembra ricorrere ad un'eccezione all'ordine naturale, lo fa però per spiegare un fenomeno che si presenta come un'eccezione rispetto all'explanandum proprio delle scienze naturali: dunque si potrebbe dire che, se la spiegazione è singolare, lo è perché singolare è il fenomeno stesso che si chiede di spiegare.

determinismo e libertà siano compatibili e che il determinismo quindi non sia affatto un ostacolo alla libertà. Il compatibilismo cerca di dare spiegazione della libertà nel contesto della visione scientifica del mondo accreditata nel momento del suo fiorire, tentando di provare che il determinismo non impedisce la libertà, come invece sostengono gli incompatibilisti secondo i quali o gli esseri umani non sono liberi in quanto determinati, oppure, ed è la tesi libertaria, in quanto liberi, gli esseri umani non possono essere determinati. Le critiche contro il compatibilismo non sono certo una novità dell'epoca contemporanea: esse hanno infatti una storia secolare e ad esempio può essere portato il feroce attacco che il vescovo John Bramhall rivolse a Hobbes. Tutta la critica di Bramhall ruotava intorno al rigido determinismo che Hobbes applicava ad ogni evento sia fisico che psicologico. In un mondo governato unicamente dalle ferree leggi di natura ogni evento è determinato. Non fa eccezione la volontà umana che può essere pensata solo come un effetto e non come una causa in grado di autodeterminarsi. Così Hobbes definisce la libertà:

La libertà è l'assenza di tutti gli impedimenti all'azione che non siano contenuti nella natura e nella qualità intrinseca dell'agente. Così, ad esempio, si dice che l'acqua discende liberamente o che ha la libertà di discendere per il letto del fiume, poiché non c'è impedimento lungo quella direzione, ma non di traverso, perché gli argini sono impedimenti. E benché l'acqua non possa risalire, tuttavia non si dice che le manca la libertà di risalire, bensì la facoltà o il potere, perché l'impedimento è nella natura dell'acqua, e intrinseco. Allo stesso modo diciamo che colui che è legato non ha libertà di andarsene, poiché l'impedimento non è in lui, ma nei suoi legami: mentre non diciamo lo stesso di colui che è malato o storpio, perché l'impedimento sta in lui stesso²⁰.

Per Bramhall la definizione che dà Hobbes di libertà è semplicemente assurda in quanto la compatibilità di determinismo e libertà implicherebbe la totale mancanza di controllo sulle azioni da parte dell'agente e l'impossibilità di fare altrimenti. Per Hobbes invece la libertà è l'assenza di tutti gli impedimenti esterni all'azione, ma egli ritiene che la volontà sia determinata internamente e che non abbia senso sostenere che un agente può compiere un'azione o astenersene, quando siano presenti tutti gli elementi necessari al compimento dell'azione stessa. Per Hobbes anche quando le azioni sono libere e volontarie, sono necessitate. In *Libertà e Necessità* Hobbes scrive infatti: «[...] ritengo che la definizione ordinaria di agente libero, cioè che un agente libero è quello che, essendo presenti tutti gli elementi necessari alla produzione dell'effetto, può nondimeno non produrlo, implichi

²⁰ T. Hobbes, *Libertà e necessità*, Bompiani, Milano, 2000, p. 111.

una contraddizione e sia priva di senso. Come se si dicesse che la causa potrebbe essere sufficiente, ovvero necessaria, e tuttavia l'effetto non seguirà»²¹. Dunque un agente, quando sono presenti tutti gli elementi perché un'azione si compia, non può evitare di produrla perché non esiste la libertà da necessità. Tutto questo per Bramhall porta a concepire solo una pseudo-libertà che egli identificava, schernendo Hobbes, con la libertà di un bastone nelle mani di un uomo. I sostenitori del compatibilismo sono dunque sempre stati oggetto di critiche feroci per la loro spiegazione della libertà, o meglio perché si contestava che quella da essi definita "azione libera" davvero avesse a che fare con la libertà tradizionalmente intesa.

Infatti il compatibilismo si propone di dimostrare che, a differenza di quello che può sembrare apparentemente, determinismo causale e libertà non sono affatto incompatibili. La definizione di libertà che danno i compatibilisti è quella secondo cui la libertà equivale all'assenza di impedimenti o di ostacoli all'azione. In questa teoria, l'opposto della libertà non è, come vorrebbero i libertari, la necessità connessa al determinismo, ma la costrizione. È dunque libero colui che non è impedito nell'agire (perché non viene ostacolato in ciò che vuole fare) né vi è costretto (perché non è obbligato a compiere un'azione che non vorrebbe fare). Così ad esempio Hobbes definisce un agente libero: «un agente libero è colui che può fare se vuole ed evitare se vuole»²². Per il compatibilismo essere liberi significa compiere l'azione che si vuole fare senza nessun tipo di ostacolo: «se io voglio un bicchiere d'acqua e nulla mi impedisce di prenderlo, allora la mia azione di prendere un bicchiere d'acqua è un'azione libera»²³. Ciò che è indispensabile per dirsi liberi è che le azioni dei soggetti discendano dalla loro volontà, senza impedimenti o costrizioni. Un'azione è libera quando è determinata da una volontà che non è né impedita né costretta. La volontà però, è completamente determinata «da fattori come le esperienze passate della gente, l'istruzione che ha ricevuto, l'ambiente circostante o ancora dal suo assetto biologico oppure dall'insieme delle variabili fisiche in gioco»²⁴. Non si ha dunque nessuna rottura nella catena deterministica delle cause e degli effetti: «la volontà dell'agente è determinata da cause su cui egli non può agire ed essa a sua volta determina, causandole, le azioni che l'agente compie; nondimeno tali azioni, in quanto discendono dalla sua volontà non costretta né impedita, sono libere»²⁵. Secondo i compatibilisti dunque il determinismo è ben lontano dall'impedire la libertà. Non solo,

²¹ Ivi, p. 113.

²² Ivi, p. 115.

²³ M. De Caro, *Il libero arbitrio. Un' introduzione*, Laterza, Bari, 2004, p. 61.

²⁴ *Ibidem*.

²⁵ *Ibidem*.

alcuni di loro si spingono ad affermare che il determinismo è anche condizione necessaria della libertà, in quanto l'indeterminismo annullerebbe la libertà, facendola coincidere con il caso. Dunque per i supercompatibilisti perché un'azione possa dirsi libera è necessario che essa sia causalmente determinata, senza ostacoli ed impedimenti, dalla volontà dell'agente. Quindi un soggetto è libero quando compie le azioni che vuole e che desidera compiere. D'altra parte però, la volontà dell'agente è interamente determinata da fattori ambientali, dalla nostra educazione o dalle leggi della fisica e della biologia. In questo contesto un'agente può volere soltanto ciò che vuole perché la sua volontà è completamente determinata da fattori che sfuggono al suo controllo. La libertà come viene intesa da i compatibilisti può dunque solo riferirsi alle azioni e non alla volontà e agli stati mentali dei soggetti.

Già alcuni compatibilisti avevano individuato questo punto spinoso della loro teoria arrivando a sostenere, per potere controbattere, che l'unica libertà possibile è quella relativa all'azione, mentre la presunta libertà della volontà è un'illusione. Scrive ad esempio Voltaire citando Locke: «al saggio Locke una volontà libera [...] sembra una chimera. Egli conosce una sola libertà: il potere di fare quel che si vuole. Il gottoso non ha la libertà di camminare, il prigioniero quella di uscire a passeggio: l'uno è libero quando è guarito, l'altro quando gli viene aperta la porta della prigione»²⁶. Per gli avversari del compatibilismo però è assurdo pensare che la libertà non si applichi alla volontà dell'agente. La libertà vuole che la volontà si autodetermini e non che sia eterodeterminata da condizioni ed eventi esterni. Ai critici del compatibilismo, questa concezione sembra presentare una prospettiva dove tutto accade automaticamente e meccanicamente: e ciò riguarda anche gli agenti dal momento che la loro volontà risulta interamente determinata da fattori che sfuggono al loro controllo. La vera libertà richiede anche la libertà di volere ed è tale libertà che i compatibilisti negano. Perché infatti la libertà sia libera deve poter sfuggire alla catena delle cause deterministiche che non lasciano spazio, secondo i critici del compatibilismo, alla libertà di scelta. Come precedentemente visto, per gli incompatibilisti, la libertà è inconciliabile con il determinismo tradizionalmente, a ragione o a torto, alla fisica classica. Ma per i compatibilisti una volontà che si autodetermina sfugge alle leggi di natura e diventa una misteriosa eccezione all'ordine naturale del mondo messo in luce dalla scienza²⁷. Contro questa idea così si pronuncia Voltaire nel suo libro *Il filosofo ignorante*: «in realtà, sarebbe

²⁶ Voltaire, *La metafisica di Newton*, in *Scritti filosofici*, Laterza, Bari, 1972, p. 215.

²⁷ A questo proposito A. Schopenhauer afferma: «l'uomo come tutti gli oggetti dell'esperienza, è un fenomeno nel tempo e nello spazio, e siccome la legge di causalità vale per tutti a priori e pertanto senza eccezioni, anche lui vi deve essere sottomesso». A. Schopenhauer, Laterza, Bari, 1997, p. 89.

ben strano che tutta la natura, tutti gli astri obbedissero a delle leggi eterne, e che vi fosse un piccolo animale alto cinque piedi che, a dispetto di queste leggi, potesse agire come gli piace solo in funzione del suo capriccio»²⁸. Secondo i compatibilisti, libertà e determinismo sono conciliabili e il fatto che le nostre azioni siano il frutto di una volontà determinata non deve essere assolutamente confuso con situazioni come la coercizione di un soggetto costretto ad agire in una certa maniera diversa da quella che vorrebbe. È questa secondo i compatibilisti la confusione che fanno i loro critici nel rivolgere loro l'accusa che il determinismo sia l'opposto della libertà.

Coercizione e costrizione comportano effettivamente la negazione della libertà e su questo punto concordano sia compatibilisti che incompatibilisti: in questi casi è giusto sostenere che le azioni non discendono veramente dalla volontà dell'agente, ma queste tipologie di situazioni niente hanno a che fare con la determinazione della volontà. È così dunque che i compatibilisti rispondono all'errore che a loro giudizio compiono i libertari affermando che libertà e determinismo sono inconciliabili. Ma altre sono le accuse che, soprattutto in epoca contemporanea, sono state rivolte al compatibilismo. Come precedentemente visto, due sono le condizioni della libertà: la prima è che perché un'azione sia libera devono presentarsi davanti all'agente corsi di azioni alternativi; la seconda è che l'agente possa autodeterminare le proprie azioni. Il compatibilismo soddisfa la seconda condizione in quanto il soggetto autodetermina le proprie azioni quando esse sono casualmente determinate dalla sua volontà. Molto più difficile per il compatibilismo è dare conto del fatto che un'azione è libera quando sono possibili delle alternative. Questa è una grave difficoltà nella quale incorre la teoria compatibilista. I compatibilisti hanno cercato di rispondere a questa critica sostenendo che «la libertà è il potere di agire in accordo con la propria volontà, nel senso che se la nostra volontà fosse diversa, anche le scelte e le azioni che essa determina sarebbero diverse»²⁹. Per i compatibilisti, allora, la libertà è la possibilità di agire secondo la propria volontà, nel senso che, se la nostra volontà fosse diversa, anche le scelte e le azioni che ne scaturiscono sarebbero diverse³⁰. Così spiega De Caro la risposta che i compatibilisti danno alla critica di non dare soluzione al problema di non poter fare altrimenti:

²⁸ Voltaire, *Il filosofo ignorante*, Bompiani, Milano, 2000, p. 71.

²⁹ M. De Caro, *Il libero arbitrio. Un' introduzione*, Laterza, Bari, 2004, p. 70.

³⁰ Così si esprime Schopenhauer a tale proposito: «l'uomo fa sempre soltanto ciò che vuole eppure lo fa necessariamente. Ciò dipende dal fatto che egli è già ciò che vuole, poiché da ciò che è segue necessariamente ciò che di volta in volta fa. Se si considera il suo fare oggettivamente, cioè dal di fuori, si vede in maniera apodittica, che, come l'agire in ogni essere di natura deve essere sottomesso alla legge di causalità in tutto il suo rigore; soggettivamente invece ognuno sente che fa sempre e soltanto ciò che vuole. Ma questo significa solamente che il suo agire è la pura manifestazione della sua propria natura». A. Schopenhauer, *La libertà del volere umano*, Laterza, Bari, 1997, p. 147.

Immaginiamo che un guidatore, non sottoposto a vincoli o costrizioni, si trovi di fronte ad un bivio: in questo caso volta a sinistra, ma se lo avesse desiderato (giacché non era impedito o costretto) avrebbe potuto girare a destra. Naturalmente, dato che in un quadro compatibilistico si ammette la verità del determinismo, si deve concedere che l'agente era determinato a voler girare a sinistra e dunque lo ha fatto; ma in senso condizionale egli avrebbe potuto fare altrimenti se i suoi desideri fossero stati determinati diversamente; e questo, secondo i compatibilisti è proprio il senso in cui la possibilità di fare altrimenti è rilevante per l'attribuzione della libertà³¹.

Il problema sta nel fatto che compatibilisti e incompatibilisti danno due interpretazioni diverse della possibilità del fare altrimenti: per i compatibilisti, la capacità di agire diversamente ha valore condizionale (l'agente potrebbe agire in maniera diversa se decidesse di agire altrimenti) mentre gli incompatibilisti ne danno un'interpretazione forte per cui l'agente in un certo momento e in una certa situazione avrebbe potuto fare altrimenti da come di fatto farà e non a condizione che la sua volontà sia determinata diversamente. Ma oggi c'è un altro e più stringente argomento che attacca il compatibilismo: il cosiddetto *Consequence Argument*. In questa argomentazione si assume la verità del determinismo e da ciò si conclude che tutte le azioni sono inevitabili, che l'agente non ha la possibilità di fare altrimenti, e che la libertà è impossibile. L'idea centrale del Consequence Argument è che per poter agire liberamente, l'agente deve essere in grado di controllare i fattori che lo muovono a compiere una certa azione. Ma se è vero il determinismo, allora tali fattori riguardano le leggi di natura e gli stati passati dell'universo: e questi sono due fattori che nessun agente può controllare³².

Dunque se è vero il determinismo, nessun agente agisce mai liberamente perché un soggetto, sottostando alle leggi fisiche non può evitare di compiere le azioni che di fatto compie. Come abbiamo visto, sono dunque molte le difficoltà in cui incorre il compatibilismo che, al pare del libertarismo, sembra inadeguato a dare giustificazione della libertà.

Entrambe le teorie del compatibilismo e dell'incompatibilismo si dimostrano insoddisfacenti a far fronte alle problematiche che ruotano intorno al libero arbitrio. Per questo sono sempre più frequenti anche in epoca contemporanea, le conclusioni scettiche

³¹ Ivi, p. 71.

³² Molti autori hanno offerto versioni formali di questo argomento. Da ricordare: Ginet C., *Might we have no choice?*, in Lehrer (a cura di), 1966; Ginet C., *On Action*, Cambridge University Press, Cambridge, 1990; Wiggins D., *Towards a Reasonable Libertarianism*, in T. Honderich (a cura di), *Essays on Freedom and Action*, Routledge and Kegan Paul, London, 1973; Van Inwagen P., *An Essay on Free Will*, Oxford University Press, Oxford, 1983; Fischer J. M., *The metaphysics of Free Will*, Blackwell, Oxford, 1994; Ekstrom L. W., *Free Will. A philosophical study*, Westview Press, Boulder, 2000.

sulla libertà da parte di numerosi autori. Alcuni sostengono che il problema della libertà sia irrisolvibile per i nostri limiti cognitivi. Vi sono filosofi che ritengono, per esempio, che il problema della libertà sia destinato a rimanere un mistero in quanto essa è causazione mentale e noi non potremo mai avere un'idea di come gli stati mentali possano causare cambiamenti nel mondo fisico. È importante sottolineare che questi autori non affermano che non siamo liberi bensì che la questione del libero arbitrio è irrisolvibile. Ciò significa che questo non è un tipo di posizione che si situa su un piano ontologico bensì, epistemico. Essi infatti ritengono che l'idea di libertà sia irrinunciabile per dare un senso ai nostri giudizi morali, ai nostri processi di deliberazione e alle nostre imputazioni di responsabilità. Dunque probabilmente non possiamo evitare di credere a qualcosa la cui esistenza pure non possiamo provare. Da un lato infatti non possiamo rinunciare all'idea di libertà, dall'altra è impossibile darne conto. Altri autori si spingono invece oltre, arrivando ad affermare che non siamo liberi e che la libertà dunque è soltanto un'illusione. Una posizione di questo tipo è stata difesa da numerosi filosofi i quali ritengono che le nostre idee concernenti la libertà siano semplicemente false. Dunque i dibattiti sulla libertà sempre più spesso in epoca contemporanea portano a soluzioni scettiche che possono essere di tipo epistemico, se si afferma che la libertà è impossibile da provare, o di tipo metafisico, se si sostiene che la libertà non esiste e che quindi è una mera illusione. Agli scettici sembrano insoddisfacenti le due grandi famiglie di teorie sulla libertà, il compatibilismo e l'incompatibilismo, perché entrambe non riescono a dare una spiegazione soddisfacente delle due condizioni della libertà: se infatti è vero l'indeterminismo allora gli agenti non possono controllare le loro azioni, se invece è vero il determinismo viene negata all'agente la possibilità di fare altrimenti. Ma perché si possa dire che un agente agisca liberamente occorrono sia l'autodeterminazione sia la possibilità di fare altrimenti. Quindi la libertà è impossibile.

Chiara Erbosì

Bibliografia

- Cini M. (a cura di), *Caso, necessità, libertà*, Cuen, Napoli 1998.
- De Caro M., *Il libero arbitrio. Un'introduzione*, Laterza, Bari 2004.
- De Caro M., (a cura di), *La logica della libertà*, Meltemi, Roma 2002.

- Di Francesco M., *Introduzione alla filosofia della mente*, Nuova Italia Scientifica, Roma 1996.
- Dorato M., *Futuro aperto e libertà. Un'introduzione alla filosofia del tempo*, Laterza, Roma-Bari 1997.
- Dorato M., *Il software dell'universo: un saggio sulle leggi di natura*, Bruno Mondadori, Milano 2000.
- Hobbes T., *Libertà e necessità*, Bompiani, Milano 2000.
- Hume D., *Trattato sulla natura umana*, in *Opere filosofiche*, Laterza, Bari 1971.
- La Mettrie J. O., *L'uomo macchina e altri scritti*, Feltrinelli, Milano 1973.
- Lanfredini R. (a cura di), *Mente e corpo. La soggettività fra scienza e filosofia*, Guerini, Milano 2003.
- Laudisa F., *Causalità. Storia di un modello di conoscenza*, Carocci, Roma 1999.
- Marconi D., *Filosofia e scienza cognitiva*, Laterza, Roma-Bari 2001.
- Marramao G., *Dopo il Leviatano. Individuo e comunità*, Bollati Boringhieri, Torino 2000.
- Mori M., *Libertà, necessità e determinismo*, Il Mulino, Bologna 2001.
- Nannini S., *Cause e ragioni: modelli di spiegazione delle azioni umane nella filosofia analitica*, Editori Riuniti, Roma 1992.
- Nannini S., *L'anima e il corpo. Un'introduzione alla filosofia della mente*, Laterza, Roma-Bari 2002.
- Sartre J. P., *L'essere e il nulla*, Il Saggiatore, Milano 1997.
- Schopenhauer A., *La libertà del volere umano*, Laterza, Bari 1997.
- Searle John R., *Libertà e neurobiologia. Riflessioni sul libero arbitrio, il linguaggio e il potere politico*, Bruno Mondadori, Milano 2005.
- Voltaire, *Scritti Filosofici*, Laterza, Bari 1972.
- Voltaire, *Il filosofo ignorante*, Bompiani, Milano 2000.

MODELLI MENTALI:UN'ANALISI EPISTEMOLOGICA

di Alberto Binazzi

“The limits of our models are the limits of our world”.

P.N. Johnson-Laird

Abstract

The aim of this paper is to investigate the concept of 'mental model' from an epistemological point of view. Since mental models theory accounts for many knowledge's domains (from semantic competence to probabilistic reasoning) we will focus on deductive reasoning and particularly on syllogistic reasoning and propositional reasoning. We will argue that a correct comprehension of Johnson-Laird's theory for deductive reasoning must include a philosophical analysis of the concept of mental model itself: a logical foundation of mental models theory. In order to understand the ontological nature of mental models we will stress the importance of the concept of logical consequence as a property that belongs to the mental models and that relies on the operations and on the structure of models. We will underline that mental models are cognitive structures, which, like theoretical models in logic, are conceived as interpretation I of a set of well-formed formulas such that each member of the set is true *in I*. *Mental* models represent, only what is true in the (mental) models. We will stress that this property, apparently abstract, can impede valid deductions. Besides, we will show that the operations of recursive manipulations of models incorporate (psychologically) the principle of logical consequence, as is described in formal logic. In order to appreciate the theoretical consequences of mental models theory for philosophy of mind as well, we will discuss the relation between mental models and the architecture of mind and we will underline the importance to consider models as embodied structures.

Keywords: mental models, logical consequence, truth, local mental models, models theory, isomorphism, embodied structures.

La teoria dei modelli mentali di Philip Johnson-Laird si applica a tre domini della conoscenza: ragionamento, percezione, linguaggio. È, però, nella sua applicazione al ragionamento deduttivo che la teoria ha dimostrato maggior efficacia esplicativa.

Per comprendere pienamente le implicazioni della teoria dei modelli mentali nella sua applicazione al ragionamento deduttivo, occorre, crediamo, indagare i fondamenti epistemologici del concetto di modello mentale. Concetti come quello di conseguenza logica, verità, validità di una deduzione, non sono soltanto i ben noti principi formalizzati in logica matematica, ma sono, altresì, gli ingredienti principali con cui il concetto stesso di modello mentale prende forma e si struttura.

Crediamo che un'analisi di questo genere possa essere utile sia per chi si occupa di psicologia cognitiva, sia per chi si occupa di filosofia della mente, di logica o di IA.

Gli studiosi della teoria dei modelli mentali hanno dato, finora, molto spazio alle applicazioni pratiche della teoria e ai suoi risultati sperimentali. Indagare, invece, i fondamenti epistemologici del concetto di modello mentale significa analizzare i rapporti che intercorrono tra i modelli come strutture cognitive e i principi logici di verità e validità come sono formalizzati in logica (classica).

Dimostreremo, anche attraverso alcuni esempi, come la teoria per la competenza deduttiva proposta dallo psicologo britannico, si fonda su due principi fondamentali: il principio di verità e il principio semantico di conseguenza logica.

Tali principi, come vedremo, sono costitutivi dei modelli. Il modo migliore di apprezzare la proposta di Johnson-Laird per il ragionamento deduttivo, è mettere in evidenza come la proprietà di conseguenza logica, almeno per i domini finiti (che si ipotizza realizzabile nella mente) diventa una proprietà *derivata* di una data proposizione, nel senso che è dipendente dalla struttura e dalle *procedure* ricorsive che costruiscono e modificano i modelli mentali *di* quella proposizione o *di* un insieme (finito) di enunciati.

I modelli mentali sono un tipo di struttura dati dinamica in grado di rappresentare, attraverso un formato semi-analogico, vari domini di conoscenza (reali, ipotetici, immaginari). Quali sono le particolarità di un modello mentale rispetto ad altre strutture cognitive? Secondo Johnson-Laird (1989) un modello mentale è caratterizzato dalle seguenti proprietà:

- 1) La struttura di un modello corrisponde alla struttura della *situazione* rappresentata nel modello.
- 2) Un modello contiene elementi corrispondenti solo ad entità percepite (perceptible entities) ma, all'occorrenza, può contenere anche elementi corrispondenti a concetti astratti (abstract notions) che dipendono, in modo decisivo, dalle procedure utilizzate per la *manipolazione* dei modelli.
- 3) È un tipo di rappresentazione mentale che non contiene variabili e che può includere soltanto un numero *finito* di occorrenze individuali.
- 4) Un modello è da considerarsi il risultato finale (output) di cinque procedure generali (START, ADD, COMBINE, VERIFY, USE) e due procedure ricorsive (REVISE, REVISE) operanti nei confronti delle occorrenze contenute nel modello. Tali procedure, secondo Johnson-Laird, sono in grado di riprodurre correttamente entità, proprietà e relazioni tra le entità *in* un modello, nonché di incorporare un metodo esaustivo per trarre inferenze deduttive.

Luccio (1998) descrive i modelli mentali come strutture rappresentazionali costituite da occorrenze e relazioni congruenti con i processi e le operazioni incorporabili nei modelli (pp. 25).

Una definizione più formale di modello mentale data da Bara, Bucciarelli e Lombardo (2001) descrive tale struttura come una tripla $\langle T, R, S \rangle$ dove T è un *insieme* non vuoto di tokens (occorrenze) R è un insieme di relazioni che valgono in T e S è un insieme di annotazioni su T (per esempio, il segno \neg che individua una negazione di una data occorrenza).

Recentemente, Johnson-Laird (2007) ha proposto di considerare i modelli come un tipo di struttura cognitiva 3D (un medium rappresentazionale tridimensionale) suscettibile di manipolazione interna.

Prendiamo, per esempio, l'ormai classico caso delle relazioni spaziali. Si consideri questa coppia di premesse (Johnson-Laird, 2006):

Il triangolo è a destra del cerchio.

Il cerchio è a destra del quadrato.

Nel caso della prima premessa, il dispositivo cognitivo costruisce il seguente modello mentale:

(START, ADD):

● ▲

La seconda premessa 'elicita' la costruzione di un secondo modello:

(START, ADD):

■ ●

I due modelli possono, a questo punto, essere integrati in un *unico* modello che conterrà i tre 'tokens', cioè, le tre occorrenze corrispondenti a 'cerchio', 'triangolo', 'quadrato', in una precisa disposizione spaziale.

(COMBINE):

■ ● ▲

Per il teorico dei modelli mentali, tali formati devono essere analogici (iconic as far as possible) e occupare, come gli oggetti dello spazio fisico, uno spazio mentale finito dove possono essere suscettibili di manipolazione geometrica. Inoltre, tali modelli rappresentano lo stretto necessario (le occorrenze precedenti possono essere bidimensionali o tridimensionali, ma questo è superfluo per l'economia della rappresentazione). Di certo, per rappresentare fedelmente la transitività della relazione

precedente, è necessario rappresentare gli elementi allineati in un piano euclideo e non, per esempio, posizionati a formare un triangolo.

Il terzo modello contiene adesso tre elementi e con essi, anche la relazione spaziale tra il quadrato e il triangolo, non esplicitamente *menzionata* nelle premesse. Tale relazione (il triangolo è alla destra del quadrato) può essere letta, estrapolata dal modello, senza premesse aggiuntive che definiscono la transitività della precedente relazione.

In logica, però, una deduzione come quella precedentemente inferita, non sarebbe valida perché la transitività della relazione deve essere esplicitamente aggiunta come premessa della derivazione:

Per ogni x,y,z , se x è alla destra di y e y è alla destra di z , allora x è alla destra di z .

Si noti che da tale 'premessa' aggiuntiva non è consentito accedere al significato della relazione: dal punto di vista della stringa considerata, infatti, non vi è né una destra, né una sinistra. Nella stringa, localmente, non c'è nulla che individui il significato di 'alla destra di'. Tale significato è colto, invece, dal modello mentale precedente, dove la relazione di transitività 'emerge' dalla struttura del modello.

Le due procedure ricorsive possono essere considerate le funzioni più caratteristiche dei modelli: infatti, tali procedure, se applicate al modello precedente (■ ● ▲) cercano di modificare (REVISE) il modello, così da rendere, compatibilmente con le premesse, *falsa* la conclusione provvisoriamente accettata. Se tale modello esiste, allora esso rappresenta un controesempio significativo. A tale modello, si applica, allora, la stessa procedura (REVISE) così da invalidare nuovamente la conclusione e così il processo continua, fino a quando non è più possibile trovare modelli alternativi delle premesse, cioè, nuove disposizioni spaziali delle occorrenze compatibili con la verità delle premesse.

Come sarà chiarito successivamente, secondo i teorici dei modelli mentali, il ragionamento può essere considerato il risultato dei particolari processi mentali implicati nella costruzione, manipolazione, integrazione ricorsiva di tali strutture cognitive.

Un altro elemento di divergenza rispetto alle strutture ricordate è che i modelli mentali sono, per definizione, isomorfici a ciò che rappresentano, nel senso che nel modello viene mantenuta 'la stessa struttura di rapporti' che esiste tra gli elementi del dominio rappresentato (si osservi l'esempio precedente). Strutture rappresentazionali come frames, scripts e schemi, inoltre, non descrivono in che modo le parole si riferiscono al mondo (Johnson-Laird, 1988). La teoria dei modelli mentali assume, invece, che la

comprensione di un enunciato corrisponda ad afferrare le sue condizioni di verità (grasping truth conditions).

Compatibilmente con una concezione internista della verità, Johnson-Laird sostiene che il riferimento di un dato enunciato viene individuato dallo stato di cose immaginato che *rende vera* la situazione descritta dall'enunciato in oggetto: si immagina uno stato di cose (si costruisce un modello mentale dell'enunciato) che corrisponde a "come sarebbe il mondo se l'asserzione venisse considerata vera" (ibidem pp.110).

Lo stato di cose immaginato corrisponde alla sua rappresentazione mentale, mentre il valore di verità dell'asserzione considerata può essere accertato comparando la rappresentazione mentale con lo stato di cose del mondo (percepibile) che essa descrive (ibidem pp.110).

Cosa significa riconoscere che un dato enunciato è vero? Si tratta, crediamo, di effettuare un'*immersione* (in un senso matematico, darne un'interpretazione relativa) del modello mentale dell'enunciato considerato *nel* modello percettivo originario.

L'ipotesi che formuliamo si basa sull'accezione matematica del termine 'immersione'. In matematica, infatti, tale termine indica una relazione tra due strutture tali che una delle due contiene all'interno una 'copia' dell'altra: in questo senso la struttura di un modello percettivo 'contiene' al suo interno una copia del modello mentale, dato che questo è considerato come la versione schematica del modello percettivo originario.

Se, però, è possibile immergere un modello in un altro, data una certa compatibilità strutturale, in che senso possiamo immergere un enunciato in un modello? Facciamo notare che per Johnson-Laird i modelli percettivi costituiscono l'interfaccia privilegiata con il nostro ambiente sensoriale, sono, cioè, rappresentazioni 'compatte' del mondo (Schwamb, 1990). Senza tale 'incorporazione', un dato enunciato risulta privo di significato, inaccessibile alla mente.

Si tratterebbe di costruire e recuperare tali rappresentazioni 'compatte' del mondo, internalizzarle (ecco che diventano mentali) e utilizzarle per ragionare e per comprendere un qualsiasi insieme di enunciati. Tale processo consentirebbe alla mente un accesso naturale alla verità, intesa come meta-proprietà di un enunciato che si 'afferra' solo se è possibile immaginare la struttura percettivo-spaziale di quell'enunciato.

Secondo tale concezione, la verità è una proprietà costitutiva dei modelli, interna, sostanziale (condivisa con la struttura dei nostri corpi, così come dei nostri ambienti di vita). In tal senso, comprendere la verità di un enunciato equivale, cognitivamente, ad ancorare percettivamente l'enunciato. Supponiamo, con Johnson-Laird, che sia così: si

potrebbe dire che a *parti* di modello corrispondono *parti* di mondo proiettati nel modello. Le proprietà geometriche della struttura del vivente (quindi anche dei nostri corpi) saranno mantenute nei modelli mentali: le caratteristiche topologiche di un oggetto dovranno essere mantenute nel modello di tale oggetto. Un modello mentale di un pallone che ruota nel campo visivo, non avrà la struttura di un toro, per intenderci; se, infatti, diamo a qualcuno il compito di disegnare un albero, con molta probabilità, disegnerà una figura che manterrà il principio di continuità nella rappresentazione (mentale): la persona, pur deformando a piacere la figura, utilizzerà rappresentazioni continue della figura rappresentata. È interessante notare come la continuità sia mantenuta anche nella rappresentazione 'mimetica' dell'oggetto che ruota: la mia mano simula una cicloide!

Quando diciamo che un modello mentale (che è una versione più schematica di un'immagine mentale e che può contenere contrassegni simbolici) è *isomorfo* al dominio rappresentato, facciamo alcune ipotesi sul tipo e sul formato di rappresentazioni mentali manipolate.

Geometricamente, tale isomorfismo sarebbe la realizzazione mentale di un'iniezione, nel preciso senso matematico, tra i 'punti' del modello percettivo e i 'punti' del modello *mentale*: in una fotografia, per esempio, dobbiamo poter riconoscere un'immagine rappresentata ed è necessario che siano conservate determinati invarianti strutturali. Punti allineati dell'oggetto devono andare in punti allineati della fotografia, le lunghezze e le dimensioni non saranno alterate eccessivamente, non saranno, cioè, tutte rappresentate in un punto di singolarità e dovrà essere conservata la relazione di vicinanza tra punti: nel modello, come nella fotografia, verrà rispettata la struttura topologica del dominio rappresentato.

Tali assunzioni, precisiamolo, riguardano la natura dei modelli concepiti come strutture cognitive astratte (model theoretic): non sappiamo se ad un livello corticale o sub corticale vi sia un rapporto in scala tra mappe neurali o se esista una corrispondenza punto a punto tra ipotetiche mappe corticali in grado di implementare i processi mentali coinvolti nella genesi di un modello mentale.

Johnson-Laird recupera e sviluppa l'idea introdotta da Wittgenstein e da Peirce di una 'struttura pittografia della forma logica di un enunciato' soggiacente all'enunciato (il quale, aggiungiamo, ne emergerebbe come una figura dallo sfondo). Tale struttura, abbiamo detto, si manifesta nella possibilità logica e quindi psicologica, di poter

immaginare (e quindi interagire con) un modello mentale della situazione descritta dall'enunciato.

Recupera, inoltre, l'idea tarskiana di verità come proprietà relazionale sempre riferita a interpretazioni (modelli) (Mangione, Bozzi 2001).

Se per Tarski la verità è sempre riferita ad interpretazioni insiemistiche, per Johnson-Laird è sempre riferita a specifiche strutture cognitive, i modelli mentali, appunto, che non sono oggetti matematici astratti, ma rappresentazioni mentali originate da input di tipo motorio, percettivo, spaziale, sensoriale e linguistico.

Consideriamo questo esempio. Se diciamo: 'La mela è sul tavolo' (ψ) comprendiamo l'enunciato se siamo in grado di immaginare una situazione che *rende vera* tale asserzione.

-----O-----

Immaginare una situazione che rende vera questa asserzione equivale a 'costruire un modello mentale' dell'enunciato. In questa accezione, la verità non è una proprietà interna all'enunciato, dato che questo (la sua struttura proposizionale) non viene messo in corrispondenza *diretta* con il mondo. Semmai viene messo in corrispondenza con il modello corrispondente, ma soltanto dopo questa fase si può dire che (ψ) è vero.

Facciamo notare che le procedure ricorsive precedentemente illustrate (REVISE) non possono applicarsi alla struttura proposizionale dell'enunciato, ma solo al suo modello mentale.

Se alterassimo, infatti, la struttura superficiale dell'enunciato, come in:

la è tavolo mela sul

non staremmo modificando la struttura di un modello e né saremmo in grado di inferire niente di significativo; sarebbe possibile percepire unicamente l'alterazione sintattica degli elementi grammaticali che compongono l'enunciato o individuare le relazioni spaziali tra gli elementi grammaticali come in 'la è *a sinistra di* mela': ma a questa relazione non corrisponderebbe niente che si trova al di fuori della struttura dell'enunciato. Una caratteristica vitale del linguaggio, invece, è proprio quella di riferirsi ad elementi estrinseci alla propria architettura interna (si esamini il confronto con le rappresentazioni proposizionali).

Se operiamo a livello del modello dell'enunciato come in:

----- ○

Siamo in grado, facilmente, di affermare che 'la mela *non* è più sul tavolo'.

La rappresentazione schematica dei modelli descritti, ripetiamolo, non corrisponde ai modelli così come sono prodotti nei cervelli, né corrisponde all'esperienza cognitiva soggettiva. Ciò che è degno di nota, invece, è che la struttura dei modelli percettivi viene tradotta fedelmente nei corrispondenti modelli mentali, ancorando, così, gli enunciati al mondo (to the models of the *real* world).

Gli esseri umani, infatti, sono in grado di comprendere un dato enunciato per procura (by proxy), perché capaci, potremmo aggiungere, di guardare *nel* linguaggio e individuarvi – recuperare il mondo (e viceversa). Sono, cioè, in grado di percepire attraverso il linguaggio: si potrebbe dire che il linguaggio possieda un paio di occhi (chiasma logico-linguistico). Tale visione, portata all'estremo, consentirebbe di inferire che linguaggio, ragionamento, percezione condividono tutti un livello rappresentazionale comune.

I modelli mentali, lo abbiamo già accennato, si candidano come interfacce cognitive in grado di interpretare domini logici. Crediamo che impostare una riflessione epistemologica nei confronti dei modelli mentali, risulta proficuo anche per chi è interessato alla comprensione del rapporto tra logica e psicologia.

Come mai la logica risulta così ostica a molte persone e perché determinate inferenze risultano più semplici di altre? Cos'è che nei nostri processi mentali fa sì che una forma argomentativa invalida finisca per essere tranquillamente accettata e utilizzata nella pratica quotidiana?

Ci sono delle limitazioni costitutive nella architettura cognitiva della mente che impediscono, del tutto naturalmente, la correttezza di determinati processi mentali? Come è possibile attribuire condizioni di correttezza al nostro stesso 'fenotipo' mentale? Come facciamo a isolare le forme argomentative tipiche della logica dai processi e dalle strutture psicologiche che le implementano fisicamente? È corretto ipotizzare strutture mentali isomorfe a calcoli logici per spiegare i fenomeni mentali implicati nei calcoli logici stessi? Se la scala che conduce verso la logica parte dalla mente, come possiamo, una volta raggiunta la logica, gettare via la scala? In che modo i nostri cervelli sono in grado di implementare processi mentali in accordo o in disaccordo a principi logici? Ipotizzare relazioni logiche tra enti mentali significa concepire, inconsapevolmente, gli enti mentali stessi, siano essi modelli o immagini o altro, come proposizioni. Una relazione come quella di conseguenza logica non è qualcosa che esiste tra i nostri lobi parietali e la percezione

sensoriale centrata sull'ego; ma se la conseguenza logica è una proprietà che si applica (solo) a enunciati, come Frege ci ha indicato, com'è possibile che si possa ragionare per mezzo di strutture cognitive che incorporano tale principio? In che senso possiamo guardare alla struttura logica senza guardare alla forma di quella struttura?

Come vedremo, l'architettura dei modelli è cruciale per ragionare correttamente. È possibile, inoltre, intervenire ricorsivamente su tale architettura interna. La scansione di un modello percettivo o cinematografico, per esempio, può essere naturalmente effettuata da sinistra a destra o da destra a sinistra e le operazioni eseguibili nel modello, come spostare un'occorrenza o eliminarla e sostituirla con un'altra, hanno un 'costo computazionale' misurabile in bit, dipendente dal numero e dal formato dei modelli (Ragni, Knauff, Nebel, 2005). Inoltre, il modo in cui i modelli vengono integrati (eseguendo trasformazioni isometriche) risulta essere fondamentale per leggere dai modelli le informazioni pertinenti (relative alle relazioni logiche tra gli elementi del modello).

La teoria dei modelli mentali nasce per contrastare l'idea che le nostre menti siano dotate di un insieme finito di schemi formali di regole di derivazione che permetterebbero di eseguire inferenze deduttivamente valide, grazie a proprietà logiche insensibili al contesto e innate (Rips, 1994, Braine, o Brien, 1998).

Supponiamo che esista un apparato logico minimale 'naturale' (un modulo fodoriano per la logica): da dove originerebbe la precisa struttura di una forma come $((\psi \rightarrow \phi) \& \psi) \rightarrow \phi$? Dall'innata configurazione delle connessioni neurali o dalla successione temporale di precisi meccanismi biochimici? Dalla struttura logica di derivazioni sintattiche su determinate rappresentazioni mentali? Se qualsiasi schema logico innato esistesse, dovrebbe anche essere, citando Fodor, "incapsulato informazionalmente" e questo, stando alle prestazioni deduttive umane, non si verifica quasi mai.

Un dato schema inferenziale potrebbe applicarsi anche a rappresentazioni mentali che possono, così, venir composte secondo le usuali funzioni di verità. Se assumiamo che i pensieri abbiano un formato proposizionale, simile ad enunciati, e diciamo che la conseguenza logica è un principio che si applica anche ai pensieri e non solo agli enunciati, dovremmo osservare un'ottima competenza deduttiva nei soggetti sperimentali. Se i nostri dispositivi cognitivi fossero dotati di schemi formali operanti su rappresentazioni mentali, allora, nel caso di una deduzione classica come il modus ponens $((\psi \rightarrow \phi) \& \psi) \rightarrow \phi$, il nostro modulo per le deduzioni dovrebbe comprendere una serie di istruzioni come la seguente: se la prima rappresentazione ψ è identica a (sintatticamente unificabile) alla seconda rappresentazione ϕ , e tra ψ e ϕ c'è una 'freccia', allora, scrivi

(deduci formalmente) ϕ . Tali 'istruzioni', completamente automatiche, data la loro natura puramente sintattica, dovrebbero applicarsi a *qualsiasi* insieme di rappresentazioni mentali.

Rips (1994) ha sostenuto che la mente è equipaggiata di un insieme (innato) di regole formali di inferenza. Ma, sostiene Johnson-Laird, il teorico degli schemi formali mentali deve rispondere ad una serie di domande imbarazzanti: che tipo di logica o schemi logici possiede il ragionatore medio? Quali teoremi considerare come schemi? In più, dovremmo legittimamente supporre che esista una qualche logica minimale, un apparato logico completo e decidibile e insensibile al contesto. Ebbene, non sembra che il ragionamento quotidiano e, più in generale, una buona competenza deduttiva, dipendano dall'utilizzo di schemi formali di ragionamento, pronti per l'uso (come la serie di scariche neurali necessarie per la contrazione muscolare) e indipendenti dal contesto.

Consideriamo questo esempio:

Solo una delle seguenti frasi è vera:

Se fra le mie carte c'è un re, allora c'è un asso.

Se fra le mie carte non c'è un re, allora c'è un asso.

Fra le mie carte c'è un re, cosa segue?

Tutti i soggetti sperimentali al quale è stato posto questo problema, hanno risposto che "c'è un asso tra le mie carte" (Johnson-Laird e Savary, 1999). La risposta sembra corretta, ma si tratta di un'inferenza illusoria: infatti, uno dei due condizionali è falso e un condizionale è falso se il suo antecedente è vero e il suo conseguente è falso. Siccome sappiamo dalle premesse iniziali che uno dei due condizionali è necessariamente falso, deve essere necessariamente vero che 'non c'è un asso tra le mie carte'. Se il comportamento del condizionale materiale fosse incapsulato informazionalmente, sostiene Johnson-Laird, i nostri soggetti sperimentali non dovrebbero compiere inferenze invalide: ma è proprio questo che fanno!

Per la verità, la circolarità dell'argomentazione è evidente già a livello della logica pura: siamo in grado di dedurre validamente perché adoperiamo schemi di assiomi simili a quelli formalizzati (appunto!) nel calcolo proposizionale. È piuttosto sorprendente (e imbarazzante) constatare che sono stati necessari numerosi esperimenti di psicologia cognitiva per sbarazzarsi di questa idea.

Ciò che sembra sfuggire sia ai teorici dei modelli mentali, sia ai sostenitori della logica naturale, come vedremo, è il fatto che uno schema formale come quello precedentemente espresso, non prende forma dal vuoto: tale struttura è formale rispetto ad una costruzione già data, cioè all'interno di un universo semantico *costruito* (Longo, 2002).

L'ipotesi fondamentale della teoria dei modelli mentali postula che il ragionamento deduttivo si basa sul processo di costruzione, integrazione, modifica di modelli mentali semi-analogici di stati di cose (Bara, 1990).

Tale definizione, sosterrò, si può ritenere 'l'implementazione architettonica', del principio di conseguenza logica (per domini finiti) che trova la sua formalizzazione matematica all'interno della teoria dei modelli. Come un punto appartenente ad un cerchio in movimento disegna una cicloide, così i modelli mentali vanno concepiti come oggetti dinamici che disegnano la forma strutturale (cinematica) di un ragionamento.

Il principio di conseguenza logica, strumento dimostrativo di logici e matematici, stabilisce che un dato enunciato ψ deriva da una teoria T se e solo se non c'è un modello di T che non sia *anche* modello di ψ , così per provare che ψ non deriva da T, basta *trovare* un modello di T che non sia modello di ψ , cioè un modello che soddisfi T e $\neg\psi$ (Bozzi, Mangione, 2001).

La procedura corretta per il ragionamento deduttivo offerta da Johnson-Laird, prevede tre tipi di 'operazioni di alto livello':

- 1) costruire un modello mentale provvisorio compatibile con le premesse e formulare una conclusione che sia vera nel modello.
- 2) ricercare un ulteriore modello che sia compatibile con un'interpretazione alternativa delle premesse e formulare un'altra conclusione che sia vera adesso nei due modelli fin qui considerati.
- 3) decidere se accettare questa conclusione come definitiva o cercare di interpretare nuovamente le premesse per verificare la verità delle conclusioni formulate o dedurre la loro falsità rispetto ai modelli fin qui presi in considerazione.

Facciamo notare che trovare un modello che soddisfa T e $\neg\psi$ equivale 'cognitivamente' a cercare un modello mentale alternativo delle premesse e formulare, così, una nuova conclusione vera nel modello. Il processo diventa ricorsivo, perché la ricerca di modelli ulteriori continua (si tenga presente, però, che bastano tre modelli per sovraccaricare la memoria di lavoro).

La teoria dei modelli è considerata come la parte più strettamente matematica della logica matematica:

Nel famoso testo di Chang-Keiser (*Model Theory, North-Holland, Amsterdam-London 1973*) tale disciplina veniva sinteticamente descritta come algebra universale + logica, ovvero un'estensione dello studio generale delle strutture matematiche in cui il linguaggio predicativo sostituisce quello equazionale [...]. Un paragone forse più illuminante, espresso da Hodges in *A Shorter Model Theory*, afferma che la teoria dei modelli è una generalizzazione della geometria algebrica, dove si descrivono proprietà e relazioni geometriche attraverso equazioni e sistemi di equazioni algebriche (su corpi, negli sviluppi più astratti). Analogamente, in teoria dei modelli la struttura di riferimento è una qualsiasi struttura matematica e il linguaggio non è equazionale ma predicativo.

[...] Quello che la logica predicativa ha permesso di fare con la teoria dei modelli, è un passo ulteriore nella linea cartesiana. La generalizzazione della teoria dei modelli è duplice: strutture qualsivoglia *hanno sostituito lo spazio* (corsivo nostro) e il linguaggio predicativo ha sostituito quello algebrico-equazionale.¹

Ora, la costruzione di opportuni modelli per una teoria consente di studiarne la struttura logica: si può dimostrare, ad esempio, che una teoria non è contraddittoria (è soddisfacibile) se possiede almeno un modello o che una data teoria è categorica se tutti i suoi modelli sono isomorfi tra loro (Odifreddi, 2003). Lo studio di operazioni e relazioni strutturali porta, così, al concetto di traduzione tra sistemi simbolici diversi e alla *rappresentazione logica* di relazioni entro modelli e teorie (Manca, 2001).

Consideriamo il quinto postulato di Euclide: come dimostrare che non è conseguenza logica degli altri? Interpretando gli assiomi in modelli alternativi a quello classico. Le geometrie iperboliche di Lobacëvskij e di Bólyai, la geometria ellittica di Riemann, il modello di Klein o di Poincaré costituiscono dei modelli nei quali non vale il postulato delle parallele. Se consideriamo, per esempio, 'piano' una superficie sferica, i suoi punti euclidei come punti del piano, e le circonferenze massime della superficie sferica come 'rette', allora, *in questo modello*, l'assioma delle parallele non è più valido: tutte le circonferenze massime di una superficie sferica si incontrano in due punti diametralmente opposti. Interpretando diversamente i postulati, collocandoci, cioè, al di fuori del modello euclideo, il quinto postulato non segue più logicamente dagli altri.

Si consideri il seguente esempio (Manca, 2001): siano dati una segnatura con 4 predicati, P, R entrambi unari e ε , \perp , binario il primo e ternario il secondo: "Essi esprimono le

¹ Manca (2001) pp.152.

proprietà e relazioni geometriche di punto, retta, appartenenza (di un punto ad una retta) e di stare tra due punti (\perp a, b, c se b sta tra a e c)” (pp. 65).

Manca elenca (senza dimostrarle) sei proposizioni che esprimono determinate proprietà geometriche tra gli enti presi in considerazione:

- 1) Per ogni due punti distinti esiste una e una sola retta che passa per entrambi.
- 2) Ogni retta contiene almeno due punti distinti.
- 3) Esistono una retta e un punto fuori di essa.
- 4) La relazione di ‘stare tra’ è simmetrica rispetto al primo e al terzo argomento.
- 5) Dati tre punti distinti allineati, uno di questi è sempre tra gli altri due.
- 6) Data una retta e un punto fuori di essa, vi è sempre una e una sola retta che passa per il punto e che non ha punti in comune con la prima retta.

Come facciamo a dimostrare che l’ultima proposizione *non* è conseguenza logica delle altre?

“Basta prendere come modello i tre vertici di un triangolo, che verificano il predicato P e i tre lati del triangolo che verificano il predicato R, imponendo che ciascun vertice appartenga ai lati di cui è estremo (il predicato di ‘stare tra’ non vale per nessuna terna di punti)” (pp.65):

I lati del triangolo “corrispondono” al predicato R :

P

P Δ P

Tale modello verifica tutte le proposizioni dell’esempio precedente, tranne l’ultima.

Sulla nozione di modello si basano le dimostrazioni di indipendenza: pensiamo all’indipendenza dell’assioma di scelta dagli altri assiomi insiemistici dimostrata da Cohen negli anni quaranta o alle suddette geometrie non euclidee.

Bene, ora pensiamo alle procedure offerte da Johnson-Laird: non è interessante notare che la teoria proposta sembri incorporare la stessa dinamica logica?

Potremmo cogliere il senso della proposta di Johnson-Laird utilizzando la nozione di validità in un modello, così come è formalizzata, appunto, in teoria dei modelli:

$$(M \models \psi \rightarrow \phi) \leftrightarrow (M \models \psi \rightarrow M \models \phi)$$

I modelli mentali, diversamente dai modelli logico-algebrici, incarnano la possibilità di intervenire ricorsivamente nei confronti dei propri processi mentali in atto: si va alla

ricerca di modelli (finiti) compatibili con interpretazioni alternative delle premesse, si valuta la validità delle conclusioni fin qui accettate, si decide se andare alla ricerca di modelli alternativi per mettere alla prova la correttezza delle conclusioni raggiunte. Il processo si svolge ricorsivamente, dato che i modelli incorporano anche le procedure in grado di modificarli e sostituirli.

Consideriamo il seguente sillogismo:

Nessun A è B

Tutti i B sono C

Cosa segue?

Si costruisce un primo modello mentale che realizza le premesse e si formula una conclusione che sia vera nel modello (o meglio, che in base alla disposizione delle occorrenze nel modello, risulta vera, cioè, leggibile dalla situazione immaginata).

Ci fermiamo qui? No, se cerchiamo un altro modello in grado di falsificare la conclusione da noi considerata (che diventa provvisoria). Esiste il modello cercato? Bene, se siamo in grado di disporre le occorrenze in un modo alternativo, allora trarremo una seconda conclusione, che dovrà valere anche nel primo modello costruito per essere considerata valida e così il processo continua:

Ecco i *tre* modelli compatibili con le premesse del sillogismo precedente (Johnson-Laird, 2005).

(START-ADD-COMBINE) → (VERIFY-READ)

[A] ¬ B

[A] ¬ B

[A] [B] C

[B] C

...

(REVISE) → (VERIFY-READ)

[A] ¬ B C

[A] ¬ B

[A] [B] C

[B] C

...

(REVISE) → (VERIFY-READ)

[A] ¬ B C

[A] ¬ B C

[B] C

[B] C

...

Il simbolo “¬” denota il simbolo per la negazione, ciascuna linea denota una determinata relazione logica tra individui, le parentesi quadre indicano che l’insieme degli elementi A e B è stato rappresentato esaustivamente e i tre puntini (...) indicano la possibilità di inserire, a discrezione del solutore, altri tokens che si ipotizza non siano rappresentati esplicitamente nel modello.²

Perché tali strutture si chiamano *modelli* mentali e non immagini o schemi mentali? Semplicemente perché, lo abbiamo già accennato, come i corrispettivi modelli logico-algebrici, i modelli mentali rappresentano soltanto ciò che è *vero*: per definizione, un modello *M* di una teoria *T* contiene soltanto gli elementi di *T* che sono *veri in M*.

Tale principio, osserviamo, impone delle restrizioni operative decisive ai modelli stessi: nel caso dell’applicazione alla logica proposizionale, per esempio, i modelli di una disgiunzione inclusiva, non rappresentano lo stato (o gli stati) di cose in cui la disgiunzione è falsa. Inoltre, “il principio stabilisce che in ogni modello i letterali presenti nelle premesse siano rappresentati solo quando sono veri in un caso possibile” (Johnson-Laird, 2001).

Per esempio, nel caso di una disgiunzione inclusiva:

$(\psi \vee \phi)$

nello stato di cose in cui ψ è vero e ϕ è falso, il modello non conterrà contenuto semi analogico per ϕ . A causa di questa limitazione, non viene conseguentemente rappresentata e percepita internamente l’asserzione “è falso lo stato di cose rappresentato da ϕ ”.

Data, per esempio, la disgiunzione esclusiva precedente, i ragionatori elencano *due* possibilità che corrispondono a *due* modelli mentali:

² A dire il vero, i tre modelli rappresentati sono soltanto l’output dei processi mentali implicati nella deduzione. I modelli vengono generati uno alla volta secondo l’ordine rappresentato. In base a numerosi esperimenti con i sillogismi, è emerso che i soggetti, di norma, costruiscono un primo modello e non prendono in considerazione la possibilità di disposizioni *alternative* delle occorrenze. Chi fosse interessato, invece, al programma per calcolatore sviluppato da Johnson-Laird per la risoluzione dei sillogismi può visitare il sito web: <http://webscript.princeton.edu/~mentmod/models.php/> dove sono presenti, tra l’altro, modelli computazionali (in Lisp) di modelli mentali per la logica proposizionale, il ragionamento modale e il ragionamento temporale.

- 1) ψ
- 2) ϕ

Il primo modello mentale non rappresenta ϕ , che è falso in questo primo modello ($\neg\phi$) e il secondo modello mentale non rappresenta ψ , che è falso in questo secondo modello ($\neg\psi$) (quindi non viene rappresentata internamente l'asserzione che ' ψ è falso'). "Le persone tendono ad ignorare questi casi" (Johnson-Laird 2005).

Gli individui possono però produrre delle note mentali (mental footnotes) dei casi falsi e utilizzare queste note mentali per trasformare (to flesh out) i modelli *mentali* in modelli *espliciti* (fully explicit models).

Per esempio, i modelli mentali del condizionale ($\psi \rightarrow \phi$) sono inizialmente due:

- 1) ψ ϕ
- 2)

dove i puntini denotano un modello implicito (implicito, perché non rappresentato in un primo momento) dei possibili stati di cose in cui l'antecedente del condizionale è *falso*.

Il modello esplicito del condizionale, una volta 'sviluppato' (fleshing out) risulta essere:

- 1) ψ ϕ
- 2) $\neg\psi$ ϕ
- 3) $\neg\psi$ $\neg\phi$

Si consideri questo esempio (Johnson-Laird, 2005):

Ann is in Alaska or Beth is in Barbados, or both.
Beth is in Barbados or Cath is in Canada or both.
What follows?

Nell'esempio considerato, la prima premessa prevede la costruzione di tre modelli, così come la seconda. Il termine 'medio' Beth, permette di integrare i modelli, tramite sovrapposizione *isometrica*, riducendone, così, il numero a tre.

Modelli mentali:

- a) Ann Cath
- b) Beth
- c) Ann Beth Cath

Conclusioni: (((Ann & Cath) v (Beth) v (Ann & Beth & Cath))).

Secondo la teoria dei modelli mentali, un’inferenza deduttiva valida dipende dalla “proporzione di modelli” di volta in volta presi in considerazione dal ragionatore: infatti, ciò che conta è il *numero* di modelli che il ragionatore costruisce e manipola. Questo è un punto controverso: alcuni teorici hanno focalizzato l’attenzione sul numero e sulla complessità di *operazioni* effettuabili nei modelli, o sul tipo di conoscenze (in formato proposizionale o analogico) che possono interferire nello svolgimento corretto della procedura, piuttosto che sul numero di modelli (Ragni, Knauff, Nebel, 2005). Tali studiosi, che hanno proposto un loro modello computazionale basato sui modelli mentali per il ragionamento spaziale, (SRM, Spatial Reasoning by Models) considerano i modelli mentali come ‘mappe spaziali bidimensionali della memoria di lavoro’. Il loro modello computazionale, infatti, dispone le occorrenze in un assetto bidimensionale (2D array) e manipola le occorrenze in modo da ‘inferire’ relazioni di tipo spaziale non espressamente dichiarate nelle premesse.

Altri studiosi si sono soffermati, invece, sulle difficoltà incontrate dai soggetti durante il processo di ‘model editing’, ovvero, sul tipo di operazioni eseguibili all’interno di un modello, in relazione alle informazioni del solutore nei confronti di un dato dominio di conoscenza.

Johnson-Laird ha implementato diversi programmi per calcolatore in cui le conoscenze dichiarative del solutore, incorporate nella struttura dati della memoria, influenzano il corso della deduzione: il fenomeno chiamato ‘modulazione’ (semantic modulation of models).

Si consideri, infatti, tale condizionale:

Se è un gioco, allora non è il calcio (Johnson-Laird, 2005).

La cui rappresentazione ‘modellistica’ (fully explicit) è la seguente:

- gioco – calcio
- gioco – calcio
- gioco calcio

Il significato del termine 'calcio' implica che il calcio sia anche un gioco e sostiene Johnson-Laird, a causa di questo naturale fenomeno, il terzo modello non verrà costruito. Il terzo modello non viene così né immaginato né utilizzato per ragionare: semplicemente 'non esiste' nella mente del solutore.

Secondo Bazzanella, Bouquet e Warglien (2005) i soggetti sperimentali costruiscono modelli piuttosto statici, fissi, non facilmente sostituibili 'in toto' durante il processo di ragionamento (una contraddizione tra modelli dovrebbe generare l'eliminazione dei modelli, ma questo non sembra avvenire 'concretamente'). I soggetti preferiscono, invece, basandosi spesso su un solo modello, modificare *parti* del modello, piuttosto che andare alla ricerca di modelli ulteriori. Un altro elemento divergente con la formulazione di Johnson-Laird è relativo alla confusione, presente nel sistema di credenze dei soggetti, tra soddisfacibilità (verità in almeno un modello) e validità (verità in tutti i modelli considerati).

Altri teorici, mettendo in luce maggiormente gli effetti di natura pragmatica nel ragionamento, evidenziano come i soggetti sono abituati, nella vita quotidiana, a 'interagire' con elementi discorsivi (utterances) piuttosto che con enunciati (sentences). Per questo, non sarebbero in grado di immaginare *sempre* stati di cose alternativi, se tale operazione non viene loro espressamente motivata (Sacco, Bucciarelli, Adenzato, 2001).

Recenti risultati sperimentali (Hagen, C., Knauff, M., Rauch, R., Schlieder, C., Strube, G., 2005) evidenziano come, in realtà, soggetti alle prese con deduzioni di tipo spaziale prendono in considerazione un solo modello e realizzano delle trasformazioni locali del modello considerato. Tale procedura contrasta con la visione classica offerta da Johnson-Laird relativa alla costruzione di modelli alternativi delle premesse 'letti' uno per volta (si osservi i tre modelli corrispondenti alle due premesse del sillogismo precedente: il secondo modello non è una trasformazione del primo, ma è un altro modello, così come il terzo non è una modificazione del secondo).

Johnson-Laird fa spesso esempi in cui sembra considerare il soggetto stesso come *parte* del processo di costruzione – integrazione - modifica dei modelli mentali. Il soggetto dovrebbe decidere di cercare un modello alternativo alle premesse, dovrebbe essere in grado di ripercorrere a ritroso i passi mentali svolti, dovrebbe essere capace di 'guardare' il modello da destra a sinistra, dovrebbe spostare i tokens in modo da invalidare le sue conclusioni provvisorie, per esempio. Benissimo, ma come fa a sapere di poter essere in

grado di realizzare tutta questa ricchezza di operazioni, senza qualcuno che lo istruisca a tal fine?

Insomma, il ragionatore proposto da Johnson-Laird dovrebbe essere capace di effettuare, spontaneamente, una consistente quantità di riflessioni meta-logiche. Come diventa naturale questo processo?

Secondo Yang e Bringsjord (2001) la teoria dei modelli mentali e la teoria della logica naturale (natural logic) non sono altro che le versioni psicologiche, rispettivamente, dell'approccio semantico e sintattico alla logica simbolica. Entrambi i partiti dimenticano di essere, in realtà, due facce della stessa medaglia. È per questo che i due autori sopra citati credono che sia necessaria una terza teoria: la 'meta-logica mentale' (Mental Meta Logic). Tali autori considerano tale progetto come un vero e proprio nuovo settore di ricerca in psicologia cognitiva.

A dire la verità, non c'è bisogno di scomodare teorie cognitive 'apparentemente' risolutive. Già a livello della logica pura la distinzione tra sintassi e semantica è considerata molto labile: le tavole di verità, infatti, vengono *costruite* per mezzo di regole sintattiche codificate dalle tavole dei connettivi e le regole sintattiche per i sequenti sono introdotte in base a considerazioni semantiche legate ai valori di verità (Odifreddi, 2003).

Il problema, sosteniamo, non è nella distinzione tra metodi semantici e sintattici in logica, quanto in una loro 'versione speculare' in psicologia. In logica, infatti, la differenza tra il livello sintattico e il livello semantico è molto meno netta di quanto i due partiti dimostrano di conoscere: quali regole siano sintattiche e quali siano semantiche è una nozione relativa, non sono, cioè, nozioni assolute (Lolli, 2000). In logica, così come in IA, "si lavora con due domini di conoscenza, uno dei quali computa l'altro; il secondo dà la semantica al primo, ma potrebbe essere sintattico rispetto ad un terzo" (pp. 86).

Supponiamo, però, che i teorici di una metalogica mentale abbiano ragione: enunciati metalogici potrebbero essere incorporati in un modello mentale? E nel caso sia possibile, come potrebbe diventare meno ostico per i soggetti incorporare all'interno dei modelli uno specifico contenuto proposizionale in grado di 'orientare' la costruzione di modelli più adeguati? La teoria, cioè, dovrebbe essere dotata di euristiche o criteri di interpretazione in grado di selezionare i modelli pertinenti (Evans, Over 1996). Inoltre, aggiungiamo, una data credenza o è un modello mentale o è una rappresentazione proposizionale. Nel caso sia una rappresentazione proposizionale, rimane da capire come possa essere ancora salvaguardata la natura *analogica* dei processi inferenziali. Per i sostenitori della teoria dei modelli mentali, ricordiamolo, il ragionamento è un'attività cognitiva non soggetta al

controllo sintattico di un ipotetico sottoinsieme 'naturale' di regole come quelle usate nei cosiddetti sistemi di derivazione naturale (Rips, 1994).

Le nostre perplessità sono relative al tipo di ontologia attribuibile ai modelli: si dice che derivano da dati di tipo percettivo, e allo stesso tempo si ammette che sono 'più astratti delle immagini mentali' (Johnson-Laird, 2007). Benissimo, ma come fa un modello mentale, che è un tipo di rappresentazione analogica 3D, a incorporare contrassegni come '-' o '[]'? Il nostro dispositivo computazionale può fare a meno di sapere a cosa tali simboli si riferiscono, ma il soggetto sperimentale *non può*.

È interessante mettere in evidenza come l'attuale dibattito in psicologia cognitiva nei confronti del ragionamento deduttivo rifletta la contrapposizione tra l'impostazione della logica data da Frege, Russel e Whitehead, Post, Hilbert, e la visione della logica del primo Wittgenstein. Secondo Wittgenstein, infatti, in logica, regole di inferenza come il modus ponens, sono da considerarsi inessenziali: "Leggi d'inferenza che come in Frege e Russel giustificano le conclusioni, sono prive di senso e superflue" (5.132). Con le dovute distanze, tale posizione è ancora viva nei sostenitori della teoria dei modelli mentali: infatti, i modelli, come le proposizioni del *Tractatus*, non 'rappresentano' internamente schemi formali d'inferenza. Le relazioni di tipo logico tra date occorrenze 'emergono' dalla struttura dei modelli, così come, per Wittgenstein, la 'derivabilità' di un enunciato non poteva essere letta o inferita dall'enunciato stesso, cioè, dalla sua forma (superficiale) enunciativa. Per Wittgenstein, infatti, le proposizioni di una teoria non potevano essere interpretate per parlare di qualcos'altro, tanto meno per le proprie formule (Lolli, 2004). Tale eredità, poi superata dagli sviluppi della logica matematica moderna con gli importanti risultati limitativi di Tarski, Gödel, Turing, è stata raccolta, con le dovute distanze, ripetiamolo, da Johnson-Laird con i suoi modelli mentali. Nei modelli, infatti, non c'è spazio per regole formali di inferenza: uno schema come il modus ponens non è una regola che guida il succedersi di una proposizione da un insieme di proposizioni. Lo schema modus ponens è da considerarsi, semplicemente, come il risultato dell'immaginabilità e della combinazione di determinati modelli mentali. Immaginiamo ψ , immaginiamo ϕ e 'leggiamo' dalla struttura delle rappresentazioni mentali di (occorrenze mentali di) ψ e ϕ una specifica relazione logica tra le due proposizioni.

In che senso dal modello (e soltanto da questo) possono essere *estrapolate* relazioni logiche?

Ripensiamo ai contrassegni utilizzati per rappresentare i modelli dei sillogismi: operazioni logiche come la quantificazione o la disgiunzione richiedono una qualche

rappresentazione di tipo simbolico, e come ha sostenuto Wilks (1988), anche i modelli mentali non sfuggono alla fallacia simbolica.

Secondo Marco Colombetti (2000) ricercatore presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano, i modelli mentali non possono essere concepiti come strutture completamente analogiche:

Per inferire che due enunciati sono contraddittori, dobbiamo utilizzare assiomi logico-formali o un meccanismo equivalente, dato che l'ancoraggio causale realizzato dall'interfaccia percettiva *non compensa* la natura formale dei simboli (pp. 167).

In logica, i simboli extralogici non hanno un'interpretazione fissa e ammettono infinite interpretazioni (modelli). È necessario, allora, vincolare le possibili interpretazioni di tali simboli e assumere, a seconda dei casi, un bel po' di assiomi. Anche i modelli mentali, sostiene Colombetti, non sono completamente analogici, dato che incorporano contrassegni per la negazione, si uniformano al comportamento delle tavole di verità dei connettivi e a quello dei quantificatori per la sillogistica: per rappresentare correttamente domini logici, è necessario includere assiomi logico-formali nei modelli mentali.

Cosa significa, infatti, rappresentare determinate relazioni logiche in un modello mentale? Immaginiamo una scena (costruiamo un modello) di una particolare disposizione di oggetti in un determinato spazio metrico come in:



Supponiamo di possedere un insieme finito di assiomi incapsulati in un modulo per il ragionamento deduttivo. Che formato avrebbero tali schemi? E, soprattutto, perché dovrebbero avere un formato simbolico? Supponiamo che abbiano un formato proposizionale (D'Mello, S. Grasser, A., Ventura, M., 2005):

Left (▲,■) if Left (▲,●) and Left (●,■).

Left (▲,●) = Right (●,▲).

Right (■,▲) if Right (■,●) and Right (●,▲).

.....

I sostenitori della teoria dei modelli mentali spesso affermano che alcune inferenze, se vi fosse una logica mentale 'naturale', non dovrebbero essere nemmeno *psicologicamente* possibili. Noi tutti, però, nella vita quotidiana, come i soggetti sperimentali, eseguiamo inferenze invalide. Come spiegare questi fenomeni, se nei nostri 'data base' mentali vivessero dei sistemi formali pronti per l'uso? Come garantire la possibilità di ragionare in

modo deduttivamente valido se disponiamo solo di modelli mentali, che, a causa di valori di default, sono il più delle volte, incompleti? I modelli, infatti, non rappresentano tutto lo spazio logico, ma, soltanto un *sottoinsieme* possibile di stati di cose compatibile con un insieme di premesse (Bouquet, 2005).

In uno scenario spaziale come quello precedentemente raffigurato (che si suppone sia neuropsicologicamente localizzabile) siamo in grado di inferire, del tutto naturalmente, relazioni spaziali tra le occorrenze, ma nessuno sano di mente si sognerebbe di dire che, nel ricavare queste conclusioni, noi siamo guidati da assiomi: il fatto è che esistono condizioni naturali (embodied) delle 'restrizioni operative', come ha sostenuto il filosofo Putnam (2003), per le quali alcune inferenze risultano spontanee e altre no (come quelle sillogistiche). Ciò significa, secondo Putnam, che "certi predicati devono applicarsi a certi oggetti ogniqualvolta gli input del computer neurale (vale a dire, gli output dei moduli percettivi) sono di un certo tipo" (2003, pp. 31).

Allo stesso tempo, i modelli mentali, a causa delle limitazioni computazionali del dispositivo inferenziale umano, non sono sempre in grado di implementare conclusioni deduttivamente valide.

Come ha affermato Lolli:

Perché l'autore non conclude che non vi è nessuna cosa come la competenza logica nei suoi soggetti sperimentali? Le persone fanno errori sistematici prima di imparare la nozione di conseguenza, di riconoscere la sua utilità e la necessità di impadronirsene attraverso metodi opportuni. Johnson-Laird ha cercato un modello del comportamento effettivo che potesse estendersi ad un calcolo logico corretto e in ciò ha avuto largo successo; ma l'esperienza fatta con la matematica ci avverte che i più efficienti algoritmi che impariamo e usiamo molto spesso non hanno niente a che fare con le manipolazioni approssimative e sperimentali con le quali incominciamo a trattare i numeri e lo stesso potrebbe essere vero per gli algoritmi logici³.

In lavori recenti, Johnson-Laird ha introdotto all'interno dei modelli mentali varie nozioni logico-matematiche: secondo Vosgerau (2006) risulta molto arduo armonizzare tali nozioni logiche con il formato percettivo dei modelli. Cosa ne sarebbe allora del principio di 'naturalzza' dei modelli? Essi si strutturano a partire da dati percettivi e allo stesso tempo, contengono astratti contrassegni simbolici, come, per esempio, il contrassegno logico della negazione '¬'.

³ Lolli (2000) pp. 97.

Autori come Rips, Wilks, Lolti, appunto, hanno sostenuto che i modelli mentali non sono meno ‘formali’ di altri tipi di strutture dati: infatti, la manipolazione di tali strutture dipende in maniera sostanziale dalle procedure algoritmiche, via via applicate a seconda del dominio di entità che viene modellizzato. Per esempio, nel caso della logica proposizionale, esistono una serie di vincoli alle procedure di integrazione dei modelli alla base dei processi cognitivi implicati.

In *Mental Models and Thought* Johnson-Laird illustra particolari procedure per il trattamento della congiunzione di una coppia di modelli. Le procedure in oggetto non riguardano solo modelli espliciti, ma anche quelli impliciti (non rappresentati).

Elenchiamo le regole espresse da Johnson-Laird (valide per la logica proposizionale):

La congiunzione di due modelli impliciti dà come risultato un modello implicito.

La congiunzione di un modello implicito con un modello che rappresenta proposizioni dà come risultato un modello nullo (null model) come nel caso dell’insieme vuoto.

La congiunzione di due modelli che rappresentano rispettivamente una proposizione e la sua negazione, danno come risultato un modello nullo ($A \neg B$ and A yields nil).

La congiunzione di due modelli in cui la proposizione B, rappresentata in un modello ma non nell’altro, dipende dall’insieme di modelli (set of models) di cui quel modello è membro. Se B occorre in almeno uno di questi modelli, allora la sua assenza nel modello corrente viene trattata come una negazione ($A B$ and A yields nil).

La congiunzione di due modelli espliciti (fully explicit models) aggiorna il secondo modello con tutte le nuove proposizioni del primo modello ($\neg A B$ and $\neg A C$ yields $\neg A B C$)⁴.

Lolti (2000) ha osservato che è opinabile modellizzare “il ragionamento logicamente corretto” assumendo, come fa Johnson-Laird, che i soggetti si uniformino ai principi di verità e validità, e affermando, allo stesso tempo, che “è possibile ragionare validamente senza ricorrere alla logica” (Johnson-Laird, 1983).

Abbiamo sostenuto, infatti, che Johnson-Laird trasporta la nozione di conseguenza logica all’interno di una teoria della mente: ora, per ironia della sorte, questa viene proposta per contrastare la teoria ‘apparentemente’ opposta della logica mentale (mental logic), la quale postula l’attivazione di schemi formali insensibili al contesto o al contenuto di determinare argomentazioni.

Benissimo, ma se questo avviene proprio per mezzo di un principio (semantico) metalogico, come facciamo a riconoscere che è possibile ragionare validamente senza ricorrere alla logica?

⁴ Johnson-Laird (2005) pp. 191.

Inoltre, secondo Lolli, dato che la procedura di Johnson-Laird si basa sulla restrizione ai modelli finiti, non è in grado di applicarsi a modelli infiniti. Ora, siccome l'indecidibilità della logica del prim'ordine è dovuta proprio all'esistenza di modelli infiniti, in che senso la procedura offerta da Johnson-Laird pretende di essere più efficiente di qualsiasi metodo sintattico (pensiamo ai tableaux analitici)?

Lolli chiarisce questo punto delicatissimo:

In realtà la logica dei modelli mentali è la logica dei modelli finiti 'piccoli'; tuttavia, 'piccoli' è un confine disperatamente vago e non computabile. E per ogni confine che si fissa si potrebbe facilmente inventare un esempio che richiede un elemento in più. [...] Così, se la competenza del dominio non è nota, o non è computabile, e per cominciare scriviamo un programma, non siamo in grado di sapere se quel programma è un programma per una funzione ricorsiva che modella la competenza o piuttosto sappiamo che non lo è. Allora, di che cosa è un programma il nostro programma? Possiamo dare come risposta solo l'insieme dei casi praticamente coperti dall'uso del programma e un numero finito di casi non determina un'unica funzione computabile. In conclusione a venir prima è l'algoritmo, cioè l'esecuzione, non la competenza.⁵

Se concepiamo il ragionamento deduttivo come un insieme di funzioni che mappano un insieme di premesse in un insieme di conclusioni (e gli algoritmi di Johnson-Laird realizzerebbero le funzioni suddette) che garanzie abbiamo, nella vita reale e non in una situazione sperimentale, che gli algoritmi selezionati implementino *solo* le conclusioni valide? Nessuna garanzia. Si prenda, per esempio, questo esempio:

If a pilot falls from a plane without a parachute, the pilot dies.
This pilot did not die, however. Why not? (Johnson-Laird, 2005).

Molti soggetti rispondono che l'aeroplano volava vicino a terra, o che l'aereo era finito nella neve o nel mare ecc, ma soltanto una piccola parte risponde in modo logicamente valido:

The pilot did not fall from the plane without a paracadute.

Questo esempio è veramente significativo e sorprendente: la conclusione valida (il pilota non si è buttato senza paracadute) non viene automaticamente immaginata. Si direbbe che sta lì sotto il naso, ma che non la vediamo, perchè preferiamo spiegazioni di tipo

⁵ Lolli (2000) pp. 96.

causale. La maggior parte di noi, assume, infatti, che il pilota si sia effettivamente buttato, come descritto dall'antecedente del condizionale. Costruiamo, in modo del tutto inconscio, modelli mentali cinematografici di un uomo (non pensiamo mai ad un pilota donna) che si butta da un aereo senza paracadute. Non abbiamo immediatamente accesso al modello mentale che descrive, per esempio, il pilota seduto nell'aeroplano, in tutta tranquillità. È incredibile, ma questo tipo di fenomeni, così ben descritti da Johnson-Laird, si presentano continuamente con soggetti alle prese con problemi di logica. In verità, le deduzioni effettivamente valide nella vita quotidiana sono una piccola parte.

È sicuramente merito di Johnson-Laird aver messo in luce come anche il ragionamento deduttivo sia, in parte, guidato da processi mentali automatici che, in prima battuta, non sono del tutto 'penetrabili'.

Abbiamo detto i modelli mentali, a causa dei ben noti limiti della memoria operativa, costituiscono la versione schematica, in un formato simbolico, dei corrispondenti modelli percettivi, nel senso che sono rappresentati da minor grado di specificità e dettaglio. Abbiamo visto, inoltre, che il modello percettivo di una situazione e il corrispondente modello mentale devono essere isomorfi e che tale isomorfismo si conserva anche tra modelli mentali.

Che cosa significa postulare tale 'somiglianza strutturale' nel caso in cui i modelli mentali si applicano alla logica? Analizziamo, per esempio, il comportamento del connettivo \rightarrow (implicazione materiale) in L0 (logica proposizionale). Si tenga in mente la tavola di verità per il connettivo suddetto: la tavola individua tre stati di cose veri su quattro (si considerino due formule ψ e ϕ): (v,v), (f,v), (f,f). Questi sono esattamente gli stati di cose rappresentati nel modello.

Modelli mentali:

| | |
|------------|------------|
| ψ | ϕ |
| $\neg\psi$ | ϕ |
| $\neg\psi$ | $\neg\phi$ |

Ora, di tale tavola è possibile dare la cosiddetta forma normale disgiuntiva (considerando i casi in cui la tavola di verità produce il valore vero): $(\psi \& \phi) \vee (\neg\psi \& \phi) \vee (\neg\psi \& \neg\phi)$.

Notiamo una qualche somiglianza con la descrizione offerta da Johnson-Laird? Questi tre casi non sono forse gli stati di cose rappresentabili in un modello mentale totalmente esplicito? Nel caso di una forma argomentativa con due premesse disgiuntive (due

disgiunzioni inclusive) per esempio, integrare i modelli mentali *corrisponde* cognitivamente ad una 'composizione isometrica' di tavole di verità.

Consideriamo, infatti, tale formula (consequentia mirabilis) : $(\neg\psi \rightarrow \psi) \rightarrow \psi$

Quali sono i modelli mentali corrispondenti?

- 1) $\neg\psi \quad \psi \quad \psi$
- 2) ...

E i modelli espliciti?

- 1) $\neg\psi \quad \psi \quad \psi$
- 2) $\psi \quad \psi \quad \psi$
- 3) $\psi \quad \neg\psi \quad \psi$
- 4) $\neg\psi \quad \neg\psi \quad \psi$
- 5) $\neg\psi \quad \neg\psi \quad \neg\psi$

In logica si distingue tra connettivi e tavole di verità: i primi sono considerati come operatori sintattici su formule, le seconde come matrici che incorporano operazioni semantiche su valori di verità (Odifreddi, pp. 78). Ciò che conta, in logica, è il comportamento vero-funzionale delle formule. Osserviamo, adesso, la struttura del modello dell'implicazione materiale: esso è isomorfo alla tavola di verità dell'implicazione (tranne per la riga corrispondente al caso (v,f) non rappresentato nel modello). Johnson-Laird sostituisce ai valori di verità, formule e negazioni di formule, corrispondenti a formule vere o false. Ricordiamoci, però, che a enunciati di un calcolo logico, corrispondono modelli mentali di quelle formule. Per Johnson-Laird, tale struttura:

$\psi \quad \phi$

corrisponde ad un modello mentale in cui ψ e ϕ sono veri (la prima riga della tavola di verità).

Tale interpretazione, ripetiamolo, deriva direttamente dalla concezione della logica esposta da Wittgenstein nel *Tractatus* (4.441):

È chiaro che al complesso dei segni 'f' e 'v' non corrisponde alcun oggetto o complesso di oggetti, così come nessun oggetto o complesso di oggetti corrisponde alle righe orizzontali o verticali (della tavola di verità). *Oggetti logici non vi sono* (corsivo nostro).

Anche nei modelli mentali, 'oggetti logici' non sono rappresentati: i modelli nascono 'particolarizzati' e corrispondono sempre a *un* caso particolare tra gli infiniti stati di cose rappresentabili (Johnson-Laird, 1983).

In logica, se a formule sintatticamente differenti corrisponde la stessa tavola di verità, allora, tali formule risultano indistinguibili da un punto di vista semantico. Consideriamo due formule equivalenti da un punto di vista semantico come: $\psi \leftrightarrow \phi \approx (\psi \rightarrow \phi) \& (\phi \rightarrow \psi)$ e $(\psi \vee \phi) \approx \neg(\psi \& \neg\phi)$.

Avranno gli stessi modelli mentali, cioè, risulteranno indistinguibili anche da un punto di vista cognitivo? Tali formule avrebbero modelli mentali isomorfi tra loro?

La risposta a tale interrogativo da parte del teorico dei modelli mentali, sarebbe, però, negativa: a causa delle assunzioni di default, dei limiti fisici della memoria di lavoro e della modulazione semantica, i modelli mentali di tali formule nascono incompleti e dunque non possono, cognitivamente, essere logicamente isomorfi tra loro (rileviamo che, se le assunzioni della teoria sono corrette, allora, questa non è una conclusione da poco per chi è interessato ad una proficua didattica della logica).

Ogni tavola di verità può essere rappresentata da un albero in cui le alternative sono rappresentate dalle lettere proposizionali, le biforcazioni dai loro possibili valori di verità e i rami dalle righe della tavola. Ogni ramo è dunque una congiunzione di lettere o loro negazioni e l'intera tavola è la disgiunzione di rami corrispondenti alle righe 'vere'; se pensiamo di rappresentare i modelli come tavole di verità 'incomplete' e postulando tale struttura ad albero come 'forma logica' comune anche ai modelli, possiamo visualizzare il processo di ragionamento (valido) come l'attraversamento di alcuni rami (quelli che corrispondono alle righe vere) dell'albero in questione.

La struttura ad albero di un modello mentale sarebbe, però, a causa delle limitazioni a carico della memoria di lavoro, incompleta (alcuni rami sarebbero *assenti*) e non sarebbe, così, sempre possibile individuare tutti i modelli mentali di una data formula.

I modelli mentali per la logica proposizionale possono essere considerati come il risultato della *proiezione* della struttura geometrica della forma logica (rappresentata, wittgensteinianamente, dalla tavola di verità): tale proiezione conserverebbe la forma logica in piani al massimo 3D (realizzabili nel cervello) di uno spazio logico che può avere infinite dimensioni (spazi logico-vettoriali).

Un essere ipersferico in grado di vedere la profondità dei nostri corpi e dei nostri cervelli, costruirebbe, probabilmente, modelli mentali molto diversi dai nostri. Lo farebbe anche nei confronti di un sillogismo? Svilupperebbe anch'esso una o molteplici logiche? Ma se lo spazio logico è infinito, come è possibile rappresentarci mentalmente tale ricchezza per mezzo di modelli mentali che possono essere al massimo 3D? Al variare dello spazio logico variano anche le relazioni tra gli enti di tale spazio? Pensiamo all'impostazione della

geometria data dal matematico tedesco Felix Klein: secondo Klein, le geometrie studiano le proprietà delle figure geometriche che risultano *invarianti* rispetto ad un certo gruppo di trasformazioni. La stessa impostazione può essere applicata alla logica (o alle logiche). Immaginiamo di immergere lo spazio logico in una superficie elastica, le proprietà logiche tra gli enti di tale spazio subiranno anch'esse delle trasformazioni. Al variare dei modelli, cosa rimane *invariante* nei modelli, rispetto a un certo gruppo di trasformazioni? Tale ricchezza logica potrebbe essere, cognitivamente, rappresentabile in una classe di modelli mentali? In che senso il formato (simbolico) dei modelli può essere considerato causalmente indipendente dal tipo di operazioni effettuabili nel modello e tra i modelli? Consideriamo, adesso, un'applicazione della teoria dei modelli mentali nei confronti della competenza semantica.

Immaginiamo di avere una lista T di infinite rappresentazioni proposizionali (di enunciati). Siano T_1, T_2, T_3 , la lista di tutti i sottoinsiemi (finiti) di T . Supponiamo di riuscire a trovare un modello per ogni sottoinsieme considerato e siano M_1, M_2, M_3 ... tali modelli mentali (locali, che soddisfano, cioè, parti di T , finite). Ora, per il teorema di compattezza che qui estendiamo anche ai modelli mentali, l'esistenza di tali modelli mentali, garantisce l'esistenza di un *modello unico* che soddisfa tutte le formule di T (che può essere diverso da ciascuno dei modelli locali).

La teoria dei modelli mentali afferma che tale modello, qualunque esso sia, debba essere isomorfo a un modello percettivo. Allora, dato un qualsiasi insieme di rappresentazioni proposizionali, è sempre logicamente possibile individuare un modello percettivo di tale insieme. Ora, se tale modello esiste ed è unico, a meno di isomorfismi, che relazione esiste, da un punto di vista cognitivo, tra tale modello e gli altri modelli locali come i modelli motori, sensoriali, percettivi, spaziali e concettuali?

Supponiamo che sia *logicamente* garantita, così, l'esistenza di un modello mentale unico. Adesso, utilizzando lo stesso procedimento adoperato dal matematico John von Neumann nel generare i numeri naturali a partire dall'insieme vuoto, possiamo, partendo da tale modello, generare, ricorsivamente, un modello di questo modello, e così via. Se però, tale modello è isomorfo ad un modello percettivo, esso deriva, necessariamente, da dati di tipo interno (ai nostri sistemi nervosi e ai nostri corpi) e da dati di tipo esterno (relativi alla struttura della realtà). L'incapsulamento di modelli in modelli e la loro successiva memorizzazione potrebbe aver generato un medium rappresentazionale dove ancorare i simboli linguistici?

Supponiamo che un qualsiasi insieme di enunciati attivi il reclutamento automatico di modelli locali (motori, sensoriali, spaziali, percettivi, ecc.). Tali modelli sarebbero suscettibili di connessione solo perché dispongono di un 'modello' che guida il collegamento tra i modelli locali. Tale unità (prototipica?) sarebbe necessaria, ipotizziamo, non solo per comprendere un qualsiasi insieme di enunciati, ma anche per possedere una propria psicologia. Mente e linguaggio, secondo tale visione, sarebbero due occhi in uno stesso volto.

Nei confronti della mente, siamo come un astronomo bidimensionale 'appiattito' su un palloncino (l'universo) che vede intorno a sé le galassie allontanarsi. Mentre il pallone si gonfia, l'astronomo vede le galassie fuggire ad una data velocità, ma non è in grado di vedere il centro di questa espansione, che si trova all'interno del palloncino. Per lui, concetti come 'all'interno di' non avrebbero senso, non sarebbero concepibili, sarebbero, semplicemente, assenti; ma anche per un neurone o per un modulo o per una rete neurale, concetti come 'conseguenza logica', 'verità', 'modello' sono assenti: non troveremo mai un modello mentale guardando dentro una scatola cranica!

Per completezza espositiva, alcune considerazioni più generali sul rapporto tra modelli mentali e architettura della mente.

Jerry Fodor (2001) ha proposto di considerare due tipologie di spiegazioni utilizzate in scienza cognitiva: "le spiegazioni computazionali in senso stretto" e "le spiegazioni architettoniche". Le prime sono teorie che generano derivazioni, ovvero sequenze causali di rappresentazioni mentali, la cui ultima riga è "un'istruzione che comunica ad opportuni meccanismi effettori di produrre il comportamento" (pp. 56). Le seconde invece, "rispondono a questioni su come (in virtù di quale processo causale) la mente passa da una riga all'altra della derivazione in questione" (pp. 57).

"Data una derivazione che include formule della forma A e $A \rightarrow B$, un processo architettonico separa B automaticamente se A e B sono espressioni primitive" (pp. 57). Secondo Fodor, o tali "decisioni di ordine superiore" sono realizzate in modo computazionale (localmente) oppure sono realizzate in modo non computazionale ma automatico: ossia sono conseguenze *causali* del modo in cui le occorrenze delle rappresentazioni mentali interagiscono con l'architettura cognitiva.

Cos'è che in un modello mentale prende 'decisioni di ordine superiore'? Il programma su calcolatore che implementa i modelli mentali 'realizza' tali decisioni localmente, citando Fodor. Ma cosa dire degli esseri umani? Se le forme logiche sono proprietà locali, sostiene Fodor, allora, dovrebbero essere insensibili alle proprietà globali dei sistemi di credenze.

Ma questo, come ha dimostrato Johnson-Laird per la logica, non avviene. Ora, però, se concordiamo con Fodor sul fatto che anche determinate computazioni su rappresentazioni sono anch'esse proprietà locali, allora, anche tali computazioni (le procedure ricorsive nel caso dei modelli) dovrebbero essere insensibili alle proprietà globali dei sistemi di credenze: ma questo, ipotizzando la natura computazionale dei processi mentali, di fatto, non avviene. Cosa dedurre allora? Dobbiamo ammettere l'inadeguatezza di qualsiasi teoria computazionale del ragionamento? Se ammettiamo che i processi cognitivi implicati in un'inferenza deduttiva sono influenzati dalle proprietà globali del sistema di credenze, com'è possibile spiegare le relazioni causali tra tali processi cognitivi per mezzo di computazioni *nei* modelli (locali)? Se la costruzione di un dato modello mentale dipende da computazioni, in che senso, allora, il soggetto è in grado di agire intenzionalmente nei confronti di tali computazioni, modificandole? Com'è possibile mantenendo *fisso* il modello percettivo, manipolare un maggior numero di modelli mentali di questo? Una volta costruito un modello mentale, siamo in grado di ritornare, a partire da questo, al modello percettivo originario o il processo è unidirezionale? Perché questo 'sganciamento' dal dominio basico, rappresentato dalla costruzione dei modelli mentali, dovrebbe mantenere una somiglianza strutturale (schematica) con i modelli percettivi? Esistono prove di natura neuropsicologica in grado di spiegare tale fenomeno?

I nostri sistemi nervosi possiedono un'intricatissima architettura tridimensionale originatasi da un tessuto bidimensionale, il neuroepitelio, la cui struttura si modifica durante lo sviluppo. Le neuroscienze contemporanee sembrano concordare sull'esistenza, nei nostri cervelli, di moduli funzionali spazialmente separati e operanti in parallelo. Il tipo di informazione elaborata da ciascun modulo (visiva, auditiva, olfattiva, tattile, ecc.) può essere analizzata in parallelo, in modo simultaneo, grazie a circuiti neurali specializzati, concepiti come unità locali distribuite in parti diverse del sistema nervoso centrale. L'attivazione congiunta di tali unità attraverso le fibre nervose forma differenti tipologie di sistemi funzionali (Schlosser, Wagner, 2004).

I modelli mentali, come Johnson-Laird a più riprese ha sostenuto, originano da input di tipo percettivo. Abbiamo proposto di considerare una maggiore varietà sensoriale come originaria dei modelli (non solo di tipo visivo). Se l'attivazione congiunta di determinati modelli mentali locali risulta essere determinante per l'ancoraggio cognitivo di un'insieme di enunciati, risulta fondamentale comprendere, neuropsicologicamente, come tale 'combinazione coesiva' avvenga.

Facciamo notare, però, che armonizzare tale richiesta con l'assunto del formato *simbolico* dei modelli mentali diventa problematico. Da una parte, Johnson-Laird sostiene che i modelli contengono soltanto entità di tipo percettivo, da un'altra, sostiene che all'occorrenza possono contenere anche contrassegni simbolici, da un'altra parte ancora, sottolinea, come elemento distintivo della sua teoria, l'assunto della naturalezza (*naturalness*) dei modelli.

Il punto è che se vogliamo comprendere come i modelli mentali sono 'computati' dal cervello, non possiamo assumere il formato simbolico dei modelli: il cervello non è composto da simboli o da occorrenze di enunciati! Noi assumiamo che durante la comprensione di un dato enunciato, accade realmente qualcosa dentro la scatola cranica: ma se esiste qualcosa come la combinazione coesiva dei modelli mentali locali, come abbiamo congetturato, questa non può essere caratterizzata da un'insieme di processi di tipo simbolico sconnessi da qualsiasi esperienza neurologica. Se un illustre scienziato cognitivo come Johnson-Laird, è disposto a rigettare qualsiasi ipotesi di tipo dualistico nei confronti della mente, allora, è in debito di una spiegazione che non sia esclusivamente computazionale, relativa al formato simbolico dei suoi modelli mentali.

Ricapitolando, abbiamo cercato di dimostrare, anche con alcuni esempi, che la proprietà di 'essere conseguenza logica' di una proposizione, diventa, nella teoria dei modelli mentali, una proprietà *dei* modelli mentali (delle procedure incorporare nei modelli mentali).

Tale proprietà non è qualcosa che si può scorgere da fuori, per così dire, guardando la forma proposizionale di una data formula. Soltanto dopo aver costruito il modello corrispondente, e soltanto a quel livello, possiamo dire che una data proposizione è conseguenza logica di un insieme di premesse. A differenza dei computer, gli esseri umani, però, possono generare modelli mentali direttamente da dati di tipo percettivo-spaziale facendo a meno di dati di tipo proposizionale come input. Sono, inoltre, in grado di costruire rappresentazioni mentali come i modelli 3D da dati linguistici in ingresso (e viceversa). Se Johnson-Laird ha ragione, i processi di ragionamento sono fortemente condizionati dal tipo di modelli costruiti (che incorporano sempre assunzioni di default) e dal tipo di operazioni che è possibile effettuare nei modelli e tra i modelli (*scanning, updating, flashing out, integrating, superimposition, ecc.*).

Diventa necessario, allora, studiare le tipologie di operazioni che è possibile effettuare nei modelli e tra i modelli, oltre che impegnarsi in una tassonomia dei modelli.

Come abbiamo precedentemente affermato, crediamo che occorra ridimensionare l'assunto molto limitante del formato *esclusivamente* simbolico dei modelli mentali. Proponiamo di considerare i modelli (percettivi, sensoriali, spaziali, motori, ecc.) come rappresentazioni interne derivate dai nostri sistemi sensoriali e motori. Tali modelli interni sarebbero, altresì, attivati a seconda dei predicati, così interpretati: predicati motori susciteranno la costruzione di modelli cinematici, predicati spaziali e percettivi favoriranno la costruzione di modelli spaziali e percettivi, e così via.

È opportuno evidenziare che i modelli mentali, considerati solamente come strutture 3D suscettibili di manipolazione interna, non sono in grado di incorporare processi mentali di tipo deduttivo. Per fare ciò, ricordiamolo, occorre incorporare nei modelli determinate procedure ricorsive operanti in conformità al principio di conseguenza logica; ma che questo sia realizzabile in un programma per calcolatore, non significa *affatto* che la mente umana o il cervello umano eseguano proprio quel tipo di computazioni (si ricordi l'argomento di Lolli).

I modelli *mentali* potrebbero essere considerati come le 'tracce' (termine mutuato dal vocabolario della topologia) dei modelli (embodied, cioè motori, spaziali, percettivi, sensoriali, ecc) originari. Resta problematico, a nostro avviso, conciliare l'ipotesi del contenuto simbolico dei modelli, con le caratteristiche bio-fisiche dei nostri cervelli e dei nostri corpi. Detto questo, come si 'interconnettono' a livello sub-simbolico i vari modelli? Da un punto di vista computazionale, ci sono delle differenze tra modelli mentali motori e spaziali, per esempio? Tali differenze tra modelli si manifesterebbero anche a un livello linguistico? In che senso le procedure ricorsive che utilizzano i modelli stessi come input, possono applicarsi a modelli motori oltre che a modelli spaziali?

Immaginiamo di ascoltare questo enunciato: "muovi il braccio ritmicamente una volta verso il petto e una volta verso il lato opposto". In tale enunciato, è già contenuto il modello mentale (motorio) dell'enunciato: il modello sarebbe, centrato sull'ego, cinematico, e 'cognitivamente penetrabile', infatti, possiamo modificarlo volontariamente in un'istantanea, in un'immagine statica. Si può parlare, qui, di un livello di ricorsività sub-simbolica, intesa come la possibilità di utilizzare le *stesse* computazioni che hanno generato il modello motorio anche per modificarlo? O sono computazioni qualitativamente diverse? La mente stessa sarebbe il risultato del modo in cui i modelli (locali) si 'collegano' tra loro in un procedimento continuo e tale 'combinazione coesiva' sarebbe mantenuta anche *in* un insieme di enunciati.

Veniamo, adesso, alle conclusioni finali. I modelli mentali incorporano il principio di verità in questa accezione: i modelli mimano il comportamento dei modelli logici come insiemi o strutture che rendono veri tutti gli assiomi di un dato sistema formale.

Possiamo, inoltre, considerare i modelli mentali come ambienti virtuali dove è possibile eseguire determinate operazioni su determinati oggetti in conformità a principi logici (verità e validità). Johnson-Laird recupera, inoltre, l'idea che la verità di un enunciato dipende dal tipo di ambiente matematico (il modello) che realizza l'enunciato. La versione cognitiva di tale principio considera la verità di un enunciato come una proprietà (accessibile dalla mente) dipendente dal tipo di modelli (l'interfaccia tra percezione e linguaggio) che costruiamo, compatibili con l'*immaginabilità* dell'enunciato. Non si tratta di strutture insiemistiche o algebriche, ma di strutture percettivo-spaziali che i nostri cervelli sono in grado di computare. Recupera, come abbiamo visto, alcune idee sulla natura della logica espresse da Wittgenstein nel *Tractatus*.

In secondo luogo, la teoria postula che proprio a causa del principio di verità (il modello rappresenta, per definizione, soltanto i casi veri) i ragionatori non andranno sempre alla ricerca di controesempi (di modelli aggiuntivi) che, appunto, falsificano le premesse e rendono invalide le conclusioni e fa delle previsioni sul numero di modelli che i soggetti prenderanno in considerazione.

Il modello di soggetto epistemico che ne deriva è, come ha sostenuto Legrenzi, (2001, pp.40) quello di un soggetto che si accontenta di rappresentazioni incomplete, di un 'satisfacer', molto simile al soggetto tipico descritto dalla teoria della razionalità limitata di H. Simon.

A differenza della teoria della razionalità limitata, però, la teoria dei modelli mentali, cerca di penetrare nei mondi rappresentazionali dei soggetti e ne descrive il funzionamento. Che questi mondi siano, per la maggior parte, fondati su rappresentazioni di natura percettivo-spaziale, anche quando il contenuto delle rappresentazioni assume un formato simbolico, significa che il prossimo passo della ricerca sarà lo studio comparato dei modelli mentali (motori, spaziali, percettivi, sensoriali, concettuali, ecc) compatibili con le differenti architetture del cervello umano, animale e artificiale.

Bibliografia

- Adenzato, M., Bucciarelli, M., Sacco, K. (2001) Mental Models and the meaning of connectives: A study on children, adolescents and adults, Proceedings of the XXIII Conference of the Cognitive Science Society. In J.D. Moore and K. Stenning (eds.). Mahwah, NJ, Lawrence Erlbaum Associates, pp.875-880.
- Bara, B. (1990) *Scienza cognitiva: un approccio evolutivo alla simulazione della mente*, Bollati Boringhieri, Torino.
- Bara, B.G., Bucciarelli, M., Lombardo, V. (2001) Model theory of deduction: a unified computational approach. *Cognitive Science*, 25, pp.839-901.
- Bazzanella, B., Bouquet, P., Warglien, M. (2005) What do external representations tell about Mental Models? An exploratory study in deductive reasoning, Proceedings of the XXVII Annual Conference of the Cognitive Science Society (Cogsci, 2005) 21-23, July 2005, Stresa, Italy.
- Bouquet, P., Warglien, M., (1998) Mental models and Local models semantics: the problem of information integration, Proceedings of the European Conference of Cognitive Science (AICS'98) University College Dublin (Ireland) 19-21th August, 1998.
- Bozzi, S., Mangione, C. (2001) *Storia della logica: da Boole ai nostri giorni*, Garzanti, Milano.
- Braine, M.D.S. (1978) On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85, pp.1-21.
- Cherubini, P., Giarretta, P., Mazzocco, A. a cura di (2000) *Ragionamento: psicologia e logica*, Giunti, Firenze.
- Chiu, C.Y, Lau, Y.Y.M., Lee, S. L., Kiesler, S. (2004) Human Mental Models of humanoid robots, IEEE, Computer Society, International Conference on Robotics and Automation Proceedings, 2005, pp.2767-2772.
- Colombetti, M. (2000) Ragionare senza significato, in *Ragionamento: psicologia e logica*, a cura di Cherubini, P., Giarretta, P., Mazzocco, A., Giunti, Firenze, 2000, pp. 157-169.
- Cosmides, L., Toby, J. (1996) A logical design for the mind? (Review of the *Psychology of Proof*, Lance Rips, MIT Press, 1994) *Contemporary Psychology*, 41, pp.448-450.

- Davidson, D. (2005) *Truth and Predication*, The Estate of Donald Davidson (trad.it. *Sulla Verità*, Laterza, Roma-Bari, 2006).
- Dennet, D.C. (1996) *Kinds of Minds*, Basic Books, New York (trad.it. *La mente e le menti*, Sansoni, Milano, 1997).
- D’Mello, S., Grasser, A., Ventura, M. (2005) A computational approach to Mental Models, Proceedings of the XXVII Annual Conference of the Cognitive Science Society (pp.2289-2294) Stresa, Italy.
- Eco, U., Santambrogio, M., Violi, P. a cura di (1988) *Meaning and mental representations*, Indiana University Press, Bloomington.
- Evans, J.St.B.T., Over, D.E. (1996) *Rationality and reasoning*, Psychology Press, Hove.
- Fagin, R., Helpert, J.Y. (1988) Belief, awareness and limited reasoning, *Artificial Intelligence*, 34, pp.39-76.
- Ferretti, F., Gambarara, D., a cura di (2005) *Comunicazione e scienza cognitiva*, Laterza, Roma-Bari.
- Fodor, J. (2001) *La mente non funziona così: la portata e i limiti della psicologia computazionale*, Laterza, Roma-Bari.
- Frege (1923) *Ricerche Logiche*, a cura di Michele Di Francesco, Angelo Guerini e Associati, Milano, 1988.
- Frixione, M., (1994) Intelligenza artificiale e semantica: il problema degli atteggiamenti proposizionali. Apparso sugli Atti del Congresso Annuale dell’Associazione Italiana Calcolo Automatico (AICA) Palermo, Settembre 1994, Vol. I, pp.425-437.
- Frosch, C.A, Johnson-Laird, P.N. (2006) The revision of beliefs about causes and enabling conditions, Proceedings of the XXVIII Annual Meeting of the Cognitive Science Society, pp. 1329-1333.
- Hagen,C., Knauf, M., Rauh, R., Schlieder, C., Strube, G. (2005) Preferred and alternative Mental Models in spatial reasoning, *Spatial Cognition and Computation*, Vol.5, N° 2,3, pp.239-269.
- Johnson-Laird, P.N. (1983) *Mental Models*, Cambridge UP, Cambridge (trad.it. *Modelli Mentali*, il Mulino, Bologna, 1988).
- Johnson-Laird, P.N. (1988) How is meaning mentally represented? In *Meaning and mental representations* (a cura di) Eco, U., Santambrogio, M., Violi, P., Indiana University Press, Bloomington, pp.99-119.

- Johnson-Laird, P.N. (1989) *Mental Models*, in Posner, M.I. (Ed) *Foundations of Cognitive Science*, Cambridge, MA: Bradford, MIT Press.
- Johnson-Laird, P.N. (1989) *Pensiero e modelli mentali*, in *Mente umana, mente artificiale* (a cura di) Riccardo Viale, pp. 260-278, Feltrinelli, Milano, 1989.
- Johnson-Laird, P.N. (1993) *Human and Machine Thinking*, Erlbaum, Hillsdale (NJ) (trad.it. *Deduzione, Induzione, Creatività. Pensiero umano e pensiero meccanico*, il Mulino, Bologna, 1994).
- Johnson-Laird, P.N. (1993) *The computer and the mind. An introduction to cognitive science*, London, William Collins Sons e Co.Ltd. (trad.it. *La mente e il computer. Un'introduzione alla scienza cognitive*, Il Mulino, Bologna, 1997).
- Johnson-Laird, P.N. (1996) *Mental Models, deductive reasoning and the brain*, in *The Cognitive Neuroscience* (Eds) Michael S.Gazzaniga, Bradford Book, MIT Press, Cambridge Massachusetts, London, England, pp.999-1008.
- Johnson-Laird, P.N., Savary, F. (1999) *Illusory inferences: a novel class of erroneous deductions*, *Cognition*, 71, pp.191-229.
- Johnson-Laird, P.N. (2000) *Perché la logica non è un buon modello del ragionamento deduttivo*, in *Ragionamento: psicologia e logica* (a cura di) Cherubini P., Giarretta, P., Mazzocco, A., Giunti, Firenze, 2000, pp.17-68.
- Johnson-Laird, P.N. (2001) *Mental Models and deduction*, *Trends in cognitive science*, 5, pp.434-442.
- Johnson-Laird, P.N. (2004) *The history of Mental Models*, in Manktelow, K. and Chung, M.C. (Eds) *Psychology of reasoning: theoretical and historical perspectives*, NY: Psychology Press, pp.179-212.
- Johnson-Laird, P.N., Lee Louis, N.Y. (2004) *Creative strategies in problem solving*, in Forbus, K. Gentnend and Regier, T. (Eds) *Proceedings of the XXVI Annual Conference of the Cognitive Science Society*, Mahwah, N.J.:LEA, pp.813-818.
- Johnson-Laird, P.N. (2005) *Mental Models and Thought*, in Holyak,K., Sternberg, R.J. (Eds) *The Cambridge Handbook of Thinking and Reasoning*, Cambridge University Press, pp.179-212.
- Johnson-Laird, P.N., Yang,Y. (2007) *Mental logic, mental models and computer simulations of reasoning*, in Sun, R. (Ed) *Cambridge Handbook of Computational Psychology*, Cambridge University Press, in press.

- Kandel, E.R., Schwarz, J.H., Jessel, T.M. (1995) *Essentials of Neural Science and Behaviour*, Appleton and Lange (trad.it. *Fondamenti delle neuroscienze e del comportamento*, Ambrosiana, Milano, 1999).
- Knauff, M., Nebel, B., Ragni, M. (2005) A computational model for spatial reasoning with mental models, in Bara, B., Barsalou, B., Bucciarelli, B. (Eds), *Proceedings of the XXVII Annual Cognitive Science Conference (Cogsci 2005)* pp.1064-70, Erlbaum.
- Knauff, M., Johnson-Laird, P.N. (2007) Preferred Mental Models in spatial reasoning, *Memory and Cognition*, 35, pp.2075-2086.
- Lakoff, G., Nunez. (2000) R. *Where mathematics comes from: how the embodied mind brings mathematics into being*, Basic Books.
- Legrenzi, P. (2001) *Come funziona la mente*, Laterza, Roma-Bari.
- Lolli, G. (2000) Un logico esamina i modelli mentali, in *Ragionamento: psicologia e logica*, a cura di Cherubini, P., Giarretta, P., Mazzocco, A., Giunti, Firenze, 2000, pp.81-93.
- Lolli, G. (1989) *Mente e matematica*, in *Mente umana, mente artificiale* (a cura di Riccardo Viale, Feltrinelli, Milano, pp.84-93.
- Lolli, G. (2004) *Da Euclide a Gödel*, Il Mulino, Bologna.
- Lombardo Radice, L. (1970) Un nuovo livello di astrazione: la teoria delle categorie, in *Insiemi, Gruppi, Strutture, Le scienze Quaderni*, n.21.
- Longo, G. (1999) Lo spazio, i fondamenti della matematica e la resistibile ascesa della metafora: il cervello è un calcolatore digitale. Apparso in *L'uomo e le macchine*, a cura di M.Bresciani Califano, Leo S. Olschki, Firenze, 2002.
- Luccio, R. (1998) *Psicologia generale: le frontiere della ricerca*, Laterza, Roma-Bari.
- Manca, V. (2001) *Logica matematica: strutture, rappresentazioni, deduzioni*, Bollati Boringhieri, Torino.
- Odifreddi, P. (2003) *Il diavolo in cattedra: la logica da Aristotele a Gödel*, Einaudi, Torino.
- Palladino, D., Palladino, C. (2007) *Logiche non classiche: un'introduzione*, Carocci, Roma.
- Peruzzi, A.(2004) *Il significato inesistente: lezioni sulla semantica*, Firenze University Press.
- Peruzzi, A. (2005) Il lifting categoriale dalla topologia alla logica, *Annali del Dipartimento di Filosofia* 2006, Firenze University Press.

- Putnam, H. (2003) *Mente, corpo, mondo*, Il Mulino, Bologna.
- Rips L.J. (1994) *The psychology of proofs*, MIT Press, Cambridge, MA.
- Schlosser, G., Wagner, G.P., ed. by (2004) *Modularity in development and evolution*, The University of Chicago Press, Chicago and London.
- Schwamb, K.B. (1990) Mental Models: a survey, url: citeseer.nj.nec.com/schwamb90Mental.html
- Viale, R., a cura di (1989) *Mente umana, mente artificiale*, Feltrinelli, Milano.
- Vosgerau, G. (2006) The perceptual nature of Mental Models, in *Mental Models in cognitive psychology, neuroscience and philosophy of mind*, (eds) Held, C., Knauff, M., Vosgerau, G. Elsevier B.V., pp. 225-275.
- Wilks, Y. (1988) Reference and its role in computational models of mental representations, in *Meaning and Mental Representations* (ed.by) Eco, U., Santambrogio, M., Violi, P., 1988, pp.221-236.
- Wittgenstein, L. (1922) *Tractatus Logico-philosophicus e Quaderni 1914-1916*, Einaudi, Torino, 1994.
- Yang, Y., Bringsjord, S. (2001) Mental MetaLogic: a new paradigm in psychology of reasoning, in L.Chen, Y.Zhou (Eds) *Proceedings of the Third International Conference of Cognitive Science (ICCS 2001)* Press of the University of Science and Technology of China, Hefei, China, pp.199-204.

EVOLUZIONE E LIBERTÁ

di Riccardo Furi

Mai come oggi i risultati delle scienze naturali, come la biologia evolutiva, la biologia molecolare, la neurofisiologia e le scienze cognitive, ci hanno permesso di considerare un quadro generale coerente riguardo l'origine, lo sviluppo e il funzionamento del nostro sistema organico. Tuttavia, i progressi delle neuroscienze e le attuali conoscenze in campo genetico sulla natura dell'essere umano e sul funzionamento dei meccanismi che permettono di svolgere normalmente la nostra vita, non sono diffuse, né accettate come sarebbe auspicabile. Al contrario, sono numerosi i tentativi di limitare l'importanza degli attuali dibattiti scientifici man mano che questi, in seguito alle nuove conoscenze acquisite, lambiscono (e spesso mettono in crisi) i concetti fondamentali che costituiscono la nostra immagine di *individuo agente*.

Seguendo la lezione di Wilfrid Sellars¹ possiamo inserire questo conflitto "divulgativo" nel contrasto dicotomico tra un'*immagine manifesta del mondo*, tipica del senso comune e la sua *visione scientifica*. Non c'è nessun errore, o contraddizione, nel ritenere l'immagine manifesta del senso comune una vera e propria teoria generale sul mondo, costruita attraverso una metodologia e osservazioni oggettive che hanno radici in epoche ancestrali dello sviluppo umano.

Noi viviamo *nella* visione del mondo del senso comune, una visione tutt'altro che ingenua, affinata nel corso della storia e che coinvolge gran parte del pensiero scientifico; l'immagine manifesta, oggi dominante quale cornice concettuale in cui pensiamo noi stessi e il mondo, guida le nostre azioni e deve il suo radicamento ai successi pratici che quotidianamente ci consente di raggiungere. È, indubbiamente, una strategia di successo: fornisce una chiave di lettura del mondo diretta e coerente, plasmata attraverso secoli di esperienze e, a parte alcune inadeguatezze, è ampiamente sufficiente allo svolgimento normale della nostra vita. Esistono vari motivi a difesa dell'utilità di una visione del mondo del senso comune, prima di tutto c'è un vantaggio di comunicazione: non è pensabile una completa sostituzione dell'immagine manifesta con le teorie scientifiche a causa dell'esclusiva specificità terminologica.

¹ Wilfrid Stalker Sellars (1912 – 1989), insegnò filosofia all'Università del Minnesota, a Yale, e poi a Pittsburg. È considerato uno dei più grandi filosofi americani del dopoguerra. I suoi contributi principali riguardano la filosofia della mente, del linguaggio e della scienza.

Le esemplificazioni del senso comune funzionano in quanto descrivono stati noti, su cui c'è un accordo sociale di massima; tuttavia in molti casi le due immagini si sovrappongono e l'immagine del senso comune mutua metodologie e concetti dalle teorie scientifiche, inglobandoli e rielaborandoli: pensiamo alle nuove applicazioni in campo tecnologico o medico i cui concetti sono ormai diffusi e posseduti da tutti gli strati sociali. In effetti il senso comune costituisce il contesto generale al cui interno vanno rintracciati gli esordi della scienza, si può ipotizzare che il senso comune abbia dato inizio all'impresa scientifica, ritenendo che essa gli avrebbe fornito indicazioni preziose sul rapporto dell'uomo con il mondo, anche il senso comune, infatti, ha bisogno di rispondere alle fondamentali domande circa la struttura della realtà.

La scienza viene promossa proprio perché il nostro bisogno di conoscenza è determinato dal bisogno di *agire* nel mondo, un agire sempre più sofisticato, ed è qui che l'immagine manifesta ha perso terreno rispetto all'immagine scientifica, in quanto non è in grado di elaborare teorie capaci di fornire una spiegazione e una previsione dei fenomeni paragonabile a quella delle teorie scientifiche. Oggi nessuno penserebbe di negare l'esistenza degli atomi, o di alcune altre entità rinvenute dal progresso scientifico che hanno contribuito allo slittamento, o all'eliminazione, di numerosi concetti che descrivevano interpretazioni false della realtà. Nonostante la revisione dei vecchi concetti fisici non sia stata indolore, oggi possiamo constatare i notevoli miglioramenti che la nostra immagine del mondo ha beneficiato grazie alle scoperte della fisica.

Auspiciabilmente ci possiamo augurare che la continua ricerca possa svelare le questioni ultime riguardanti la vita sulla terra e la nostra peculiare posizione, domande a cui nei secoli le speculazioni umane più eterogenee hanno cercato di dare delle risposte che sono risultate inadeguate alla costruzione di un quadro concettuale completo; per fortuna attualmente possiamo contare su numerose discipline scientifiche impegnate a fare luce proprio su questi interrogativi di questo genere: la biologia e la genetica, a fronte di una valanga di risultati sperimentali stanno definendo una esauriente cornice esplicativa sulle nostre origini. In seguito alla cornice concettuale esplicitata dalla *teoria dell'evoluzione neo-darwiniana*, si sono delineate varie correnti filosofiche impegnate nella sintesi dei dati scientifici all'interno di nuovi concetti che esprimono correttamente con maggiore rigore empirico ciò che siamo.

Le ipotesi presentate, però, non solo sono a volte svalutate, ma anche pubblicamente osteggiate, perché prevedono una rivoluzione concettuale rispetto alla presunzione di possedere particolari caratteristiche che fanno di noi entità completamente diverse dalle

altre specie, ad esempio l'essere umano come agente morale, responsabile delle proprie scelte. Si teme, nello specifico, che le teorie riduzioniste di stampo materialistico, che descrivono i processi decisionali del cervello come processi fisici, determinati da un rapporto causale quantitativo, se vere, eliminerebbero dal nostro orizzonte concettuale la convinzione di essere *soggetti liberi di agire*.

A fronte delle scorrette e fuorvianti interpretazioni date alle teorie riduzioniste; la tesi sostenuta in questo articolo è che, *nonostante un necessario processo di revisione o eliminazione del concetto comune di libertà, non è limitato, in nessun caso, il nostro potere di effettuare scelte soggettive*.

1. Cosa si intende per libertà. L'incompatibilità col determinismo e gli errori di valutazione del senso comune.

Prima di capire perché l'evoluzione, e il materialismo in generale, non sono una minaccia per la nostra libertà, dobbiamo introdurre il significato di questo concetto, così come viene inteso dal senso comune.

Nella nostra vita di tutti i giorni siamo abituati a prendere continue decisioni, alcune semplici ed elementari che non comportano alcun conflitto interiore, mentre altre, invece, più significative e importanti, sono tormentate e hanno un tempo più lungo di elaborazione per determinare l'azione che riteniamo più giusta: è qui, in quello che possiamo definire il tempo della decisione, che nasce il conflitto con i dati scientifici.

Ognuno di noi, infatti, non dà importanza solo all'esito della decisione, ma anche al modo con cui tale esito viene raggiunto; questo modo ci determina come soggetti morali e ci differenzia dalle altre specie animali.

Certo esistono vari gradi di libertà per cui, alcuni sicuramente, sono appannaggio di molte specie animali: un uccello ha la possibilità di volare dove vuole e i maschi adulti di molte specie possono decidere di accettare o meno lo scontro con un rivale; tutte queste decisioni però non sembrano libere fino in fondo, ma condizionate dalle possibilità che natura e ambiente hanno implementato nell'animale, che fa ciò che per lui è più utile ai fini della sua sopravvivenza, e non sceglie tra ciò che è giusto o sbagliato.

Una scelta in questo senso è volta esclusivamente al soddisfacimento di imperativi biologici: non è libero arbitrio, non è una scelta morale. Osservazioni di questo tipo rischiano di circoscrivere il concetto di libero arbitrio esclusivamente verso quelle scelte che prevedono un atteggiamento morale. Queste, però, rappresentano solo una parte

delle scelte che facciamo e non sono pochi i filosofi che considerano il problema del libero arbitrio come un problema ai margini dell'approccio scientifico.

Tuttavia le scelte morali che prendiamo durante la nostra vita, seppur emotivamente molto sentite, si perdono nella miriade di piccole decisioni quotidiane; da quando ci svegliamo non facciamo che prendere decisioni: come vestirsi, che strada fare per andare a lavoro, cosa mangiare. Possiamo considerare anche gli atti motori, o la preferenza di un sapore piuttosto che un altro come esempi di scelte soggettive e libere: scegliere di muovere il dito indice della mano destra implica comunque una decisione individuale e molte di queste sono comuni a tutte le specie, ma non saremmo disposti ad ammettere che un gatto o un cane scelgano come noi. È prerogativa diffusa intendere le nostre scelte non vincolate a nessuna determinazione, sia genetica e quindi naturale, sia ambientale che culturale; questi possono essere elementi valutativi, ma perché ci sia *vera libertà* l'agente deve decidere in piena autonomia e indipendenza dai condizionamenti, ovvero deve essere pienamente responsabile della sua azione.

Per citare William James²:

Il punto importante è che...le possibilità ci sono realmente...in quei momenti in cui...le bilance del destino sembrano oscillare da una parte o dall'altra...siamo disposti ad ammettere che la decisione non viene presa se non qui ed ora. È questo che dà una palpitante realtà alla nostra vita morale e la pervade...di un entusiasmo così strano e complesso.³

Dovremmo quindi postulare un tempo e uno spazio decisionale in cui il soggetto è virtualmente isolato da qualsiasi condizionamento: dalla sua natura filogenetica, e dai suoi vissuti ontogenetici in modo che possa elaborare gli elementi adeguati a piacimento e, indeterminatamente, scegliere.

Dalle conoscenze in nostro possesso non funziona così, almeno in parte: le nostre azioni sono diretta conseguenze del lavoro del nostro sistema nervoso e il cervello è un sistema deterministico che segue leggi deterministiche.

La maggior parte dei riduzionisti aderisce a una visione del mondo deterministica, cioè causalmente chiusa; il determinismo sostiene che tutta la realtà fisica (escludendo la fisica dei quanti) deriva la sua attuale manifestazione da una serie concatenata di precisi

² William James (New York, 11 gennaio 1842 – Chocorua, 26 agosto 1910), psicologo e filosofo statunitense.

³ William James, *La volontà di credere e altri saggi di filosofia popolare* (*The Will to Believe, and Other Essays in Popular Philosophy*, 1897).

eventi causali e, così, la realtà sarà il prodotto del verificarsi dell'insieme delle cause che lo precedono.

La tesi conclusiva di questa teoria è molto forte, sottolineando come ogni realtà implica un solo futuro possibile, determinato dalle condizioni precedenti alla sua realizzazione. Con una tesi così radicale, se il determinismo fosse vero, perderemmo qualsiasi speranza d'averne libertà di scelta: l'universo è un meccanismo così ben funzionante e concatenato a tutti i livelli che il prodotto delle nostre azioni non può essere diverso da quello che è. Per quanta forza di volontà crediamo d'averne è impossibile opporsi alle leggi dell'universo che ci governano. Il determinismo radicale, così come viene presentato, giustifica il timore di considerarsi meri meccanismi biologici che svolgono inevitabili compiti determinati, non esisterebbe nessun libero arbitrio e la nostra convinzione di poter scegliere liberamente sarebbe una semplice illusione.

La presente tesi viene infatti definita *incompatibilista* nei riguardi del libero arbitrio, è indifendibile e rappresenta le caratteristiche del determinismo portato alle sue estreme conseguenze. La tesi radicale infatti può essere presa in considerazione solo avendo l'opportunità di conoscere lo stato effettivo dell'universo in ogni suo particolare (microfisico e macrofisico), solo così la realtà ci si può dispiegare davanti come un libro già scritto; non a caso il punto di vista radicale è quello, cosiddetto, di "Dio". Quella che interessa a noi è, invece, una versione più moderata, compatibile con il libero arbitrio, o con una facoltà ad esso molto simile.

Il *determinismo compatibilista*, innanzitutto, non osserva la realtà da una prospettiva onnisciente, ma riduce l'orizzonte conoscitivo al punto di vista del soggetto agente, il che implica un notevole scarto predittivo sul futuro. Per come siamo fatti e per ciò che ci è dato conoscere ci è impossibile prevedere con esattezza ciò che avverrà ma possiamo provarci; quando ci troviamo davanti a una scelta l'unica nostra possibilità è selezionare la migliore tra le alternative che siamo in grado di elaborare, in base alle quali misuriamo l'esito futuro dell'azione.

Una limitazione notevole se pensiamo al concetto di libero arbitrio come azione autonoma, incondizionata e pienamente responsabile dell'individuo, ma che ci lascia comunque ampie possibilità di scelta, come questo intervento di J.C.C. Smart⁴ chiarisce esplicitamente:

⁴ J.C.C. Smart è professore emerito di Filosofia presso la Research school of Social Science della Australian National University.

Sostengo un punto di vista filosofico molto familiare: il compatibilismo; ciò equivale a dire che nonostante il determinismo sia vero possiamo comunque conservare un grado di libertà sufficiente per dare all'uomo semplice abbastanza per i suoi scopi pratici.

Non possiamo dare all'uomo semplice tutto ciò che vuole dal concetto del libero arbitrio, perché vuole una contraddizione, vale a dire che le nostre azioni debbano fluire dalle nostre credenze e desideri e parimenti non essere determinate. Agire per puro caso non è agire liberamente: sarebbe orribile!⁵

Possiamo quindi ora includere, a corollario della tesi, questo nuovo concetto di libero arbitrio, così definito in accordo con il determinismo: la nostra libertà di scelta, dunque, si esaurisce nella facoltà di selezionare la migliore azione fra un elenco di azioni alternative elaborate dal nostro sistema, il quale è in grado di manipolare solo determinati elementi rinvenibili sia internamente che esternamente al sistema.

Il libero arbitrio secondo me dipende dal numero spaventoso di connessioni sinaptiche del nostro cervello; noi abbiamo 10^{11} cellule, ciascuna ha circa 10000 connessioni, quindi 10^{15} connessioni. Se dunque esiste il libero arbitrio, e secondo me esiste una maggior o minor libertà d'arbitrio, esso è racchiuso nel numero delle connessioni. I geni sono in grado di determinare alcune delle proprietà dei neuroni, della loro organizzazione in strutture macro e microscopiche e una quota rilevante ma non totale delle connessioni: tutto quello che la nostra biologia e la nostra eredità non specificano, può essere specificato sia dalla nostra storia (ma questo è un altro condizionamento), sia in una grossa percentuale dal caso, sia forse da quello che noi chiamiamo "libero arbitrio".⁶

Seguendo quella che è l'opinione di un noto biologo come Edoardo Boncinelli⁷, adesso non rimane che dare una descrizione euristica di come funziona il nostro sistema cognitivo, e qui entrano in gioco le responsabilità del processo evolutivo. Attraverso la chiave di lettura fornitaci dal neo-darwinismo e sostenuta dai dati della biologia molecolare, possiamo arrivare a comprendere la complessità del cervello umano.

⁵ Benzoni – Coppola, *Nove domande sulla coscienza*, McGraw Hill, 2000. p. 68.

⁶ Benzoni – Coppola, *op. cit.*, p. 65.

⁷ Edoardo Boncinelli (Rodi, 1941) è un genetista italiano. Laureato in fisica presso l'Università di Firenze con una tesi sperimentale di Elettronica Quantistica (Relatore: Giuliano Toraldo di Francia). Nel 1968 una borsa di studio presso l'Istituto Internazionale di Genetica e Biofisica, CNR di Napoli, lo ha introdotto nel mondo della genetica. È stato direttore del Laboratorio di Biologia Molecolare dello Sviluppo presso l'Istituto Scientifico dell'Ospedale San Raffaele e Direttore di Ricerca presso il Centro per lo Studio della Farmacologia Cellulare e Molecolare del CNR di Milano.

2. Eredità e ambiente, i fondamenti dell'evoluzione

Nell'intervento appena riportato, il biologo italiano Edoardo Boncinelli, allude a uno degli argomenti più importanti della neurobiologia contemporanea, secondo cui la variabilità fenotipica neuronale supera notevolmente quella genotipica; secondo i dati a disposizione infatti, le informazioni contenute nel DNA non sono sufficienti a creare la fitta distribuzione spaziale delle sinapsi presenti nel sistema nervoso centrale.

Le istruzioni contenute nel codice genetico bastano giusto a determinare le funzioni strutturali del sistema nervoso, mentre la formazione, stimata, di circa 55 miliardi di neuroni, tra loro interconnessi è dovuta a una peculiare plasticità delle configurazioni neuronali, grazie alla quale innumerevoli connessioni si rimodellano sistematicamente.

La plasticità neuronale permette al sistema nervoso di riconfigurarsi in risposta alle interazioni con l'ambiente. Questa capacità è alla base di ogni processo d'apprendimento e di trasmissione della conoscenza manifestato dalle specie animali; la variabilità fenotipica garantisce, inoltre, notevoli differenziazioni nella risposta da un individuo a un altro. Una maggiore espressione della libertà va dunque ricercata nella maggiore complessità del sistema nervoso e nelle diverse configurazioni di risposta agli stimoli dell'ambiente, e nel nostro caso della società.

Una risposta che in gran parte deriva dalla nostra eredità genetica espressa dal fenotipo, effettiva manifestazione fisica di un organismo, in opposizione al suo genotipo, la totalità delle istruzioni ereditate che porta nel DNA, le quali possono essere o non essere espresse. Questa distinzione genotipo-fenotipo fu proposta da Wilhem Johannsen nel 1911 per rendere chiara la differenza tra l'eredità di un organismo e cosa dell'eredità viene selezionato. Il fenotipo è determinato soltanto in parte dal genotipo e i fattori ambientali possono avere una maggiore o minore influenza a seconda dei casi. Estendendo ulteriormente questo concetto, vengono a far parte del fenotipo di un organismo anche qualità ereditabili più complesse, come ad esempio il suo sviluppo, o il suo comportamento.

L'ereditarietà genetica e l'ambiente sono aspetti fondamentali della teoria evolutiva. Per mezzo di questi due strumenti fondativi la biologia evolutiva sta ricostruendo con minima approssimazione lo sviluppo della vita sulla Terra, dalla più piccola e semplice entità unicellulare, a organismi pluricellulari funzionalmente più sofisticati come i mammiferi.

Molti dei geni attualmente classificati dalla biologia molecolare sono uguali per tutti gli organismi, tanto nell'uomo che nel topo, che nel moscerino. Attraverso la ricostruzione dell'albero filogenetico e dell'albero dei geni che lega le specie in esame, e attraverso la

stima dei geni posseduti dalle specie ancestrali (gli antenati comuni di queste specie) e da "archi-MRCA" (Most Recent Common Ancestor, il più recente antenato comune che lega tutte le specie in esame), è stata ricostruita la storia dell'evoluzione, nell'ultimo miliardo di anni.

Premettendo che la ricerca è tutt'oggi impegnata a esaurire molte incertezze, senza entrare in dettagli che richiederebbero una più completa esposizione delle teorie e di dati precisi ai quali fare scientificamente riferimento, è utile inquadrare un contesto generale che fornisca alcuni strumenti interpretativi utili.

Ai primordi della comparsa della vita sulla Terra, le cellule procariote, sistemi primitivi relativamente semplici, si sono moltiplicate per circa tre miliardi di anni, fino a che, in seguito a una mutazione riuscita, emerse un tipo di cellula, quella eucariote, più grande, costituita al suo interno di organi e parti funzionalmente organizzate. La mutazione che ha permesso l'emergenza del sistema eucariote viene descritta e diffusa da Lynn Margulis. Nel suo lavoro del 1981 *Symbiosis in Cell Evolution (La simbiosi nell'evoluzione cellulare)*, Lynn Margulis sostiene che le cellule eucariotiche si sono originate come comunità di entità interagenti tra loro da cellule procariote invase da un simbionte unicellulare (come batteri e virus); la cellula ospite invece di morire, o eliminare l'invasore, è mutata, riorganizzando la sua struttura in un rapporto di dipendenza simbiotica, diventando una "super cellula". Questa teoria, nota come teoria dell'endosimbionte si fonda sulle osservazioni comparate della struttura cellulare eucaristica e l'azione dei virus: sia i mitocondri che i plastidi contengono DNA diverso da quello del nucleo cellulare e simile a quello dei batteri (per la forma circolare e le dimensioni). Nel corso delle successive proliferazioni la cellula ha perfezionato l'endosimbiosi e alcuni geni all'interno dei mitocondri sono trasmigrati nel nucleo formando quella che è l'attuale unità costituente minima di tutti gli organismi animali.

Usando un'immagine derivata dalle teorie di tipo computazionalista potremmo assimilare la cellula ad un piccolo processore.

Il nucleo è l'elaboratore centrale che manipola le informazioni contenute nel codice genetico, dove specifici geni all'interno della stringa di informazioni, come i geni Hox, detti "geni architetto", contengono tutto il progetto dell'organismo, così il nucleo, elabora le specifiche istruzioni (come potrebbe fare un qualsiasi calcolatore che segue procedure algoritmiche), e attraverso sintesi di proteine e amminoacidi, permette alla cellula di ricoprire il ruolo che le compete, nei pesci, rettili, uccelli o mammiferi, sia questa una cellula epiteliale, o un neurone.

Dalla comparsa dei primi cordati, organismi dotati di scheletro, circa 530 milioni di anni fa i pesci si svilupparono in una vera e propria “corsa agli armamenti”; per quasi trecento milioni di anni il graduale cambiamento dell’ambiente ha favorito innumerevoli mutazioni come se il nostro pianeta fosse un’enorme laboratorio sperimentale da cui emergevano specie filogeneticamente sempre più sofisticate e adatte. Con l’emersione della terra molte specie marine hanno mutato il loro apparato polmonare per poter respirare anche fuori dall’acqua: è dagli anfibi che poi si sono sviluppati gli abitanti della terraferma.

Tutto grazie alle diverse istruzioni che per trasmissione genetica sono contenute nel DNA, decodificato dal nucleo della cellula. Tutti i mammiferi placentati vissuti circa 140 milioni di anni fa, all’inizio del Cretaceo, mostrano di possedere un unico antenato rappresentante da un piccolo insettivoro dal quale differenti strade evolutive portano nell’Alto Cretaceo, 85 milioni di anni fa, fino all’ingresso sulla scena di un gruppo molto vario di mammiferi, detto dei *Laurasiateri*. e dopo altri 10 milioni di anni la comparsa dei piccoli roditori, quali il topo e il castoro, nostri vicini co-antenati. Già in questo periodo è evidente l’enorme evoluzione strutturale e cognitiva del cervello e del sistema nervoso in generale, che permette l’emergere di nuove funzioni in aggiunta a quelle del cervello rettiliano, come l’apprendimento dall’interazione con l’ambiente e le sue modificazioni, nuove interazioni sociali (come il riconoscimento della prole e del gruppo) hanno permesso la trasmissione orizzontale (in opposizione alla trasmissione genetica detta verticale) delle conoscenze: pensiamo all’abilità del castoro nel modificare l’ambiente costruendo dighe, e come questa abilità viene insegnata alle nuove generazioni.

Per circa 40 milioni di anni, fino a circa tre milioni di anni fa, lo sviluppo dei primati ha portato alle specie oggi conosciute come gli scimpanzè, i gorilla, gli orang-utan: esperti nel bipedalismo e nella brachiazione. Tra 3 e 6 milioni di anni fa incontriamo tutta una serie di ominidi, apparentemente bipedi, che via via si avvicinano al genere *Homo*; comincia la “rivoluzione della corteccia cerebrale”: due milioni di anni fa oggi incontriamo *Homo habilis*, cosa lo distingue dalle australopithecine? Il volume encefalico su cui si basa lo studio della correlazione tra massa corporea e massa encefalica, per arrivare poi al Quoziente di Encefalizzazione (EQ), che rivela una letterale esplosione della nostra massa cerebrale negli ultimi 3 milioni di anni.

La formazione della corteccia cerebrale è evolutivamente recente e solo l’uomo la presenta in maniera così sviluppata; ma, se la pensiamo nel contesto di tutto il processo evolutivo, il progetto del nostro sistema nervoso si è sviluppato da quasi un miliardo di anni. Nel modello del “cervello tripartito” descritto negli anni cinquanta dal

neurofisiologo Paul McLean (MacLean P.D, 1964 *Man and his animal brains*, *Mod. Med.*, 32: 95-106), si possono distinguere chiaramente le “fasi di costruzione”.

Secondo questa teoria il cervello è organizzato in tre strutture sovrapposte ereditate nelle successive fasi evolutive: la parte più arcaica, collocata all'estremità della colonna vertebrale, costituisce il tronco encefalico e altre formazioni neuronali come l'Amigdala. Definito come cervello rettiliano, questo primordiale *cerebrotipo* possiede funzioni elementari e necessarie per la sopravvivenza, l'espressione comportamentale è fissata in schemi ripetuti e memorizzati a seguito di esperienze ataviche, fissate nel codice genetico e trasmesse ereditariamente come la lotta per la sopravvivenza dell'individuo e il soddisfacimento dei bisogni biologici propri e della specie.

Al cervello rettiliano si sono sovrapposte altre strutture organizzate in un sistema (che corrisponde al sistema limbico dei mammiferi moderni) con una maggiore quantità di corteccia primitiva; questo secondo *cerebrotipo*, possessore del cervello limbico (la corteccia limbica ha caratteristiche simili in tutti i mammiferi ma strutturalmente meno sofisticata della neocorteccia), è dotato di maggiore sensibilità emotiva. Il cervello limbico regola infatti i processi viscerosomatici ed emotivi, è il primo sistema omeostatico, in grado di elaborare le informazioni che arrivano dall'ambiente, tramite le afferenze dei sensi, e quelle somatiche interne.

Il comportamento del secondo *cerebrotipo* è già notevolmente meno ripetitivo e stereotipato del possessore del cervello rettiliano, la maggiore sensibilità e la spinta alle relazioni emotivo-sociali permette una maggiore variabilità alle risposte, un più alto grado di apprendimento e di trasmissione orizzontale della conoscenza. Test su scimmie e gatti, sottoposti a stimolazioni del cervello limbico dimostrano decise modificazioni nel comportamento in relazione alla zona stimolata.

La terza e più recente struttura formatasi è la neo-corteccia, il cervello neo-corticale si è evoluto parallelamente alle nuove funzioni cognitive dei primati e degli ominidi, appena evidente in questi, durante gli ultimi due milioni di anni si è notevolmente sviluppato rendendo possibili tutte le funzioni razionali dell'uomo moderno. Lobo frontale, lobo parietale, lobo occipitale e lobo temporale, sono alcune delle zone specifiche della neo-corteccia che ricopre le altre due strutture ed è deputata alle funzioni cognitive complesse, come il linguaggio, l'apprendimento e l'elaborazione concettuale.

Questa nuova struttura, somma degli sforzi evolutivi, rinvenibile solo nell'uomo, è strettamente collegata da numerose connessioni con le strutture sottostanti, ma non possiede connessioni dirette con il mondo esterno tramite le afferenze sensibili. Per

sfruttare ancora l'analogia computazionalista, la neo-corteccia ha la funzione di elaborare i dati registrati dalle altre due strutture, possiamo considerarla come un processore centrale estremamente potente in grado di correlare in parallelo le informazioni a sua disposizione, permettendoci di considerare, molto rapidamente, in base all'analisi concettuale, gli esiti possibili di molte alternative d'azione.

Lo sviluppo dell'encefalo nel terzo cerebrotipo ha permesso la nascita del linguaggio e del ragionamento logico, due caratteristiche che hanno spinto la nostra specie alla ricerca della conoscenza; grazie alla comunicazione interindividuale e l'analisi logica l'uomo ha potuto perfezionare l'abilità manuale e la capacità di creare utensili e strutture in grado di modificare l'ambiente in cui viveva.

L'innovazione tecnologica e la conseguente fine del nomadismo, legata all'invenzione dell'agricoltura e la capacità di comunicare e tramandare le conoscenze acquisite alle generazioni successive, ha consentito agli uomini di costituirsi in comunità. Sono state le prime comunità e le loro relazioni che hanno determinato lo sviluppo culturale (e quindi etico ed estetico) che sotto altre forme influenza ancora oggi le nostre esperienze quotidiane, ed è grazie alle notevoli capacità logico-deduttive di prevedere futuri possibili in base alle nostre conoscenze, che siamo liberi.

3. Le obiezioni alla spiegazione deterministica ed evolutiva

La lunga concatenazione deterministica dello sviluppo evolutivo trova nel nostro sistema nervoso centrale la sua più sofisticata manifestazione. Un sistema anch'esso descrivibile secondo le leggi naturali; tuttavia, per chiarire i restanti dubbi che il punto di vista del riduzionismo inevitabilmente solleva è necessario considerare anche le ragioni degli anti-riduzionisti.

Ci sono numerose ricerche sperimentali, come ad esempio gli studi di Benjamin Libet, che mostrano come il tempo di svolgimento del processo cerebrale che porta al compimento dell'azione, preceda il momento in cui l'agente è consapevole di produrlo, in breve la decisione avverrebbe prima della nostra consapevolezza della stessa.

La conclusione, per molti riduzionisti dell'ala radicale, è ritenere la nostra credenza di guidare la scelta un fenomeno illusorio. Il nodo filosofico maggiormente dibattuto dagli anti-riduzionisti riguarda il fenomeno che soggettivamente tutti esperiamo, considerato reale, irriducibile e causalmente efficace, ma che analizzato con gli strumenti della scienza, della cosiddetta ricerca in terza persona, è completamente determinato dal cervello.

Come è possibile conciliare i risultati sperimentali della neurofisiologia con un fenomeno il cui esclusivo accesso soggettivo ci viene dato con ferma evidenza?

Thomas Metzinger, dell'Università di San Diego ha proposto una teoria che solleva alcuni dubbi nei confronti delle conclusioni riduzioniste compatibili col determinismo: l'esperienza interna del fenomeno che rappresenta la libertà di scelta sarebbe un *processo agente allucinato (hallucinated agency)*.

Credo che ciò che debba essere attentamente analizzato sia l'esperienza soggettiva di processo agente attuale (*agency*). Perché ci esperiamo come liberi? (...) Ma che cosa è la volizione consciamente esperita? Secondo la mia teoria, essa è il contenuto fenomenico attivato da un modello trasparente del processo di selezione interna che guida l'organismo a eseguire, per esempio, uno specifico schema di comportamento motorio scelto tra un numero di possibili comportamenti, che sono stati preventivamente simulati internamente.⁸

Il fenomeno della volizione, continua Metzinger, non è introspektivamente rinvenibile finché non diventa cosciente, oggetto cioè di pensieri di ordine superiore, (ovvero, aggiungerei, dato di elaborazione della neo-corteccia). Al momento in cui tale fenomeno viene rinvenuto dalla "lente d'ingrandimento" della coscienza noi lo esperiamo come un *nostro* atto libero e autonomo, poiché viene integrato nella costruzione della nostra immagine del sé (*self-model*). L'emergere alla coscienza, di una o più alternative, avviene, però, successivamente alla selezione delle strategie eseguibili, così, se certamente questa può avere un ruolo, in qualità di supervisore, nella flessibilità della scelta o di veto, non ha alcuna responsabilità causale. In poche parole la nostra coscienza né crea, né causa la scelta.

Concesso allora che la scelta sia determinata dal cervello, com'è possibile definire questa "razionale" o "autonoma"? La risposta formulata dagli anti-riduzionisti è affidata in questa sede a Robert Kane, anti-riduzionista e naturalista, magistralmente riportata da D. Dennett nel suo "*L'evoluzione della libertà*".

Nel suo "*Significance of Free Will*" (1996), Kane parte dalla concezione comune di libero arbitrio, sostenendo che perché ci sia libertà è necessario che la lunga catena causale a cui siamo legati si interrompa in qualche punto e questo, per il nostro filosofo anti-riduzionista, non può che collocarsi all'interno del nostro processo decisionale; è nella nostra facoltà di ragion pratica, continua Kane, che si manifesta l'unico momento di *indeterminazione* dell'universo (meccanica quantistica a parte).

⁸ Benzoni – Coppola, *op. cit.*, p. 71.

C'è almeno un tipo di libertà che vale che è incompatibile con il determinismo, ed è un tipo significativo di libertà che vale la pena desiderare. Si tratta del potere di essere i creatori ultimi e i sostenitori dei nostri fini o scopi.⁹

Il modello del processo decisionale descritto da Kane, è determinato in tre fasi che elaborano gli input (interni ed esterni) da cui conseguono gli output: la prima fase, detta *volontà desiderativa*, fornisce gli input che vengono sintetizzati in output dalla *volontà razionale* se il processo è trasparente e non presenta intoppi. Quando, al contrario, il processo decisionale si trova a considerare più alternative in conflitto tra loro, ecco che entra in gioco la *volontà combattiva*. Durante la fase “combattiva”, quando il soggetto è indeciso tra più alternative contrastanti, si apre una finestra in cui tutto è possibile, anche violare le leggi del determinismo:

se c'è indeterminazione nel libero arbitrio, dal mio punto di vista, deve emergere da qualche parte tra l'input e l'output.¹⁰

Nella situazione conflittuale descritta dalla volontà combattiva, secondo Kane, si formano molteplici reti neurali ricorrenti e interconnesse che rinviano serie di impulsi che interferiscono tra loro finché una non prevale. L'attività caotica che deriva da queste reti non lineari permette una soluzione “creativa” al problema, in accordo con l'estrema plasticità e flessibilità neuronali; questo però non basta, come il nostro filosofo sa bene, le dinamiche non lineari, descritte dalla *teoria del caos*, riguardano ancora fenomeni deterministici.

L'agente, così, si trova in uno stato conflittuale volendo compiere l'azione sulla base di più motivazioni e quando il miglior output non è definibile il sistema introduce variazioni “casuali”. Anche le variazioni di tipo casuale, come il lancio di una moneta sono deterministiche, legate cioè, a implicazioni impercettibili che generano uno “pseudocaso” non rinvenibili per il nostro limitato punto di vista; ma ciò che Kane vuole è una genuina casualità, così sostiene che in un punto tra l'input e l'output deve esserci un istante di *indeterminazione quantistica*, nel momento in cui, l'ipotetico ago che oscilla tra l'azione A e l'azione B, deve esserci un istante in cui entrambe le azioni contrastanti si presentino, insieme, ugualmente disponibili.

Solo in questo caso sarebbe lecito dire che il soggetto poteva avere un'alternativa da scegliere, solo se la scelta non è deterministica, solo così, secondo Kane, è possibile

⁹ D. Dennett, *L'evoluzione della libertà*, Raffaello Cortina Editore, 2004, p. 138.

¹⁰ D. Dennett, *op.cit.* p.146.

eludere il principio deterministico per cui date tutte le informazioni sullo stato della realtà, esiste un unico futuro possibile e quindi, per il nostro agente, una sola azione possibile.

Kane pensa in questo modo di aver salvato il libero arbitrio, al contrario potrebbe aver complicato la situazione, pur di togliere l'individuo dalle proprie influenze lo ha lasciato scegliere a caso.

Ammettiamo comunque che la sua ipotesi sia plausibile e che decisioni particolarmente sofferte, quelle più importanti, siano affidate ad eventi indeterminati, più del lancio di una moneta: innanzitutto, considerata la natura dell'indeterminazione quantistica, non sarebbe possibile stabilire quale decisione sia o non sia deterministica; inoltre è impossibile che una tale facoltà del nostro sistema resti isolata da tutti i condizionamenti genetici, ambientali e culturali. Infine, in considerazione di quanto esposto, sorge spontanea la riflessione su un dubbio che si poneva William James, a cui Kane, e coloro che come lui sono spaventati dalle implicazioni del determinismo dovrebbero rispondere:

Se un atto libero fosse qualcosa di assolutamente nuovo, che non viene da me, dall'io che ero prima, ma ex nihilo, e semplicemente si attacca a me, come posso io, l'io che ero prima, essere responsabile?¹¹

Conclusioni

Il concetto di libero arbitrio, condiviso dal senso comune, ha una valenza pragmatica confermata induttivamente dal successo nelle pratiche quotidiane, ma è falso: la libertà di scelta deriva direttamente dal numero delle alternative che il nostro cervello è in grado di selezionare; il nostro cervello è un sistema deterministico progressivamente perfezionato dall'evoluzione nell'ultimo miliardo di anni; tutte le tappe, dagli organismi unicellulari agli uomini, fino ai meccanismi di selezione dell'azione da eseguire sono causalmente determinate da eventi macrofisici e microfisici.

Descritto in questa forma succinta e schematica il potenziale rivoluzionario del concetto scientifico di libero arbitrio turba sicuramente il sonno di coloro che sono alla guida di quelle istituzioni che si arrogano il diritto di indirizzare l'individuo nelle sue scelte morali, come le istituzioni religiose. Inoltre lo spettro dell'anarchia derivante dalla possibilità della completa deresponsabilizzazione dell'azione influenza scelte politiche conservatrici che non aiutano a familiarizzare con questi nuovi concetti. Senza entrare nel dibattito tra scienza e fede, forse oggi è giunto il momento di smettere di negare e demonizzare

¹¹ William James, *op. cit.* p. 210.

ipotesi provenienti dalla ricerca scientifica. L'allarmismo, al contrario, rischia di far penetrare queste idee nell'immagine manifesta in modo non corretto e superficiale. Siamo ciò che possiamo essere, in base alla nostra natura e alla nostra cultura, una cultura perfezionabile e rivedibile nei suoi concetti. Il nostro sistema cognitivo, la nostra natura, ci rende capaci di elaborare scenari predittivi, astrarre e combinare informazioni: immaginiamoci i primi insediamenti umani, culla del primo sviluppo culturale, e come l'uomo del neolitico poteva spiegarsi accadimenti quali la morte. In mancanza di strumenti conoscitivi l'invenzione del soprannaturale sopperiva alle lacune esplicative presenti nel crescente bisogno degli uomini di sapere. La nostra vita e pure la società sono fondate sullo scambio dell'informazione e solo l'uomo è capace di manipolarla in un senso così rivoluzionario per la natura, ma che a noi pare scontato. L'unico sistema per dipendere meno dai condizionamenti che regolano la nostra esistenza, e quella di tutti gli organismi, è, quindi, disporre di più informazione possibile in modo da poter selezionare un maggior numero di scelte. Non è affatto limitata la nostra libertà di scelta, finché noi stessi non limitiamo le possibilità di scegliere.

Riccardo Furi

Bibliografia

- William James, *La volontà di credere e altri saggi di filosofia popolare (The Will to Believe, and Other Essays in Popular Philosophy, 1897)*.
- Benzoni – Coppola, *Nove domande sulla coscienza*, McGraw Hill, 2000.
- D. Dennett, *L'evoluzione della libertà*, Raffaello Cortina Editore, Milano 2004.
- W. Sellars, "Philosophy and the Scientific Image of Man", in W. Sellars, *Science, Perception, and Reality*, Routledge & Kegan Paul, London, 1963.
- Richard Dawkins, *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution*, 2006.
- Robert Kane, *Significance of Free Will*, 1996.
- Paul McLean, *Man and his animal brains*, *Mod. Med.*, 32, 1964.
- Lynn Margulis, *Symbiosis in Cell Evolution*, 1981.

LOGICA, COMPUTAZIONE E TEORIE QUANTISTICHE DELLA MENTE

di Giulia Battilotti

Il fallimento dell'intelligenza artificiale ha evidenziato un fatto: i processi logici fino a ora considerati sono insufficienti, se non inadatti, a rendere conto dei processi mentali umani. Come la ricerca sui neuroni specchio ha recentemente evidenziato, la comprensione del significato dell'azione altrui non avviene in modo *procedurale*, vale a dire confrontando l'informazione appena acquisita con quelle già presenti in un presunto database mentale. Questa era la modalità con cui si pretendeva di ricreare computazionalmente la comprensione ed erroneamente, di spiegarla. Secondo alcune interpretazioni, la ricerca sui neuroni specchio ha mostrato come, al contrario, la comprensione del significato si ottenga rivivendo l'azione come se fosse propria (Rizzolatti G., Sinigaglia C., 2006).

Ciò pone la logica di fronte a una difficile sfida: studiare e proporre legami logici alternativi, che possano riprodurre i nostri processi mentali in modo più attendibile, "imitando la natura". Per questo si fa riferimento al ruolo di specifici processi logici e non in modo generico a processi computazionali, supponendo che la nostra mente elabori, in primis, informazioni che essa ritiene vere, e che entrano a far parte della nostra vita personale. Queste possono essere informazioni acquisite dai nostri sensi, comunicate da qualcun altro che riteniamo attendibile, o raggiunte per via simbolica. In altre parole la nostra mente compone asserzioni (giudizi), cioè dati di informazione considerati non semplicemente in quanto tali, ma in quanto veri. Si noti che, componendo asserzioni, essa elabora ulteriormente ipotesi riguardo ciò che potrebbe essere vero, ciò che gli altri pensano sia vero, ciò che si ipotizza sia vero. La mente arriva dunque ad elaborare il concetto di proposizione, a cui si può attribuire un valore di verità, ma che non è necessariamente vero. Sfruttando un elemento di cui si tratta nel seguito, si può dire che la proposizione rappresenta "la presa di coscienza" dell'asserzione. Il livello-base è quello assertivo.

Una teoria che spieghi i connettivi logici a partire dall'analisi di processi messi in atto dal cervello permetterebbe di approdare a una comprensione "naturalistica" dei processi posti alla base della composizione tra asserzioni. Tale teoria risulterebbe, inoltre, affrontabile dal punto di vista del puro calcolo logico. Come è stato evidenziato dalla logica di base, attraverso il "principio di riflessione", risulta possibile giustificare i connettivi della logica,

assieme alle corrispondenti regole di calcolo, pensando i connettivi come provenienti da legami metalinguistici fra asserzioni (Sambin G, Battilotti, G., Faggian, 2000). Il problema cruciale è: a quali legami e processi mentali, finora inesplorati dal punto di vista computazionale, dovremmo riferirci? Una ulteriore domanda correlata è: perché mai tali legami e processi sarebbero rimasti oscuri fino a oggi?

Una nuova ipotesi sulla natura dei nostri processi mentali viene permessa dalle teorie quantistiche della mente. Secondo tali teorie, i processi fisici di natura quantistica (che sicuramente avvengono nel nostro cervello) concorrono a formare la nostra mente, cioè danno luogo a processi mentali. In particolare, la teoria quantistica della mente di Hameroff-Penrose (1995), colloca i processi quantistici che avvengono nel cervello nelle tubuline, proteine che formano i microtubuli, i quali sono a loro volta componenti delle cellule neurali. Le tubuline sono dei dimeri che possiedono due diversi stati che si trovano in sovrapposizione quantistica¹. Secondo la teoria di Hameroff-Penrose, i processi quantistici che avvengono nelle tubuline sono alla base della distinzione tra stati inconsci e stati consci. Mentre i primi coinciderebbero con gli stati di sovrapposizione quantistica, la coscienza coinciderebbe con il momento della decoerenza dello stato di sovrapposizione (o collasso della funzione d'onda). Ricordiamo che la sovrapposizione quantistica consiste nella compresenza di più stati diversi, ad esempio entrambi gli spin di una particella. La decoerenza è il momento in cui la sovrapposizione degli stati viene distrutta, collassando in uno solo fra gli stati inizialmente costituenti la sovrapposizione, con probabilità determinata dallo stato iniziale del sistema. Ciò può avere varie cause. Secondo Roger Penrose, nel caso della coscienza, la decoerenza non sarebbe dovuta a interventi dall'esterno del sistema, quale ad esempio è una misura dello stato quantistico, ma a un collasso spontaneo. In altre parole, il sistema inconscio ha una sua evoluzione spontanea verso la coscienza. Il collasso che determina la coscienza nella teoria di Penrose avverrebbe per motivi di gravitazione quantistica (Quantum gravitational threshold). La descrizione di questo è formalizzabile in termini matematici, ma, secondo Penrose, non traducibile in termini computazionali. Esisterebbe dunque un momento di non-computabilità intrinseco nel mondo fisico, che si ripercuoterebbe nel modello della mente da esso conseguente: il momento della coscienza sarebbe non computabile. Proprio per questo motivo, secondo Penrose, noi non potremo mai ottenere l'intelligenza artificiale.

Sembra invece che la teoria Hameroff-Penrose permetta di aumentare la nostra

¹ Per un survey sulle teorie quantistiche la mente vedi la voce "quantum approaches to consciousness", di Harald Atmanspacher, sulla Stanford Encyclopedia of Philosophy, <http://plato.stanford.edu/entries/qt-consciousness/>. Il materiale sulla teoria Hameroff-Penrose si trova nel sito web di Stuart Hameroff:

comprensione della mente dal punto di vista computazionale, fornendo i mezzi per affrontare le due domande poste sopra. In primo luogo essa obbliga a considerare le caratteristiche della computazione quantistica per descrivere “naturalisticamente” la nostra computazione mentale. Inoltre, se la computazione quantistica avviene esclusivamente a livello inconscio, non siamo consapevoli dei suoi effetti, al contrario della computazione legata alla fisica classica, di cui possiamo essere consapevoli. Ciò spiegherebbe perché una parte fondamentale dei legami fra asserzioni operata dalla mente ci sarebbe rimasta nascosta.

Come è noto, la computazione quantistica avviene sfruttando i due legami caratteristici degli stati quantistici, vale a dire la sovrapposizione e l'entanglement. La sovrapposizione consiste nella presenza di più stati diversi, ad esempio entrambi gli spin di una particella. Essa permette il parallelismo nei processi di calcolo. Il legame di entanglement si crea quando, in un sistema di due o più particelle sovrapposte, gli stati non sono separati, vale a dire non si possono descrivere come prodotto degli stati delle singole particelle. Ciò comporta che, nel momento del collasso della funzione d'onda, i risultati ottenibili per le singole particelle componenti il sistema non sono indipendenti, in termini di indipendenza statistica. L'esempio più importante è quello rappresentato dagli stati di Bell, relativi a coppie di particelle che si comportano come “gemelle”, cioè collassano nello stato non sovrapposto con identici risultati. Essi rafforzano l'effetto del parallelismo dato dalla sovrapposizione, limitando la complessità e permettendo la velocizzazione del calcolo rispetto a quello ottenibile in un ambiente classico. Gli effetti della sovrapposizione e dell'entanglement non sono riproducibili in ambiente non quantistico. Si noti che questo sarebbe un forte motivo a favore dell'ipotesi di una teoria quantistica della mente: infatti, perché mai la natura non dovrebbe sfruttare le potenzialità uniche dell'entanglement anche per la nostra mente?

Al di là della velocità computazionale, il legame di entanglement offre anche altri vantaggi nel trattamento dell'informazione. Esso crea un legame fra informazioni di tipo olistico, in cui il tutto non è equivalente alla somma delle sue parti. Ciò può essere letto in due modi: la somma delle parti non è sufficiente ad ottenere il tutto e/o la somma delle parti non è necessaria ad ottenere il tutto. Come nella comprensione di una frase, sappiamo che conoscere il significato delle singole parole (ammesso che questo sia possibile), non è sufficiente a comprendere il significato della frase, ma anche, al contrario, sappiamo che talvolta è possibile comprendere una frase senza conoscere qualcuna o molte delle parole

che la compongono. In quest'ultimo senso il legame olistico è vantaggioso per la nostra mente, perché consente una notevole economia nel trattamento dell'informazione, quando "il tutto" sia conoscibile (o ipotizzabile) per qualche via.

Come è chiaro dall'esempio del linguaggio, la mente sa padroneggiare legami olistici. Al contrario, la logica classica sembra non poterlo fare, in quanto i connettivi logici sono definiti, alla maniera delle operazioni algebriche, in modo compositivo. Effettivamente "connettivo binario" significa che si connettono due parti e che il risultato della connessione definisce completamente una nuova entità. Ma, se si guarda al caso della logica predicativa, le cose non stanno esattamente così. Ad esempio, consideriamo la proposizione quantificata universalmente "tutti i figli concepiti da una coppia di portatori sani di talassemia hanno il 25% di probabilità di essere malati". Per comprenderla, occorre pensare a una coppia generica di portatori e ragionare sulle possibili combinazioni genetiche. In questo caso, la "colla" che tiene insieme questa entità è il generico elemento variabile (in questo caso sulle coppie di portatori). Si noti che questo vale anche se gli insiemi considerati (le coppie di portatori e i loro figli) sono finiti. Considerare a una a una le effettive coppie di portatori di talassemia esistenti e contare i figli malati dà un altro significato alla proposizione (non consente di affermare una legge genetica che preveda i casi futuri, benché in origine tale legge sia stata ipotizzata contando i casi esistenti). Inoltre, la prima soluzione si rende necessaria per interpretare una proposizione quantificata universalmente su un dominio infinito. Dunque la variabile rende possibile una nuova entità logica senza necessariamente dover comporre delle parti. Per questo nella logica predicativa è possibile trovare spazio per esprimere il legame di entanglement. Un ulteriore argomento a favore della variabile per esprimere il legame di entanglement può essere formulato sulla base dell'identificazione fra processi quantistici della mente e processi inconsci, così come ipotizzato dalla teoria Hameroff-Penrose. Infatti, la variabile diventa interamente parte del nostro linguaggio-oggetto solo attraverso il raggiungimento della piena maturità e coscienza di sé. Possiamo pensare che i processi logici di cui siamo consapevoli, vale a dire quelli descritti dal calcolo dei predicati classico, interpretino quest'ultima idea di variabile.

Accanto a questa dovrebbe esistere una *gestione-bambina* in cui la variabile viene sfruttata nel processo di calcolo ma non oggettivizzata da chi la sfrutta. Ad esempio, un bambino abbastanza grande è in grado di capire il significato di una regola scolastica del tipo "l'ultimo che esce dall'aula chiude la porta" (che, come ogni regola, contiene una variabile), anche se non è in grado di comprendere il senso di un codice giuridico, né di

una teoria matematica che contiene variabili. Un bambino più piccolo non è neppure in grado di comprendere la regola. Sfruttando una specie di passaggio al limite, si potrebbe dire che in questo caso l'uso della variabile nei processi mentali è ancora più interiorizzato, tanto da non far neppure sospettare a chi la usa la possibilità di usarla. Inoltre appare plausibile che l'interiorizzazione della variabile sia tanto maggiore quanto più si è lontani dalla coscienza di sé. Quale sarebbe il vantaggio di tutto ciò? Un enorme vantaggio computazionale e cognitivo. Il bambino è un “pessimo” logico, soprattutto nella prima infanzia, ma ci batte nettamente nelle possibilità cognitive. Basti pensare che, in genere, il bambino apprende la lingua madre entro i tre anni, prima della separazione dalla madre, che l'apprendimento della lingua madre deve avvenire nella prima infanzia, altrimenti non è più possibile e che l'apprendimento di una seconda lingua, al pari della lingua madre, avviene solo entro il termine dell'infanzia.

Tutto ciò non impedisce di provare a descrivere che tipo di processi logici avvengano inconsciamente. In questo caso, la dinamica fra il legame metalinguistico e il linguaggio oggetto sarà diversa e darà luogo a regole di calcolo diverse rispetto all'usuale calcolo dei predicati. Fidandoci della teoria Hameroff-Penrose, dovremmo trovare un calcolo adatto a modellare la computazione quantistica, e, viceversa, potremmo provare ad assumere le caratteristiche delle logiche per la computazione quantistica per provare a scoprire come costruiamo le nostre verità inconse.

Fin dal loro esordio, è stato chiaro che le logiche quantistiche avevano caratteristiche peculiari che le distinguevano nettamente dalla logica classica e dalle altre logiche note. Le logiche ultimamente studiate in specifico per la computazione quantistica rafforzano questa distinzione. Lo scenario prefigurato è quello di un ambiente logico paraconsistente e lineare. In una logica paraconsistente non è valida la legge di non contraddizione, senza che per questo sia valida qualsiasi proposizione. In una logica lineare, i connettivi seguono due tipi di regole: regole “moltiplicative” e regole “additive”. Esse danno luogo a due tipi di congiunzione, moltiplicativa ed additiva, e analogamente a due tipi di disgiunzione. In logica lineare, i connettivi pensati per modellare processi paralleli di calcolo sono quelli moltiplicativi. Il parallelismo quantistico, dato dalla sovrapposizione, si può invece meglio interpretare attraverso gli additivi.

La logica lineare è stata pensata per essere computazionalmente conveniente, attraverso un'idea “geometrica” della derivazione che elimini il fattore dell'ordine sintattico delle regole. I quantificatori in logica lineare hanno un carattere esclusivamente additivo e non rientrano appieno nell'idea “geometrica” della derivazione nel caso si abbiano più

proposizioni che dipendono dalla stessa variabile. Ciò è legato ad un vincolo sulla variabile che garantisce la consistenza delle regole di calcolo, ma in ambiente paraconsistente è possibile indebolire, o togliere, il vincolo e gestire le regole del calcolo predicativo più liberamente. Di particolare interesse risulta un connettivo predicativo dal carattere moltiplicativo-additivo, che si forma in presenza della dipendenza da una variabile comune a più proposizioni. Esso “riflette”, nel senso del principio di riflessione del metalinguaggio nel linguaggio oggetto che ricordavo inizialmente, la presenza della comune dipendenza, e si può interpretare come nuovo legame fra proposizioni quantificate. Il vantaggio computazionale di questa soluzione consiste nel fatto che essa riesce a gestire il caso di dipendenza da una o più variabili in comune a più proposizioni, in modo geometrico e non legato alla sintassi. La stessa soluzione, in ambiente di calcolo predicativo consistente è possibile solo nel caso di variabili distinte, che vengono sostituite in modo indipendente con elementi del dominio, ottenendo un aumento esponenziale dei casi possibili .

In tale contesto di calcolo, se si interpretano gli stati sovrapposti attraverso formule quantificate universalmente, in cui il dominio delle variabili è dato dall'insieme degli stati della base, il legame permesso dalla variabile comune riesce a interpretare gli stati di Bell. Riprodurre gli stati di Bell in logica ha un prezzo: in presenza di un legame dato da una variabile in comune non si ottiene l'implicazione logica, in quanto essa richiede di poter separare la singola informazione da un contesto, qualunque esso sia, mentre la variabile in comune non dovrebbe permettere, in questo modello, di rompere il legame. In altre parole, il legame olistico è vincolante e alternativo al legame di conseguenza logica. La nostra mente allora potrebbe procedere in almeno due modalità: una a carattere olistico, computazionalmente vantaggiosa, attraverso libere associazioni inconsce di variabili, e una a carattere “metodologico”, che rinuncia all'associazione ottenuta per mezzo di variabili comuni e dà spazio alla conseguenza logica. Quest'ultima è una modalità di cui siamo consapevoli e che evita la contraddizione. Le due modalità sarebbero dovute al trattamento quantistico e classico dell'informazione, e sarebbero legate a diverse tipologie nella coscienza di sé. È chiaro che le due modalità sono intrecciate nella nostra vita e danno quindi luogo a diverse tipologie miste di pensiero, oltre che a scambi di ruolo con esiti dis-logici, che pure osserviamo.

Abbiamo una conferma sperimentale nella ricerca psicologica riguardo all'ipotesi che la computazione operata a livello inconscio sia più efficace. Nelle ricerche sui processi decisionali condotte per ricerche di mercato, si è appurato che la scelta del prodotto migliore viene fatta con processi inconsci nel caso vi sia una quantità elevata di variabili da

considerare per la scelta, mentre viene operata dal ragionamento cosciente nel caso le variabili siano poche. Questa è una conferma dell'ipotesi che sia il diverso trattamento della variabile a dare un vantaggio computazionale all'inconscio².

Giulia Battilotti

² L'unica teorizzazione di come sia strutturata la logica dell'inconscio, formulata in modo organico, ci è stata lasciata dallo psicoanalista cileno I. Matte Blanco, basandosi su una lunga esperienza clinica. Sarà necessario considerare le caratteristiche sottolineate da Matte Blanco nell'ottica delle regole di un sistema olistico.

Appendice

Calcolo dei sequenti e parallelismo quantistico

Cos'è un sequente e un calcolo dei sequenti

Un sequente è una scrittura del tipo $\Gamma \vdash \Delta$, dove $\Gamma = C_1, \dots, C_n$ e $\Delta = D_1, \dots, D_m$ sono liste finite di formule, fra loro separate da virgole, e dove \vdash è un segno che indica conseguenza. Per questo Γ e Δ si dicono premesse e conclusioni del sequente, rispettivamente. Un calcolo dei sequenti è un insieme di regole fissate, che trasformano sequenti in altri sequenti. Possono trasformare un sequente in un altro oppure una coppia di sequenti in un terzo, e vengono scritte nella seguente forma:

$$\frac{\Gamma_1 \vdash \Delta_1}{\Gamma \vdash \Delta} \quad \frac{\Gamma_1 \vdash \Delta_1 \quad \Gamma_2 \vdash \Delta_2}{\Gamma \vdash \Delta}$$

Gli assiomi del calcolo dei sequenti sono i sequenti della forma $A \vdash A$. Derivare il sequente $\Gamma \vdash \Delta$, in un certo sistema di calcolo, significa ottenerlo come conclusione di un'opportuna applicazione delle sue regole a partire da assiomi.

Parallelismo nel calcolo dei sequenti

Secondo la logica di base, i connettivi logici, insieme alle corrispondenti regole nel calcolo dei sequenti, sono il risultato di una traduzione di certi legami metalinguistici fra asserzioni nel linguaggio oggetto. Le asserzioni sono rappresentate da sequenti. La traduzione nel linguaggio oggetto si ottiene ponendo equazioni definitorie, la cui soluzione restituisce le regole per i connettivi corrispondenti.

Consideriamo tre legami, *e* e *comporta* per la logica proposizionale, *per ogni* per la logica predicativa.

Il legame metalinguistico *comporta* rappresenta la conseguenza fra due giudizi logici, cioè collega due asserzioni in modo *sequenziale*. Dà origine all'implicazione, che si ottiene risolvendo la seguente equazione definitoria:

$$\Gamma \vdash A \rightarrow B \quad \equiv \quad \Gamma, A \vdash B$$

Il legame metalinguistico *e* unisce due giudizi logici allo stesso livello, vale a dire considerati in modo *parallelo*. *e* si rappresenta attraverso una coppia di asserzioni (sequenti) che dipendono entrambe dalle stesse premesse nel caso additivo, e attraverso la virgola nel sequente nel caso moltiplicativo. Ciò produce i connettivi additivi e moltiplicativi della logica lineare. A destra del segno di sequente, essi sono definiti dalle equazioni:

$$\begin{aligned} \Gamma \vdash A \&B & \equiv \Gamma \vdash A \quad \Gamma \vdash B \\ \Gamma \vdash A \cdot B & \equiv \Gamma \vdash A, B \end{aligned}$$

(denotiamo la disgiunzione moltiplicativa con \cdot). Analoghe equazioni definiscono la disgiunzione additiva \oplus e la congiunzione moltiplicativa \otimes a sinistra. In seguito, ci si limita ai connettivi definiti a destra (connettivi destri) dato che gli altri sono analoghi. Così si ottengono due modi di implementare i processi paralleli nel calcolo dei sequenti: additivo, per mezzo della congiunzione $\&$, e moltiplicativo, per mezzo della disgiunzione \cdot . In logica lineare classica, vale la distributiva del connettivo moltiplicativo \cdot rispetto al connettivo additivo $\&$, cioè della disgiunzione rispetto alla congiunzione. Ciò significa poter trovare un connettivo composto moltiplicativo-additivo, dato che il risultato non dipende dalla sintassi.

Supponiamo ora che le nostre asserzioni dipendano da una variabile. Il quantificatore \forall si ottiene ponendo e risolvendo la seguente equazione definitoria:

$$\Gamma \vdash (\forall x \in D)A(x) \quad \equiv \quad \text{per ogni } z \in D \Gamma \vdash A(z)$$

che, più formalmente, diviene:

$$\Gamma \vdash (\forall x \in D)A(x) \quad \equiv \quad \Gamma, z \in D \vdash A(z)$$

In entrambi i casi, z non è libera in Γ , cioè Γ non dipende da z . Ciò significa che l'equazione ha un carattere additivo. Si ottiene $\&$ come caso particolare, quando D ha due elementi. Tuttavia, in tal caso, la caratteristica fondamentale data dalla variabile viene perduta.

Sovrapposizione quantistica nel calcolo dei sequenti

In genere, nei modelli di processi paralleli si adottano i connettivi moltiplicativi; qui proponiamo che l'additività rappresenti la sovrapposizione quantistica, responsabile del parallelismo computazionale quantistico.

Consideriamo un sistema quantistico \mathcal{A} rappresentato in uno spazio di Hilbert H con base ortonormale \mathbf{B} (finita o numerabile). Come è noto, una misura quantistica su un certo sistema produce come risultato un elemento della base con una certa probabilità.

Una misura quantistica su \mathcal{A} è un esperimento, cioè, in teoria della probabilità, una coppia (X, p) , dove X è una variabile casuale, a valori in \mathbf{B} , e $p : \mathbf{B} \rightarrow [0, 1]$ una funzione di probabilità. Nel nostro caso, p_a dove a è una ampiezza di probabilità a valori complessi. Dato che \mathbf{B} è numerabile, X e p_a sono discrete. Consideriamo l'insieme $D = D(X, p_a)$ dei possibili risultati di un esperimento con le relative probabilità: $D = \{(x, p_a(X = x)), x \in \mathbf{B}\}$. Ciò significa che, *per ogni* $z \in D$, dove z è la coppia $(z_1, p_a(X = z_1))$, si trova che "la misura quantistica su \mathcal{A} produce z_1 con probabilità $p_a(X = z_1)$ ". Se denotiamo con $A(z)$ la precedente proposizione fra virgolette, e con Γ le ulteriori ipotesi sulla misura (che non dipendono dal suo risultato $z!$), possiamo scrivere formalmente tale *per ogni* metalinguistico, ottenendo l'asserzione:

$$\Gamma, z \in D \vdash A(z)$$

dove Γ non contiene z libera. Allora, considerando l'equazione definitoria di \forall

$$\Gamma \vdash (\forall x \in D)A(x/z) \quad \equiv \quad \Gamma, z \in D \vdash A(z)$$

la proposizione $(\forall x \in D)A(x)$ rappresenta la sovrapposizione quantistica del sistema \mathcal{A} .

Esempio: se abbiamo una particella rappresentata in C^2 con base ortonormale $|0\rangle, |1\rangle$, e l'esperimento su di essa è la variabile casuale con probabilità $p_a(X = |0\rangle) = 1/2$ e $p_a(X = |1\rangle) = 1/2$, data da $a : \{|0\rangle, |1\rangle\} \rightarrow C^2$, dove $a(|0\rangle) = a(|1\rangle) = 1/\sqrt{2}$, la proposizione $(\forall x \in D)A(x)$ rappresenta lo stato $1/\sqrt{2}|0\rangle + 1/\sqrt{2}|1\rangle$.

Parallelismo quantistico nel calcolo dei sequenti

La legge distributiva compone il parallelismo additivo con quello moltiplicativo. Nel caso predicativo si scrive:

$$(\forall x \in D)(\forall x' \in D')(A(x) \cdot B(x')) \vdash (\forall x \in D)A(x) \cdot (\forall x' \in D')B(x')$$

e dunque richiede domini eventualmente diversi e variabili indipendenti che variano su di essi. Se le proposizioni $(\forall x \in D)A(x)$ e $(\forall x' \in D')B(x')$ rappresentano due particelle, domini differenti significano variabili aleatorie diverse, con risultati indipendenti. Ciò significa stati separati per le due particelle.

Se il dominio è lo stesso e viene considerata la stessa variabile, abbiamo la stessa variabile aleatoria con risultati dipendenti per gli esperimenti sulle due particelle. Ciò implica che le particelle sono nello stato entangled. In tal caso, la legge distributiva, pensata come indipendenza della composizione moltiplicativa-additiva dalla sintassi, dovrebbe avere questo aspetto:

$$(\forall x \in D)(A(x) \cdot B(x)) \vdash (\forall x \in D)A(x) \cdot (\forall x \in D)B(x)$$

Si noti che, considerando due particelle \mathcal{A} e \mathcal{B} entrambe come nell'esempio precedente, il nostro oggetto moltiplicativo-additivo, definibile grazie all'indipendenza dalla sintassi, è uno stato di Bell. Infatti, \mathcal{A} e \mathcal{B} sono entrambe nello stato $1/\sqrt{2}|0\rangle + 1/\sqrt{2}|1\rangle$, ma i risultati di ciascun esperimento su una di esse sono lo stesso per l'altra, dato che stiamo usando la stessa variabile. Allora chiamiamo il sequente dato "distributiva di Bell".

Il problema è che il sequente è falso, interpretando \cdot come semplice disgiunzione, dunque il calcolo dei sequenti non lo dimostra. Si potrebbe dimostrarlo con la seguente regola inconsistente, che introduce \forall "in parallelo", e che è impossibile, data la restrizione sulle variabili imposta nella regola del quantificatore \forall :

$$\frac{\Gamma, z \in D \vdash A(z), B(z)}{\Gamma \vdash (\forall x \in D)A(x), (\forall x \in D)B(x)} \quad \forall f \parallel$$

Si possono superare queste difficoltà in un calcolo paraconsistente originato dalla logica di base, che delineiamo nel seguito.

Come abbiamo notato, il problema è la disgiunzione \cdot , originata dalla virgola. Nell'asserzione $\Gamma \vdash A(z), B(z)$, le formule $A(z)$ e $B(z)$, essendo dipendenti dalla stessa variabile, sono più incollate di due formule A a B senza variabili comuni, come, per esempio, due fratelli sono più legati di due persone in genere. In altre parole, il legame moltiplicativo “,” è in generale più debole della virgola in $\Gamma \vdash A(z), B(z)$. Allora scriviamo la virgola come z -virgola $,_z$: $\Gamma \vdash A(z),_z B(z)$. Dato che $,_z$ è un forte collante, le formule legate da $,_z$ non si possono separare l'una dall'altra portando una di esse (negata) dall'altra parte del sequente, e non sono contesto una dell'altra. Ogni azione su una di esse si deve fare anche sull'altra, “in parallelo”.

Consideriamo l'asserzione $\Gamma, z \in D \vdash A(z),_z B(z)$, dove Γ non dipende da z . Per le ultime osservazioni, la condizione su Γ non è alterata dalla presenza di due formule in cui appare z a destra. Dunque si può applicare la regola di formazione di \forall , sia a $A(z)$ che a $B(z)$, “in parallelo”, al seguente modo:

$$\frac{\Gamma, z \in D \vdash A(z),_z B(z)}{\Gamma \vdash (\forall x \in D)A(x),_z (\forall x \in D)B(x)} \forall f \parallel$$

Notare che la \forall -regola converte il legame *per ogni* e non può alterare il legame $,_z$, che dunque è ancora presente nella conclusione, anche se lega due formule in cui z non compare più libera. Concludiamo che $,_z$ è un legame dinamico che unisce formule in cui z compariva o compare libera in comune. Ciò che elimina $,_z$ è la sostituzione di z con un termine chiuso t , dato che la variabile non è più né presente come libera né “quantificata in comune”, ma scompare. In tal caso, $\Gamma \vdash A(z),_z B(z)$ diventa $\Gamma \vdash A(z/t), B(z/t)$.

Come ogni legame, $,_z$ si traduce in un connettivo, denotiamolo con \bowtie . Esso si ottiene risolvendo l'equazione

$$\Gamma \vdash A \bowtie B \quad \equiv \quad \Gamma \vdash A,_z B$$

come in logica di base, data una opportuna regola di taglio per risolverla. Opportuno significa che, dove un certo legame, sia esso “,” oppure “ $,_z$ ”, è presente nelle premesse, viene conservato dalla sostituzione della formula tagliata nella conclusione (se nella formula/nelle formule sostituita/e alla formula tagliata non c'è dipendenza da z , l'eventuale $,_z$ va letta come semplice virgola).

Per esempio, assumendo l'assioma di riflessione di \bowtie , cioè il sequente $A \bowtie B \vdash A,_z B$, si deriva la regola di \bowtie a sinistra (regola di riflessione in logica di base) come segue:

$$\frac{\frac{A \bowtie B \vdash A,_z B \quad \Gamma_1, A \vdash \Delta_1}{\Gamma_1, A \bowtie B \vdash \Delta_{1,z} B} \text{ cut} \quad \Gamma_2, B \vdash \Delta_2}{\Gamma_2, \Gamma_1, A \bowtie B \vdash \Delta_{1,z} \Delta_2} \text{ cut}.$$

Per ulteriori spiegazioni si rimanda alla logica di base.

Grazie alle regole di \bowtie ed alla regola $\forall ||$ si dimostra la legge distributiva di Bell nella forma:

$$(\forall x \in D)(A(x) \bowtie B(x)) \vdash (\forall x \in D)A(x) \bowtie (\forall x \in D)B(x)$$

che restituisce l'oggetto moltiplicativo-additivo per l'entanglement, rappresentato indifferentemente da $(\forall x \in D)(A(x) \bowtie B(x))$ o da $(\forall x \in D)A(x) \bowtie (\forall x \in D)B(x)$. Le due scritture indicano i due modi di ottenerlo.

Qualche commento a carattere logico e fisico:

Il legame $,_z$ così introdotto produce un trattamento olistico dell'informazione, proprio dell'entanglement, poiché esso riesce a creare un tutto in cui non si può distinguere un contesto e separarne una parte. La corretta equazione per l'implicazione logica, vale a dire $\Gamma \vdash A \rightarrow B \equiv \Gamma, A \vdash B$, non vale con $,_z$, cioè nella forma $\Gamma \vdash A \rightarrow B \equiv \Gamma,{}_z A \vdash B$. Nel calcolo dei sequenti si ottengono così due modi: uno dovuto alla virgola, in ambiente consistente, che dà spazio all'implicazione e ferma l'entanglement; l'altro dovuto alla z -virgola, in ambiente paraconsistente, che ha effetto opposto. I due ambienti logici hanno vite separate (da un punto di vista formale): in ambiente paraconsistente non bisogna preoccuparsi della contraddizione, dato che non ci può essere (nel linguaggio degli studiosi di logica paraconsistente) "esplosione" del falso, dovuta alla regola di modus ponens dell'implicazione; in ambiente consistente l'entanglement, che dimostrerebbe il falso, non esiste.

Tutto ciò concorda molto bene da una parte con la difficile interpretazione della causalità (legata all'implicazione) in fisica quantistica, mentre dall'altra parte l'entanglement non esiste in fisica classica. La sostituzione di una variabile con un termine chiuso avrebbe lo stesso effetto del collasso della funzione d'onda.

Ringraziamenti: Ringrazio Stuart Hameroff per avermi aiutato nella comprensione della sua teoria della coscienza quantistica. Eventuali errori o fraintendimenti, nella breve esposizione che ne ho dato qui, sono solo imputabili alla mia imprecisione. Le opinioni che esprimo sulle conseguenze della teoria sono mie personali. Stuart Hameroff ha sempre sottolineato l'importanza di una comprensione da un punto di vista logico della materia, e suggerisce inoltre di considerare l'opera di Matte Blanco come testimone delle caratteristiche logiche dell'inconscio. I miei ringraziamenti vanno inoltre a Silvano Zipoli Caiani per i suoi accurati ed utili suggerimenti sulla prima stesura di questo scritto.

Bibliografia

- Battilotti, G., Faggian, C. , Quantum logic and the cube of logics, in Handbook of Philosophical Logic, new edition, capitolo “Quantum Logic” by M.L. Dalla Chiara e R. Giuntini, vol. 6, D. Gabbay and F. Guenther eds., Kluwer, 2002.
- Battilotti, G., Basic logic and quantum computing: logical judgements by an insider observer, Proc. FQI04, 16-19 april 2004, Camerino, Italy, in *International Journal of Quantum Information* 2005, 3,1, 105-109 - arXiv:quant-ph/0407057.
- Battilotti, G., Hameroff, S., Zizzi, P., From Quantum Mind to A.I., abstract, convegno “Computers and Philosophy, Laval, France, maggio 2006
- Battilotti, G., L'implicazione logica e il legame di attaccamento, draft.
- Blanco, M. The unconscious as innate sets, Duckworth, London, 1975.
- Dalla Chiara, M. L., Giuntini, R., Paraconsistent Quantum Logic, *Foundations of Physics* 19, 1989, 891-904.
- Dalla Chiara M.L., Giuntini R., Leporini R., Quantum Computational Logics: a survey. in V. F. Hendricks, J. Malinowski (eds.), Trends in Logic: 50 Years of Studia Logica, Kluwer Academic Publishers, Dordrecht, 2003 pp. 213-255.
- Dalla Chiara M.L., Giuntini R., Leporini R., Compositional and Holistic Quantum Computational Semantics, *Natural Computing*, 6, 2, 2007, 113-132.
- Dijksterhuis, Ap, On making the right choice: the deliberation-without attention effect , *Science*, 311, 2006, 1005-1007.
- Girard, J. Y., Linear Logic, *Theoretical Computer Science* 50, 1987, 1-102.
- Hameroff, S., Consciousness, neurobiology and quantum mechanics: The case for a connection, in The Emerging Physics of Consciousness, edited by Jack Tuszynski, Springer-Verlag, 2006.
- Hirvensalo M., Quantum Computing, Springer, 2004.
- Maietti, M.E., Sambin, G., Toward a minimalist foundation for constructive mathematics, in “From Sets and Types to Topology and Analysis: Practicable foundations for constructive mathematics”, L. Crosilla and P. Schuster eds, Oxford University Press, 2005.
- Matte Blanco, I., The unconscious as infinite sets, Duckworth, London, 1975.
- Penrose, R., The emperor's new mind, Oxford University Press, 1989.
- Penrose R., Hameroff S., Consciousness as orchestrated space-time selections, in *Journal of Consciousness Studies*, 3(1), 1999, 36-53.
- Rizzolatti G., Sinigaglia C., So quel che fai, Milano, Cortina, 2006.
- Sambin G., Battilotti, G., Faggian, C., Basic logic: reflection, symmetry, visibility, in *The Journal of Symbolic Logic*, 65 (3), 2000, 979-1013.

ALCUNE CONSIDERAZIONI SUI LIMITI SEMANTICI DELLA FISILOGIA. GLI STUDI DI BENJAMIN LIBET SUL LIBERO ARBITRIO

di Laura Beritelli

1. Introduzione. Questioni di metodo

Nel saggio *Il normale e il patologico*¹, Georges Canguilhem polemizzava così rispetto al variopinto sottobosco concettuale di assunzioni soggiacente la medicina moderna - che egli infatti considerava una pratica, piuttosto che una vera e propria scienza:

siccome gli scienziati conducono la propria esistenza di uomini entro un ambiente e una comunità non esclusivamente scientifici, la storia delle scienze non può ignorare la storia delle idee. Applicando ad una tesi la sua conclusione, si potrebbe affermare che le deformazioni che essa subisce nell'ambiente della cultura possono rivelarne il significato essenziale².

Questa riflessione, che sposa l'orientamento più proprio della cultura europea continentale, invitava allora ad uno studio sinottico dei fenomeni legati alla crescita della tecnica medica, introducendo al contempo una questione epistemologica fondamentale che, in quel caso, aveva il suo perché nella relazione essenziale tra fisiologia e patologia. Nella nostra contemporaneità, la rimozione di quel rapporto fondativo della nozione di normalità ha reso oscura la percezione della malattia e sempre più sfocata l'idea di salute; questo accade perché la ricerca ha impostato straordinari cambiamenti che, nonostante abbiano di fatto rivoluzionato il corpo³, sono rimasti impensati. Al tentativo di emancipare il concetto di salute da quello di malattia, ovvero a quello di liberarsi da una definizione della prima come negazione della seconda, ha contribuito lo stesso Oms,⁴ ottenendo che la parola 'salute' venisse declinata qualitativamente, come 'benessere'; affermare un simile modello sanitario ha significato, al contempo, riconoscere e incentivare una visione

¹ C. Canguilhem, *Il normale e il patologico*, Einaudi, Torino, 1998. Il testo, pubblicato nel 1943, è stato riproposto nel 1966 senza sostanziali modifiche. Nella prefazione tuttavia Canguilhem spiega come all'epoca non avesse letto M. Pradines e M. Merleau-Ponty, che invece ora avrebbe citati.

² C. Canguilhem, *op cit*, p 21.

³ Vedi JLNancy, *L'intruso*, Cronopio, Napoli, 2000. Il discorso è tuttavia ben più ampio, ad esempio, ciò che è cambiata è anche l'idea del rapporto del corpo con l'ambiente (oltre che del corpo con la percezione di sé e della propria vita mentale).

⁴ Organizzazione Mondiale della Sanità (agenzia ONU dal 1948). La definizione di salute dell'Oms è per l'appunto 'stato di completo benessere fisico, psichico e sociale e non semplice assenza di malattia'.

in cui l'attenzione alla condizione psichica degli individui assumeva sempre maggiore rilevanza. L'inarrestabile affermarsi, nella stessa pratica medica, di un approccio 'olistico', che intende la vita come complessità psicofisica e dinamica, ci parla oggi del fallimento economico delle varie forme di riduzionismo e rende impensabile riproporre soluzioni di tipo dualista, se non ancora eliminativista, al problema mente-corpo.

Nel tentativo di impostare un nuovo orizzonte di senso che renda ragione del progresso tecnico e scientifico, diviene difficile ignorare i solleciti richiami⁵ all'interdisciplinarietà e all'interazione tra le diverse prospettive teoriche, 'naturali' e 'umane' che dir si voglia; a fondamento di questa richiesta sta l'istanza a costruire (in un movimento continuo di decostruzione e affermazione) un'interpretazione condivisa e appoggiata dal maggior numero di saperi specifici, capace di definire i contorni della nostra cultura materiale e, soprattutto, di tramandarne l'*episteme*⁶. La differenziazione nelle più diverse discipline è stata un processo necessario, così come la specializzazione, attraverso la quale tutte loro hanno ottenuto maggiore successo nei rispettivi campi d'indagine; siamo tuttavia andati incontro ad una frammentazione che, nonostante non debba coincidere *eo ipso* con una deriva di senso e pur non inficiando direttamente i singoli ambiti di ricerca, mantiene in sospenso, come problematica, la possibilità di un discorso condiviso.

La particolare situazione della conoscenza e dell'informazione ha dato voce e spazio ad indagini che hanno talvolta offerto panorami semantici paradossali⁷ e prodotto enunciati altrettanto improbabili, che hanno ben poco a che fare con i presupposti della buona scienza.⁸ C'è, in effetti, una tendenza dispersiva che fa talvolta compiere sforzi sperimentali volti ad ottenere risposte non significative.⁹ Tuttavia, come illustreremo di

⁵ Vedi A. Binazzi, *Conferenza di George Lakoff*, e D. Manetti, S. Zipoli, *Intervista a George Lakoff*, 'Humana.mente' n°4, 2/2008.

⁶ Nell'accezione che diede a questo termine il filosofo Michel Foucault.

⁷ Un esempio di tale paradossalità potrebbero essere proprio gli esperimenti di Libet presi in esame: una volta decontestualizzati dagli studi medici, un profano non tarderebbe a chiedersi come sia possibile ottenere un'espressione di libertà 'riproducibile' e 'sotto controllo' dato che, assunti questi due ultimi requisiti, necessari all'analisi di laboratorio, sembra impossibile aver fatto salvo l'oggetto dell'analisi stessa.

⁸ Libet giunge infatti alla totale inconoscibilità riguardo all'esistenza o meno del libero arbitrio: la sua valutazione finale consisterà in un invito a credere quel che vogliamo.

⁹ Significativa l'istituzione del Premio Ig Nobel. Ogni anno vengono assegnati dieci premi ad altrettante ricerche scientifiche di dubbia costruttività nel corso di una cerimonia al Sanders Theatre di Harvard. La cerimonia è seguita a distanza di pochi giorni da lezioni informali al MIT, durante le quali alcuni laureati possono illustrare al grande pubblico i risultati e l'importanza delle proprie ricerche. I premi Ig Nobel 2007, ad esempio, sono: per la medicina, a Brian Witcombe di Gloucester, UK, e Dan Meyer di Antioch, USA, per uno studio sugli effetti collaterali di mangiare le spade; per la fisica, a Lakshminarayanan Mahadevan della Harvard University, USA, e Enrique Cerda Villablanca dell'Università di Santiago del Cile per aver studiato come le lenzuola si spiegazzano; per la biologia, a Johanna E. M. H. van Bronswijk dell'Università della Tecnologia di Eindhoven, Olanda, per un censimento di tutti gli organismi che vivono nel letto; per la chimica, a Mayu Yamamoto dell'International Medical Center del Giappone per la scoperta di un modo per estrarre vanillina, la fragranza e l'aroma della vaniglia, dallo sterco di mucca; infine, per la linguistica, a Juan Manuel Toro, Josep B.

seguito, non sono certo la singola scoperta scientifica o la ricerca che possono essere messe sotto accusa ma, in primis, la contraffazione che ne colpisce l'esposizione e, di conseguenza, la diffusione; quest'ultima, in particolare, troppo spesso beneficia di una libertà espressiva propria piuttosto dei fenomeni pubblicitari, la quale non solo non è loro consona, ma si rivela dannosa per la loro stessa credibilità. Il problema di un'attribuzione ulteriore di senso ai dati sperimentali (e quindi di interpretazioni più o meno ammissibili) ha spesso origini endogene alla letteratura scientifica. Per darsi ragione delle tesi peregrine di cui sentiamo parlare ogni qualvolta si nomina, ad esempio, la neuroscienza, non possiamo imputare ogni responsabilità all'analfabetismo scientifico; o, quantomeno, se si desidera che alle sue scoperte non vengano attribuiti significati che le oltrepassano e siano pertanto assimilate in modo più ortodosso dalla comunità, sembra più utile lavorare ad una epistemologia per le scienze e ad una teoria della comunicazione per la divulgazione scientifica, che soddisfino l'istanza di sintesi, interdisciplinarietà e narrazione comune, di cui ogni progresso ha bisogno. Una volta preso atto della necessità di una più sofisticata impostazione metodologica, che partecipi dell'esperienza cumulata dalle singole discipline, sembra sensato proporre uno sforzo di collaborazione, affinché si ottenga che le scienze esprimano correttamente l'evidenza sperimentale rinvenuta e, a partire da una maggiore chiarezza e diffusa consapevolezza, la ricerca sia portata a lavorare coerentemente sulle domande più urgenti che solleva ad ogni suo avanzamento. Una certa filosofia, o come la si preferisca chiamare, anch'essa vittima delle sue molteplici declinazioni e specializzazioni, è preposta al controllo logico e linguistico delle affermazioni della scienza, così come volle Wittgenstein; ma questo, da solo, non è sufficiente e non esaurisce il problema. Ad essa andrebbe ascritto, piuttosto, il ruolo di 'direzione dei lavori'. In particolare, per dialogare e collocare l'avanguardia scientifica, è necessario affiancare ad un lavoro di comprensione profonda uno di sintesi, che consenta di inquadrare i fatti da lontano o, per dirla con Canguilhem, leggerli in seno alla loro comunità legittimante (a maggior ragione, pertanto, è necessario che ad essa pervengano informazioni corrette). Anche se il controllo dei sottesi e delle certezze che sostengono le nostre credenze è solo una parte del lavoro critico della filosofia, esso contribuisce al continuo reinventarsi del sapere e, soprattutto, è ciò che rende il pensiero la traccia invisibile di ogni discorso.

Trobalon e Núria Sebastián-Gallés dell'Università di Barcellona per aver dimostrato che i ratti a volte non riescono a distinguere tra una persona che parla giapponese al contrario e una che parla olandese al contrario

2. Neurologia

Precursore e fautore di un'interessante modalità di dialogo tra neuroscienze e filosofia è stato Benjamin Libet.¹⁰

In *Do we have free will?*¹¹ egli spiega come, a seguito dei suoi esperimenti di cronometrando di processi cerebrali e volontà cosciente, sia giunto alla conclusione che il cervello dà luogo alla preparazione dell'atto volontario, che ha dunque origine inconscia¹². L'attivazione dell'RP¹³ cerebrale è infatti precedente l'effettivo movimento di circa un secondo (non meno di 550 msec prima dell'azione); il W¹⁴, ovvero l'emersione alla coscienza della volontà di compiere il movimento stesso, sopravviene in seguito (non meno di 250/300 msec prima dell'azione); alla fine del processo e qualora non venga esercitata l'azione cosciente di veto¹⁵ - cui Libet riduce, in sostanza, il libero arbitrio-, il muscolo compie il movimento volontario. L'assenza di un RP che precede il veto all'azione (la quale ha origine inconsapevole) convince Libet che esso ha invece un'origine esclusivamente cosciente e che, quindi, siamo liberi di scegliere.

Nonostante un investimento etologico dialetticamente limitato¹⁶, che analizzeremo in seguito, le conclusioni cui giunge Libet a proposito dell'esistenza del libero arbitrio hanno il pregio di mostrare come, da dati sperimentali ottenuti in laboratorio, si possa giungere a posizioni antiriduzioniste.

¹⁰ Benjamin Libet nasce a Chicago nel 1916. Si laurea in fisiologia nel 1939. Tra il 1945 e il 1948 è assistente presso la sua università. Nel 1956 collabora con K.A.C. Elliott e John Eccles (a Canberra, Australia) alle ricerche sui meccanismi delle sinapsi. Nel 1958 è nominato professore emerito al *Medical Center* e presso il *Center for Neuroscience* dell'Università della California (San Francisco), dove inizia i suoi studi sperimentali su soggetti umani volti a interpolare i dati dell'attività cerebrale con i resoconti verbali sull'esperienza cosciente. Ottiene il suo Ph.D. in neurofisiologia sotto la supervisione di Ralph Gerard, fondatore e presidente onorario della *Society of Neuroscience*. E' attualmente professore presso il Dipartimento di Fisiologia della stessa Università della California.

¹¹ B. Libet, *Do we have free will?*, in AA.VV., R.H.Kane, *The Oxford Handbook of free will*, Oxford University Press, NY, 2002, p 563.

¹² S'intende 'inconsapevole', ma è tradotto letteralmente come 'inconscio', al fine di sottolineare proprio le difficoltà comportate dall'uso di un linguaggio fenomenico in ambiti neurobiologici o neurologici.

¹³ *Readiness potential*, o *bereitschaftspotential* (Kornhuber e Deecke, 1965): potenziale di preparazione. Si tratta di una leggera variazione elettrica registrabile attraverso un EEG. Indica una certa attività neurale volta a predisporre e avviare un'azione.

¹⁴ *W time*: il tempo indicato dai soggetti come prima coscienza di voler agire. Per poter avere questo dato Libet inventò un dispositivo sperimentale che sembra un orologio a muro: invece che di lancette per le ore ed i minuti, un punto luminoso rosso indica i microsecondi; i soggetti riferivano verbalmente (e in seguito) la posizione esatta in cui si trovava il punto luminoso sul disco al momento in cui erano consapevoli di voler alzare un braccio, ad esempio, o compiere uno dei semplici gesti standard scelti da Libet affinché l'esperimento fosse controllabile e riproducibile. Inutile dire che la situazione costruita da Libet non ha molto a che vedere con quelle che figurava Kant per spiegare l'azione libera dal meccanismo o l'imperativo categorico! Tuttavia, se Kant potesse commentare il dato sperimentale, lo definirebbe coerente con le sue descrizioni della autonomia della volontà.

¹⁵ *Conscious veto*: in sostanza si tratta del controllo esercitato dalla mente cosciente ovvero della capacità di bloccare l'atto in potenza della cui urgenza siamo divenuti coscienti:

¹⁶ S'intenda qui l'etologia umana, quella branca della psicologia che si è ispirata agli studi dell'etologia animale (Konrad Lorenz) per dar ragione del disagio psichico (ma non solo); è una disciplina che studia i comportamenti umani prendendo in considerazione elementi sociologici, psicologico-cognitivi, biologici.

In assenza di prove a favore del determinismo o dell'indeterminismo,¹⁷ Libet ricorre al fatto fenomenico dell'esperienza soggettiva condivisa come prova *prima facie*¹⁸ che i processi mentali possano controllare causalmente alcuni processi cerebrali: si appella, cioè, al 'sentimento intuitivo del fenomeno della libertà della volontà',¹⁹ al fatto che noi tutti siamo intimamente convinti di avere il libero arbitrio: entro i limiti imposti dalle condizioni del nostro cervello e dall'ambiente, la libera scelta (*free will*) può essere considerata indipendente dal meccanismo. A difesa della sua tesi, Libet ci mette anche in guardia da ciò che 'si sente dire' riguardo a presunte scoperte che abbiano dimostrato il contrario, poiché non disponiamo di alcuna prova scientifica che possa avallare o negare il libero arbitrio.²⁰ La sua conclusione, in accordo con il neurobiologo Roger Sperry²¹, è quindi che l'esistenza del libero arbitrio sia una buona, se non ottima, opinione scientifica; e che, inoltre, dato che è su quel sentimento condiviso che ancora oggi si basa ogni interpretazione della natura umana, tale credenza possa essere demolita solo dalla più totale certezza scientifica.

Interrogando più a fondo l'assetto epistemologico da cui muovono le sue considerazioni scopriamo che esso si fonda nei precedenti studi sperimentali della psicologia dell'esperienza cosciente (1950). Il presupposto dei test condotti è dunque che i processi cerebrali (esternamente osservabili e manipolabili) e le dichiarazioni sulle relative esperienze introspettive (riferite dai soggetti) debbano essere studiati contemporaneamente, ma come categorie indipendenti; pertanto, gli esperimenti sono costruiti a partire da una tesi analogica debole. Il risultato più evidente dei suoi studi sarà dunque la mancanza di 'principi ponte' che possano collegare esaustivamente ed in maniera irrefutabile il discorso sulla coscienza con le teorie della fisica medica sul funzionamento del cervello.

La questione della possibilità di veto apre poi una questione, per risolvere la quale Libet deve appellarsi ai presupposti della sua stessa ipotesi, piuttosto che all'effettività

¹⁷ In questo caso, Libet intende per determinismo l'assunzione che la natura del mondo fisicamente osservabile possa rendere conto delle funzioni coscienti del soggetto; per indeterminismo, l'assunzione che il libero arbitrio possa produrre effetti in disaccordo con le leggi fisiche ovvero secondo l'idea che tali violazioni siano rintracciabili, ma non con le nostre tecnologie (es. ipotesi quantistica), piuttosto che secondo l'idea che non lo siano tout court.

¹⁸ B. Libet, *Op cit.*, p 551-564.

¹⁹ *Ibidem*. Si tratta di una licenza, in realtà 'free will'='libero arbitrio', come di seguito è sempre tradotto.

²⁰ Libet, 1996. Come non disponiamo di un progetto di test sperimentale che venga incontro alla domanda se siamo o meno forniti di libero arbitrio. Libet spiega di avere in mente un progetto test per verificare la sua ipotesi dei 'campi mentali di coscienza': la volontà potrebbe infatti influenzare le attività delle cellule nervose attraverso di esso, cioè senza alcuna connessione neurale che ne supporti l'effettività. La sua sarebbe una teoria iterazionista, non dell'identità.

²¹ R. Sperry (1913-1994). Medico e neurobiologo statunitense, Nobel per la medicina nel 1981 per i suoi studi sulle funzioni cognitive degli emisferi cerebrali in soggetti *split brain*.

sperimentale. Perché, infatti, contrariamente a ciò che questo esperimento mostra (una sopravvenienza della coscienza al processo cerebrale), il veto dovrebbe essere un evento causale cosciente, piuttosto che anch'esso una delibera incosciente prodotta dal meccanismo cerebrale? Nonostante in effetti altri, tra cui Velmans (1991),²² abbiano proposto che, per un principio di simmetria non contestato dai dati sperimentali, il veto debba essere considerato anch'esso avere il proprio cominciamento cerebrale inconscio, per la stessa ragione Libet arriva a concludere l'esatto opposto. In sostanza, non esistendo un tracciato EEG²³ del veto cosciente, è possibile dedurre sia che esso è parte del processo di libero arbitrio e sia dunque 'inconscio'; sia che esso costituisca la realtà mentale indipendente dal processo cerebrale in cui identificare l'effettività del libero arbitrio.

Le ragioni che spingono Libet a questa interpretazione sono dichiaratamente di ordine etico: secondo il neuroscienziato, se ammettiamo che il veto - nonostante non ce ne sia prova - sia anch'esso anticipato da un RP (che sarebbe lo stesso registrato prima dell'emergere alla coscienza della volontà di compiere l'azione che il veto inibisce), ciò significherebbe che l'individuo non è in grado di controllare le proprie azioni consciamente. In pratica, il soggetto sarebbe solo capace di divenire cosciente di meccanismi che lo riguardano. Dalla considerazione che un evento incosciente nel cervello sia responsabile causalmente sia della pulsione sia della sua censura, si arriva a negare la coscienza stessa e per Libet questo è inammissibile, perché dalla sua esistenza ci aspettiamo conseguia la responsabilità delle proprie scelte; ad esempio, coloro che soffrono di malattie neurologiche non vengono considerati responsabili delle loro azioni: nessun tribunale oggi condannerebbe un uomo affetto da sindrome di Tourette per le oscenità gridate, poiché egli non può scegliere di censurare o meno quel bizzarro comportamento; ma se anche le azioni di un individuo sano dovessero essere considerate di origine inconscia, nessuno potrebbe più essere considerato responsabile per le proprie azioni. Assunta l'origine cerebrale dell'azione volontaria, non solo non possiamo più dire che l'uomo è libero, ma nemmeno che è responsabile. Questo implicherebbe una completa ristrutturazione della nostra architettura concettuale comune, la cui chiave di volta, come nella celebre metafora kantiana, è la ragion pratica, la nostra certezza morale. La proposta di Libet è dunque di pensare il veto cosciente come l'essenza del libero

²² Max Velmans, professore presso il Dipartimento di Psicologia dell'Università di Londra, autore di *The Limits of Neurophysiological Models of Consciousness*, in *Behavioral and Brain Sciences*, 18 (4), 702-703 (commentary on Gray, J. (1995) *The contents of consciousness: a neurophysiological conjecture*. *BBS*, 18 (4): 659-676), Cambridge University Press, 1995.

²³ Elettroencefalogramma.

arbitrio e al contempo come quel qualcosa di ‘indefinito’ che ci permette di scegliere (ed essere così nominalmente padroni, quantomeno, del nostro comportamento).

È evidente che, per l’elaborazione dei test, Libet ha dovuto ridurre la definizione classica di ‘libero arbitrio’ a quella di ‘controllo volontario’: l’impostazione di ricerca, cui ha dovuto forzare la domanda, lo ha portato ad ottenere non più che delle indicazioni. Siamo dunque partiti da una definizione di libero arbitrio positiva, come possibilità di compiere volontariamente un movimento; l’anticipazione neurale, negando il carattere di assoluta initialità (che la filosofia trascendentale vuole propria della libertà, piuttosto che del libero arbitrio) della decisione cosciente per il detto movimento, avrebbe escluso che l’azione libera sia veramente tale; a seguito dell’apparente assenza di una simile anticipazione neurale prima della decisione cosciente di bloccare la stessa azione, siamo approdati, infine, ad una definizione di libero arbitrio negativa, come possibilità di controllare consciamente la volontà a compiere un movimento (che invece è anticipata ‘inconsciamente’ e dunque ‘governata dal meccanismo’).

L’interpretazione dei dati ottenuti resta incerta anche tenendo presente che tesi e ipotesi sono sviluppate all’interno di una costruzione semantica di riferimento appositamente creata, all’interno della quale non dovrebbe esistere prova sperimentale che ne contesti la validità (ed invece c’è, secondo Velmans). Come abbiamo già visto, Libet affianca ragioni d’ordine gnoseologico ed encomiabili argomenti etici a sostegno delle sue conclusioni: il problema è la loro indistricabilità dalle prove sperimentali; questo errore è una costante e si rileva sia quando si difende scrivendo che non esiste nessuna logica stringente, in nessuna teoria della mente e tanto meno in nessuna teoria dell’identità tra stati cerebrali e stati mentali, che imponga una specifica attività neurale a determinare la funzione del controllo cosciente; sia quando spiega che la sua proposta riconcilia neuroscienza e filosofia perché garantisce ancora una sorta di libertà; sia, ancora, quando, volendo riportare il senso delle sue conclusioni alla specificità della prassi scientifica, pretende invece di universalizzarlo,²⁴ sostenendo che sia caratteristico della scienza il limitarsi allo studio del processo in un sistema semplice per affermarne la validità anche in sistemi complessi: basterà dunque impegnarsi a ricordare che è stata adottata quella che chiama ‘un’extrapolazione’. Si tratta dunque di un problema d’ordine metodologico e, per il filosofo della scienza, epistemologico.

Libet sembra essere ben consapevole che le conseguenze teoretiche tratte dai suoi test non hanno il peso di una verità e che aggiungono solo un tassello alla comprensione dei

²⁴ Libet utilizza l’esempio della misura della carica dell’elettrone compiuta da Milliken, valida per ogni elettrone; l’esempio tuttavia è fizioso perché illustra un solo *modus operandi* della scienza.

movimenti volontari in generale. La sua posizione non sembra essere che la neuroscienza debba spiegare la coscienza, la volontà o la libertà così come le conosciamo attraverso la filosofia o la religione, quanto contribuire all'ideazione di modelli del mentale sempre più verisimili. Il pensiero speculativo, tuttavia, non può evitare di imporre un uso più attento degli argomenti e delle parole che, attinte dal linguaggio comune e forzate a denotare fenomeni fisici o neurobiologici, alimentano fraintendimenti: da un lato, ingannando l'immaginario comune e, dall'altro, radicando nelle nostre costruzioni concettuali idee confuse o errate, che non mancano di dar adito a disintellezioni anche gravi della realtà. La filosofia ha dunque il compito di evitare che la ricerca più avanzata crei, paradossalmente, ignoranza: privata di questo controllore delle affermazioni sulla natura umana, infatti, la neuroscienza, come tutte le innovazioni (o i pensieri difficili), rischia di divenire l'appiglio per ogni genere di conclusione o superstizione.

Ciò che stupisce di Libet, di contro al suo rigore scientifico, è difatti il paragrafo dedicato alle 'implicazioni etiche di come opera la volontà libera':²⁵ egli ritiene infatti che le sue soluzioni siano in perfetto accordo con le prescrizioni etiche e religiose più diffuse e che, d'altro canto, possano contribuire a marginalizzare quelle che restano scientificamente non comprovate:

This kind of role for free will is actually in accord with religious and ethical strictures, which commonly advocate that you "control yourself". Most of the Ten Commandments are "do not" orders. How do our findings relate to the questions of when one may be regarded as guilty or sinful, in various religious and philosophical systems?[...]President Jimmy Carter admitted to having had urges to perform a lustful act [...]But any such urges would be initiated and developed in the brain unconsciously, according with our findings.²⁶

una simile affermazione è di un'irresponsabilità imperdonabile, se non solo superficiale e riduttiva²⁷. Concessa e non ammessa l'ingenuità della compiaciuta analogia tra conclusioni arbitrarie ottenute attraverso il *Timing of Brain Processes* ed i Dieci Comandamenti, e ricordata la piaga del puritanesimo che contraddistingue la vita pubblica americana, questa riflessione incollocabile è una pericolosa premessa se usata per arbitrare questioni più delicate quali il conflitto, l'identità, la spiritualità o la politica. Volendo ricorrere alla stessa disinvoltura con cui Libet affronta le *ethical implications*, potremmo allora dire che l'idea di religioni più giuste di altre, in quanto i loro precetti trovano conferma nella

²⁵ B. Libet, *Op.Cit.*, in *Ethical implications of how free will operates*, p 560.

²⁶ *Ibidem.*

²⁷ Per lo stesso problema semantico che caratterizza tutta la produzione del neuroscienziato sulla 'coscienza' ed il 'libero arbitrio', per 'implicazione etica' si intendono solo le ragioni religiose e psicologiche.

scienza, ricorda la pretesa scientificità di cui si ammantò lo sterminio nazista. Restando su toni più neutrali, si ammetterà comunque che l'idea di un'etica fondata direttamente in questo genere di parzialità lascia sgomenti.

A nostro giudizio, l'esperimento *in vitro* sul libero arbitrio stabilisce il limite semantico cui giunge la proposta neuroscientifica e identifica il luogo in cui si apre lo spazio per il ripensamento, a partire dalla questione fondamentale di cosa sia in gioco nei test di Libet e cosa debba intendersi per *free will*.

3. Di cosa parliamo, quando parliamo di *libero arbitrio*?

Se da un lato è facile dire che la tradizione cui egli si riallaccia è quella anglosassone, poiché il dibattito sul cosiddetto libero arbitrio, nell'Europa continentale, è stato riassunto in quello più fondamentale sulla libertà con Kant,²⁸ dall'altro, non sarà scontato assumere che la tensione di tutto il suo esperimento sembra volta piuttosto ad identificare la coscienza stessa.

Per far chiarezza su questo aspetto fondamentale è necessario uscire dal linguaggio del neuroscienziato, e dalle ambiguità di metodo che abbiamo sottolineato. Per farlo, è possibile richiamarsi all'autore di uno dei lavori più esaustivi sul tema della coscienza – poiché in esso vengono trattate approfonditamente tutte le sue declinazioni -, ovvero a David Chalmers,²⁹ che ha infatti l'indubbio merito di averne catalogato in maniera ragionata tutte le accezioni (in cui ne parlano i protagonisti di cento anni di dibattito sulla mente cosciente): il suo lavoro sarà utile per confermarci che la questione del libero arbitrio in Libet è riconducibile, come controllo cosciente, alla sola coscienza psicologica.

Quel che è encomiabile, in Chalmers, è che pone un unico vincolo alla sua speculazione, quello di costruire una teoria della coscienza che funzioni: nella sua argomentazione non troveremo dunque misconosciuta la sostanziale coerenza tra la coscienza ed i processi cognitivi, come nelle posizioni indeterministiche estreme; riscontreremo, piuttosto, la stessa evidenza, già emersa con Libet, della possibilità di trarre conclusioni antiriduzioniste dalle neuroscienze, anche se ammessa per ragioni molto diverse. Il contributo più

²⁸ I. Kant, *Critica della ragion pratica*, Laterza, Bari, 1997. Il discorso percorre però tutte e tre le *critiche*.

²⁹ D. J. Chalmers, *La mente cosciente*, Mc Graw Hill, Milano, 1999. David John Chalmers (1966) è un filosofo australiano di area analitica. Si è laureato presso l'università di Alelaide, ha proseguito gli studi in varie università americane fino a divenire direttore del *Center for Consciousness* all'Università dell'Arizona. Dal 2004, è direttore dell'omonimo centro presso l'Australian National University. La sua ricerca è incentrata soprattutto sul problema della coscienza, che ha diviso in questione facile e difficile (nel suo aspetto di qualità soggettiva e irriducibile, *hard problem*).

tangibile dell'opera del filosofo australiano resta, tuttavia, la decostruzione della definizione standard di coscienza e la sua identificazione con la qualità soggettiva dell'esperienza (qualia), un 'qualcosa come l'effetto che fa'³⁰.

Definendo uno stato mentale 'cosciente' solo se possiede (è) una sensazione qualitativa (o qualità fenomeniche o *qualia*, appunto), non possiamo che concludere che l'argomentazione di Libet si espande oltre i propri limiti ed è viziata proprio per l'uso che dell'idea di 'coscienza' fa.

Esistono, per 'mente cosciente', almeno due diversi concetti: quello della mente fenomenica, che si caratterizza per ciò che prova, e quello della mente psicologica, che si caratterizza per ciò che fa; con questo, potremmo già collocare il 'libero arbitrio' ed il 'veto cosciente' nella seconda accezione. Ma, per verificare se nel lavoro di Libet sopravvive questa confusione semantica, è necessario rintracciarne l'origine.

Ne *La mente cosciente* il segno di tale adulterazione è rintracciato nell'assunto cartesiano che la mente è trasparente a sé stessa; come sappiamo, non essendoci altro, nello spirito, se non *cognitatio*, uno stato mentale inconscio è contraddittorio; la mancata rielaborazione di un simile presupposto, in occasione della fortuna delle dottrine freudiane, è il paradigma di quelle fatali distrazioni, a causa delle quali una disintellezione³¹ può sedimentarsi nel nostro linguaggio. Come ricorderemo anche in seguito, pur fondando una teoria della mente sull'intuizione che molte sue attività sono inconscie, lo stesso Freud si servì di un'analisi causale del pensiero e non di una fenomenica: in seguito, il preponderante sviluppo di teorie della psiche in cui gli stati mentali venivano considerati in base al loro ruolo causale e mai nel loro aspetto irriducibile (comportamentismo, eliminativismo, il funzionalismo stesso), ha portato rapidamente all'idea di escludere o negare la coscienza. Il problema è quindi ancora una volta di metodo, laddove non solo non sono stati rivisti, alla luce delle nuove teorie, i singoli pregiudizi di partenza (*la mente è trasparente a sé stessa*) ma, oltretutto, si è applicato un solo meccanismo, conosciuto osservando la natura, allo studio di fenomeni privilegiati ed alla sola natura irriducibili.

Pur riconoscendo la ragionevolezza e validità di quelle analisi, relativamente al caso in cui la coscienza psicologica non sia accompagnata anche da una qualità fenomenica, è tuttavia ora necessario distinguere quantomeno le proprietà mentali psicologiche (apprendimento, memoria) da quelle fenomeniche (gioia, tristezza), per capire a fondo

³⁰ Chalmers recupera quindi la definizione di Nagel: 'un essere è cosciente se c'è qualcosa come il provare la sensazione di essere quell'essere'.

³¹ Così chiama questo genere di fraintendimenti V. Jankelevitch, in traduzione dal francese in *Il non-so-che e il quasi-niente*, Marietti, Genova, 1987.

cosa si intende per ‘coscienza’. In particolare, a riguardo dell’uso che fa Libet dei concetti di libero arbitrio, coscienza e volontà, bisogna aprire un’ulteriore parentesi sugli stati mentali ‘proposizionali’, perché essi costituiscono parte integrante del suo esperimento (che si serve dei resoconti verbali dei soggetti in fase di *brain timing*): ai fini del test, i resoconti verbali dei soggetti sono chiaramente assunti come semplice adesione psicologica che orienta il comportamento; tuttavia la credenza del soggetto di voler agire resta, nel suo aspetto esperienziale, fenomenica (e come tale andrebbe studiata, se l’impiego dei resoconti fosse volto a spiegare stati mentali più primitivi quali il desiderio e la volontà inconscia così come la si studia in filosofia). Ovviamente, gli aspetti fenomenici e psicologici degli stati mentali sono quasi sempre cooccorrenti: il problema più evidente resta quindi, oltre a questo ora sollevato - che apre alla questione dello statuto ontologico della parola e dell’attendibilità scientifica dell’impiego dei resoconti verbali (Dennett) -, l’inadeguatezza del nostro vocabolario a distinguerli; è proprio parlando delle proprietà fenomeniche dell’esperienza cosciente con un lessico non-fenomenico che si è giunti a trattare in termini causali degli stati mentali.

A seguito di tutta questa disamina, possiamo ora affermare che quando Libet parla di ‘libero arbitrio’ (e soprattutto quando ne ‘salva’ l’ammissibilità come ‘veto cosciente’) parla invero di un aspetto della ‘coscienza psicologica’,³² ovvero della consapevolezza razionale. Di conseguenza, i problemi morali, religiosi e metafisici (a proposito dell’idea che abbiamo di noi come esseri umani) potrebbero esser lasciati scivolare più lontano dalla ricerca neuroscientifica e ricondotti all’ambito, loro più consono, della filosofia, della psicologia e della religione. La coscienza psicologica può infatti spiegare o aderire ad una spiegazione sul funzionamento fisico, ma lascia completamente irrisolte e sensatamente irriducibili le questioni legate al perché (e se) un esistente abbia esperienza cosciente. In sostanza, quando si considera la coscienza in senso qualitativo, scopriamo che l’esistenza o meno della nostra esperienza cosciente non è implicata logicamente da fatti inerenti la nostra organizzazione funzionale.³³

Come concluderà Chalmers, il tasto dolente di ogni altra teoria neurobiologica (a partire da Crick e Koch fino ad Edelman) o fisica (Penrose, Hameroff) che spiega la coscienza è che nelle loro proposte la correlazione tra una qualche proprietà neurobiologica e la coscienza è un fatto, un presupposto e al contempo ciò che dovrebbero dimostrare e che non

³² I tipi di coscienza psicologica sono, per Chalmers, tutti quelli che, nonostante possano essere associati al fenomenico, non possono essere assimilati alla coscienza fenomenica in generale (qualia); tra essi include: lo stato di veglia, l’introspezione, la referibilità, l’autocoscienza, l’attenzione, il controllo volontario e la conoscenza stessa. In D. J.Chalmers, *Op. Cit*, cap.I.

³³ J.Chalmers, *Op. Cit*, cap. III, *Si può dare una spiegazione riduzionistica della coscienza?*, par.1.

riescono a spiegare, in assenza di 'principii ponte'. Lo stesso G.M.Edelman³⁴, che usa in modo tale e quale a Libet il linguaggio non-fenomenico nei suoi studi sulla 'coscienza' (in particolare su quella che chiama 'coscienza di ordine superiore'), colloca il suo aspetto fenomenico al di fuori delle competenze scientifiche poiché, se anche venissero individuati dalla neurobiologia i processi soggiacenti l'esperienza cosciente, non sarebbe comunque spiegata l'esperienza stessa.

Possiamo dunque affermare che oggi la lezione di Canguilhem suggerisce qualcosa di nuovo: infatti, il tipo di problematicità sollevato da lavori come quello di Benjamin Libet concorrono a vanificare il richiamo al buon senso che egli stesso ci offre, quando chiede di non credere alle presunte scoperte che giustificherebbero la riduzione (o l'eliminazione) della coscienza. La validità delle sue inferenze si esaurisce nell'ambito della spiegazione del processo cerebrale di preparazione al movimento, in correlazione alla consapevolezza (e non al libero arbitrio o alla coscienza). Il significato e le implicazioni morali o religiose, anche riferite al solo controllo volontario (o, meglio, al meccanismo operativo che ci consente di inibire il compimento di un gesto), sono dunque da ridimensionare, quando non fuori luogo: la loro discussione compete ad altri ambiti teorici e ad altre metodologie d'indagine; essa, inoltre, quando venga condotta col metodo scientifico, concorre solo ad alimentare una falsificazione (ed una falsa idea di noi), che mortifica gli sforzi in atto per elaborare risposte soddisfacenti ad aspirazioni conoscitive fondamentali.

4. Il problema della razionalità come epilogo del discorso sulla libertà. Filosofia dell'azione

Più interessante della sola prospettiva neuroscientifica per rispondere alle perplessità etiche sollevate dello stesso Libet a seguito delle sue scoperte scientifiche (quando si debba essere considerati colpevoli o peccatori³⁵ e la questione primaria: perché dovremmo essere considerati responsabili delle nostre azioni volontarie, se procedono da un processo inconscio?³⁶) è quella impostata da Donald Davidson³⁷, che trasforma la

³⁴ G.M.Edelman (New York, 1929), Nobel per la medicina nel 1972, ideatore della teoria della selezione dei gruppi nervosi (TSGN), autore del recente *Più grande del cielo. Lo straordinario dono fenomenico della coscienza* e del celebre *Il presente ricordato*, è considerato padre del neurodarwinismo, una teoria che vuole dar ragione in modo complessivo e globale della coscienza su basi neurobiologiche.

³⁵ B.Libet, *Op. Cit.*, p. 561.

³⁶ B. Libet, *Op cit*, p. 558.

³⁷ Donald Davidson (1917-2003) è stato un celebre filosofo americano, ha insegnato nelle più prestigiose università nordamericane, tra cui Harvard, Stanford, Princeton e Berkeley. E' autore, oltre ad un centinaio di articoli scientifici, di *Verità e significato* (1967) e di *Azioni ed eventi* (1980).

domanda nella seguente: in che misura un gesto volontario deve essere considerato l'azione di una mente cosciente piuttosto che l'evento prodotto da una causa inconscia?

A nostro giudizio, Davidson costituisce un interlocutore più appropriato per Libet rispetto, ad esempio, ai già citati Foucault e Canguilhem, i quali si esprimevano in un contesto culturale troppo distante da quello del neuroscienziato. La collocazione in una diversa tradizione di pensiero, quando si trattano temi vasti e aperti come razionalità o libertà, suscettibili quindi delle più svariate articolazioni, implicherebbe la premessa di tante e tali distinzioni da renderci impossibile far dialogare gli autori.

In *I paradossi della irrazionalità*,³⁸ il filosofo americano affronta la validità di alcuni nodi cruciali e ancora attuali della teoria freudiana. Il testo spiega dunque come Freud, medico psichiatra, sia da biasimare proprio per la pretesa di trattare i fenomeni mentali come fenomeni naturali. Davidson fa quindi notare come Freud avesse attinto il suo vocabolario da scienze quali l'elettromagnetismo e la neurologia, enfatizzando l'aspetto naturalistico della sua teoria al punto da renderla oggi, per molti critici, 'confusa' (e quindi oggi inadeguata). Alla neuroscienza di Libet abbiamo opposto una critica simile, se non la medesima invertita: essa non ne inficia la validità, ma insiste sulla necessità di una messa in parentesi delle sue conclusioni speculative, fino a che non ne sia stato epurato il linguaggio che, dovendosi limitare al vocabolario scientifico, attinge invece da quello della filosofia e della psicologia, attribuendosi, per autoinganno, un peso etico che, di fatto, non ha.

Alla luce delle riflessioni di Davidson, il posizionamento dell'uomo nel mondo e in società, che Libet vorrebbe fondare direttamente nelle sue scoperte, appare subito di stampo moralista: se il libero arbitrio risiede nel veto cosciente ad una volontà inconsciente (laddove la 'prova' del carattere inconscio della volontà è fornito dall'RP cerebrale), la 'peccaminosità' o la disattesa sociale di un'azione possono infatti essere combattute con una maggiore risolutezza. L'azione 'indegna' sarebbe, infatti, da attribuirsi a 'debolezza della volontà', come da tradizione classica³⁹.

³⁸ D. Davidson, *Paradoxes of Irrationality*, in Richard Wolheim & James Hopkins eds, *Philosophical Essays on Freud*, Cambridge, CUP, 1982. Tradotto come *I paradossi dell'irrazionalità*, in AAVV, a cura di D. Meghnagi, *Studi freudiani*, Guerini, 1989.

³⁹ È il problema dell'etica così come viene trattato da Platone e Aristotele, rispettivamente nel dialogo socratico *Protagora* e nel libro VII dell'Etica Nicomachea di Aristotele (*akrasia*). Platone nega l'esistenza dell'intemperanza, ovvero dell'azione irrazionale: si tratta di 'ignoranza': essa sarebbe infatti un paradosso, perché ammette un'azione 'contro il proprio più meditato giudizio', che è conoscenza (significherebbe affermare che la 'scienza' può essere sopraffatta dalle passioni). Secondo Aristotele invece l'intemperanza esiste e va trattata all'interno del discorso sull'incontinenza solo che, al contrario di questa, in cui viene pervertito il principio etico, l'intemperante è colui che si fa vincere dalle passioni, sotto la cui malefica influenza 'agisce come ubriaco' (pur conoscendo la premessa maggiore del sillogismo pratico, non conosce la

Alla radice del problema della responsabilità, imposto alla nostra attenzione, secondo Libet, dall'idea che le azioni 'libere' siano invece determinate (in realtà, *anticipate*) da processi cerebrali di cui siamo completamente inconsapevoli, c'è quello, eviscerato per secoli in filosofia, della causalità mentale. Se infatti, come faceva Freud, consideriamo la causa mentale come una causa naturale, descrivendo gli eventi mentali come neurologici, non solo perdiamo la possibilità di definire il limite tra razionalità e irrazionalità ma, con Libet, lo trasformiamo in quello tra normalità e patologia, finendo per imporre, su base medica, l'universalizzabilità e la prescrittività⁴⁰ del principio di responsabilità.

Secondo Davidson, un modello interpretativo del mentale che viene incontro al problema è quello che recupera la teoria della 'mente a comparti'. Accettando che alcuni eventi mentali siano mere cause di altri, sarà possibile descrivere l'irrazionalità come il risultato di relazioni causali non logiche tra le parti della mente, le quali ciascuna reagisce o agisce secondo un nucleo di credenze, di cui le altre parti non sono coscienti. Il problema più fondamentale, che rende incompleta sia questa, che è una descrizione funzionale del mentale, sia la prospettiva neurobiologica, è che, come conclude Davidson, non si spiegano in nessun modo ancora innegabili fenomeni legati alla volontà, tra cui:

la riforma che tendiamo a tenere in alta stima, e che è stata considerata l'essenza stessa della razionalità e la sorgente della libertà. E tuttavia è chiaramente un caso di causalità mentale che trascende la ragione [...] Una teoria che non spieghi l'irrazionalità sarebbe una teoria che non potrebbe neppure spiegare i nostri saltuari sforzi, e occasionali successi, di auto-critica e auto-miglioramento.⁴¹

La libera volontà, pur nell'accezione limitata di cui fa uso Libet,⁴² contestualizzata in un pensiero come quello di Davidson, passa da essere l'allarmante premessa di un'etica scienziata medicalmente fondata ad una prospettiva parziale ma facilmente riconciliabile con l'aspetto fenomenico che essa stessa riconosce base della sua domanda (se pur così poco presa in considerazione nell'argomentazione 'morale').

Una volta orientata, attraverso il lavoro di questo genere di autori, la ricerca di Libet può essere riavvicinata e messa a dialogo anche con le più recenti elaborazioni della filosofia

minore, ovvero non può applicare la norma al caso particolare; in sostanza, sbaglia nel trarne la conclusione; la passione lo induce ad un errore conoscitivo.)

⁴⁰ Si tratta delle caratteristiche che R. M. Hare definisce 'le regole del gioco' per la costruzione di una scienza morale. R. M. Hare, *Libertà e ragione*, Il Saggiatore, Milano, 1973.

⁴¹ D. Davidson, *Op. cit.*, 485.

⁴² Henrik Walter, *Neurophilosophy of free will*, in AA.VV., R.H.Kane, *The Oxford Handbook of free will*, Oxford University Press, NY, 2002. In questo articolo Walter amplia le definizioni di Libet attraverso nuovi dati sperimentali ottenuti con test di diverso tipo (TAC, PET e brain screaming); propone un modello deterministico ma compatibilista.

continentale, che riconoscono nell'inizialità della libertà non tanto un tipo particolare di causalità, ma il tratto d'efficacia della causalità stessa⁴³, in una visione che unisce la tradizione trascendentale alla cultura materiale, cui la postmodernità ha esposto il pensiero.

Laura Beritelli

Bibliografia

- ❑ Binazzi A., *Conferenza di George Lakoff*, 'Humana.mente' n°4, 2/2008.
- ❑ Canguilhem C., *Il normale e il patologico*, 1998, Einaudi, Torino.
- ❑ Chalmers D. J., *La mente cosciente*, 1999, Mc Graw Hill, Milano.
- ❑ Davidson D., *I paradossi dell'irrazionalità*, in AAVV, a cura di D. Meghnagi, *Studi freudiani*, 1989, Guerini.
- ❑ Foucault M., *L'archeologia del sapere*, Rizzoli, 1971, Milano.
- ❑ Hare R. M., *Libertà e ragione*, 1973, Il Saggiatore, Milano.
- ❑ Jankelevitch V., *Il non-so-che e il quasi-niente*, Marietti, 1987, Genova.
- ❑ Libet B., *Do we have free will?*, in AA.VV, a cura di R.H.Kane, *The Oxford Handbook of free will*, Oxford University Press, NY, 2002, p 551-564.
- ❑ Manetti D., Zipoli Caiani S., *Intervista a George Lakoff*, 'Humana.mente' n°4, 2/2008.
- ❑ Nancy J. L., *L'esperienza della libertà*, 2000, Einaudi, Torino.
- ❑ Nancy J. L., *L'intruso*, Cronopio, 2000, Napoli.
- ❑ Velmans M., *The Limits of Neurophysiological Models of Consciousness*, in *Behavioral and Brain Sciences*, 18 (4), 702-703 (commentary on Gray, J. (1995) *The contents of consciousness: a neurophysiological conjecture*. *BBS*, 18 (4): 659-676), 1995, Cambridge University Press.
- ❑ Walter H., *Neurophilosophy of free will*, in AA.VV, a cura di R.H.Kane, *The Oxford Handbook of free will*, 2002, Oxford University Press, NY.

⁴³ J. L. Nancy, *L'esperienza della libertà*, Einaudi, Torino, 2000.

NEUROIMAGING: CONTINUITÀ E INNOVAZIONE¹

di Matteo Borri

Ora, come possiamo esaminare qualcosa che stiamo usando continuamente? Come possiamo analizzare i termini nei quali esprimiamo di solito le nostre osservazioni più semplici e dirette, e riportarne in luce i presupposti? In che modo possiamo scoprire il tipo di mondo che presupponiamo, dati i nostri modi di procedere abituali?

La risposta è chiara: non possiamo scoprirlo dall' *interno* .²

1. Introduzione

Il *neuroimaging* l' insieme di tecniche che forniscono informazioni *in vivo* sul sistema nervoso, è un processo complesso che si realizza all'interno di una specifica specializzazione medica: la medicina nucleare³. La complessità di questo settore è determinata in primo luogo dall'alto grado di interdisciplinarietà: i servizi offerti dalla medicina nucleare sono resi possibili da figure professionali che afferiscono a discipline diverse⁴. Risultano poi estremamente articolate anche l'organizzazione tecnologica⁵ e la sua gestione⁶. L'*imaging* è dunque un sapere altamente specialistico che per i progressi oggi ottenuti è sempre più centrale nelle prassi mediche.

Il presente lavoro intende far emergere, da alcuni degli aspetti che concorrono a delineare la complessità di questo settore delle scienze biomediche, prospettive di indagine e problematiche che sollevano riflessioni epistemologiche.

A tale scopo, verrà indagato se si possano riscontrare fattori di continuità fra le attuali ricerche e i primi studi scientifici sul cervello. Successivamente verrà valutato se e sotto quali aspetti si possa considerare "innovativo" il contributo delle tecniche di *neuroimaging*, in modo particolare in riferimento alla loro possibile dimensione euristica.

¹ Ringrazio il Centro di Documentazione per la Storia dell'Assistenza e della Sanità Fiorentina di Firenze per il permesso di consultazione dell'opera di Morselli e per la riproduzione delle immagini ad esse relative.

² Feyerabend, 2005, p. 28.

³ La natura della medicina nucleare è esplicitata nei documenti dell'*European Board of Nuclear Medicine*. La finalità specifica viene così indicata: "La mission della medicina nucleare è una disciplina la cui specificità risiede nell'impiego regolamentato di radionuclidi artificiali in forma "non sigillata" per attività sia di tipo diagnostico che terapeutico", Maffioli, 2006, p. 1. Attualmente in Italia operano 250 centri di medicina nucleare.

⁴ Gli operatori di una unità di medicina nucleare sono: medico specializzato in medicina nucleare; fisico sanitario; radio farmacista/radiochimico; tecnico sanitario di radiologia medica.

⁵ La strumentazione della medicina nucleare è molto complessa ed è sottoposta ad una normativa articolata e molto prescrittiva. Un primo strumento è quello dedicato alla produzione di radionuclidi: il ciclotrone. Il primo ciclotrone dedicato a servizi medici fu usato nel 1955 presso l'*Hammersmith Hospital* a Londra. Attualmente in Italia ne sono in funzione circa 20. I radionuclidi vengono usati in strumentazioni diverse corrispondenti a differenti tipi di visualizzazione.

⁶ Le prestazioni del Servizio di Medicina Nucleare coprono più settori: applicazioni diagnostiche *in vivo* e *in vitro*; esami speciali; ambulatorio di medicina nucleare; applicazioni terapeutiche; applicazioni di ricerca.

2. Guardare dentro il corpo: una tradizione di ricerca

Quando viene presentato un referto di neuroimaging, abbiamo un'immagine che viene direttamente dall'interno del corpo, e praticamente in tempo reale. Una domanda antica, quella riguardante il funzionamento del corpo, che nei secoli passati poteva essere ispezionato – quando non era addirittura proibito – solo *in vitro*. Oggi ispezionare il corpo *in vivo* è prassi consolidata che dà risposte a "nuove" e antiche domande.

Le linee di ricerca che guidano l'indagine fatta con le tecniche di *neuroimaging* sono quelle che da sempre afferiscono alla neurologia e alla psichiatria⁷. Se in letteratura si concorda nell'indicare le ricerche di Paul Broca e Carl Wernicke⁸ come momento iniziale degli studi neuropsicologici, è nondimeno corretto focalizzare l'attenzione sulle caratteristiche generali dell'ambiente scientifico nel quale esse si sono sviluppate.

La storia delle conoscenze scientifiche sul cervello inizia con un approccio basato sulla patologia dei disturbi mentali e interseca quella più ampia metamorfosi della medicina che ebbe inizio nel periodo a cavallo tra Settecento e Ottocento e che portò alla nascita della Clinica modernamente intesa⁹.

Nell'impianto di un nuovo processo di causazione la ricerca medica assunse come centrali due entità: l'organo e il tessuto biologico¹⁰. La psichiatria e la neurologia si introdussero in questa nuova dimensione riconoscendo come campi di studio propri un sistema – il Sistema Nervoso Centrale – e un organo: il cervello. Da questo assunto la ricerca si rivolse quindi verso la definizione dei processi etiologici delle malattie mentali, individuandoli nel concetto di "lesione nel cervello".

Dopo le prime ipotesi avanzate nel XVIII secolo da Johann Meyer¹¹ – che sviluppò uno dei primi tentativi di passare oltre la concezione di una coscienza intesa come un tutto

⁷ Il XIX secolo merita un posto particolare nella storia della medicina, non solo per progressi in settori specifici come quello della psichiatria che cominciava a studiare scientificamente le malattie mentali ma anche per lo sviluppo di nuove metodologie conoscitive. In particolare è qui da ricordare lo sviluppo dei sistemi diagnostici rivolti verso l'interno del corpo come quelli dei primi servizi di radiologia. "Rendere visibile l'interno del corpo umano senza aprirlo, effettuare una sorta di autopsia senza praticare dissezioni, questo sogno dei medici di un tempo è diventato realtà", Grmek, 1998, p. 285.

⁸ Cfr. Young, 1970 e Shallice, 1990.

⁹ Cfr. Foucault, 1998.

¹⁰ "La storia della patologia fa parte della storia della medicina clinica in senso ampio e non è meno complessa della storia delle tecniche diagnostiche. Il passaggio da una patologia umorale ad una patologia degli organi di tipo morgagnano, segna già una rottura epistemologica fondamentale, se non la rivoluzione più importante che abbia mai subito la medicina. Ma questa patologia anatomica ha essa stessa conosciuto una trasformazione capitale le cui implicazioni hanno sconvolto anche la clinica, trasformazione che ha consistito nella rifusione della patologia degli organi su una nuova base, quella della patologia dei tessuti. Numerosi storici hanno visto, ed ancora vedono, in questa elaborazione di una problematica istologica o dei tessuti, il segnale più importante della nascita della medicina clinica moderna", Keel, 2007, p. 17.

¹¹ Cfr. Meyer, 1779.

inscindibile correlando le varie facoltà psichiche ad una articolata anatomia cerebrale – e dopo quelle frenologiche di Franz Joseph Gall¹², l'ottica neurologica si spostò definitivamente verso l'“interno” del corpo. L'analisi *post mortem* del cervello accompagnerà per oltre un secolo gli studi di patologia mentale. Questo specifico indirizzo trovò un contributo significativo¹³ nell'opera di Wilhelm Griesinger. I suoi studi rappresentarono un vero punto di svolta in quanto determinarono il passaggio da una concezione etiologica basata su cause psicologiche e affettive all'assunzione di un rapporto di causazione organica¹⁴.

Gli studi di Wilhelm Griesinger cambiarono la prospettiva di ricerca psichiatrica, portando a considerare primari alcuni aspetti. Il primo di questi fu proprio la centralità dello studio del Sistema Nervoso Centrale e del cervello.

Se i fatti fisiologici e patologici ci mostrano che questo organo non può non essere altro che il cervello, dobbiamo concludere che sempre per le malattie mentali abbiamo una “affezione” del cervello.¹⁵

Questo tipo di studi aveva però un punto particolarmente critico: lo stato dei tessuti cerebrali che erano esaminati *post mortem*. Tale condizione determinava infatti cambiamenti nell'organo analizzato e comunque questi reperti risultavano ‘distanti’ dalle dinamiche di un cervello in funzione. Anche riconoscendo questi limiti oggettivi, Griesinger poteva tuttavia affermare:

Dalle analisi anatomiche *post mortem* non rileviamo solo conferme, ma possiamo dedurre certezze capaci di gettare una luce più viva sulla natura intima, sull'essenza di queste malattie, sulla fisiologia e la patologia del cervello in generale.¹⁶

Altro elemento caratteristico delle ricerche di Griesinger fu lo sviluppo di un tipo di conoscenza basata sul principio del “correlato anatomico”¹⁷.

Per Griesinger una tale prospettiva implicava esaminare in modo “scientifico” anche la

¹² Cfr, Gall, 1822-23.

¹³ Cfr. Ellenberger, 2003.

¹⁴ “Noi oggi respingiamo i grossolani errori delle scuole antiche, che studiavano la follia esclusivamente dal punto di vista psicologico e morale, e che lottavano contro chi voleva cercare per le malattie mentali le lesioni anatomiche”, Griesinger, p. 479 (la traduzione è mia).

¹⁵ Griesinger, 1865, p. 1.

¹⁶ Ivi. 477.

¹⁷ “Il ruolo dell'anatomia patologica è di ricercare quali sono le lesioni che si trovano ordinariamente nelle autopsie degli alienati e arrivare attraverso il confronto di queste lesioni con l'insieme dei sintomi che presentavano a stabilire nozioni anatomiche precise su queste malattie cerebrali”, Ivi, p. 480.

psiche, riconoscendo come momento costitutivo della diagnosi una specifica attenzione alla sintomatologia psicologica. “Per ben comprendere l’alienazione mentale dobbiamo identificarci con lo stato dell’anima degli alienati.”¹⁸

Lo psichiatra, indicato in letteratura principalmente come l’iniziatore della scuola somatica, seguì in effetti una prospettiva di studi molto articolata e di carattere interdisciplinare, descrivendo una articolata pluralità etiologica riferendosi anche a categorie psicodinamiche¹⁹.

Il nuovo indirizzo “biologico” negli studi sui disturbi mentali non rimase confinato solo ai paesi di lingua tedesca ma era diffuso in tutta Europa. Un apporto significativo nel panorama scientifico internazionale lo troviamo anche in Italia nella scuola di Reggio Emilia²⁰ e nel lavoro di uno psichiatra che influì significativamente negli studi della fine del XIX secolo: Enrico Morselli²¹.

Assunta la visione “positivista” a base della ‘nuova’ psichiatria, Morselli riconobbe come prioritario per questa disciplina aderire ai “bisogni ed alle nozioni della scienza”. Proprio la ricerca anatomica venne messa al centro del lavoro dello psichiatra, così come aveva indicato Griesinger:

La mente sovrana di Griesinger vide l’errore di mantenere la psichiatria sotto il giogo della vecchia psicologia, ed introdusse nella psicologia patologica le idee feconde di Herbart massime intorno ai rapporti tra l’intelligenza ed il sentimento. Egli per il primo pensò che non bastava studiare il pazzo dal lato sintomatico, né considerare la pazzia come un prodotto del caso, un *casus naturae*: ma che occorreva investigarne le leggi e determinarne almeno la genesi individuale.²²

¹⁸ Ivi, p. 270. Qui il concetto di “anima” corrispondeva alla psiche. Il termine tedesco *Seele* può essere tradotto in italiano sia come psiche sia come anima, riconoscendo comunque a questo termine l’indicazione dell’insieme delle facoltà psichiche.

¹⁹ Cfr. Ellenberger, 2003.

²⁰ Nel periodo fra il 1860 e il 1880, si sviluppò a Reggio Emilia una scuola basata sulla interdisciplinarietà degli studi (laboratori di batteriologia, chimica, istologia, psicologia) che usando l’esame psicologico come strumento diagnostico sviluppava parallelamente ricerche sulle alterazioni del cervello. In questo percorso di studi fu ottenuto un interessante modello di correlazione fra funzioni psichiche e circuiti cerebrali. Un’opera significativa fu la monografia di Luciani e Seppilli del 1885, ristampata nel 2002 a cura di Carmela Morabito.

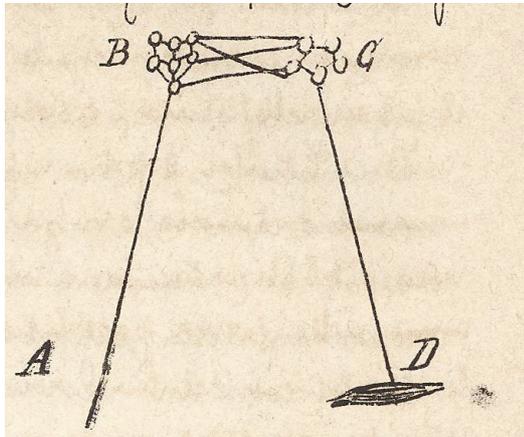
²¹ Enrico Morselli, costituì un punto di riferimento non solo nello specifico campo della psichiatria ma anche nel più ampio contesto culturale italiano. Come psichiatra egli lavorò presso l’ospedale psichiatrico di Reggio Emilia, e successivamente fu direttore dell’Ospedale psichiatrico a Macerata (1877) e poi a Torino (1880). Autore di un *Manuale di semeiotica della malattia mentale*, sviluppò ricerche significative sulla genesi corticale dell’afasia. Docente di psichiatria a Torino dal 1880 e a Genova dal 1889 si interessò in modo specifico alla psicologia sperimentale. Fondò e fu direttore della *Rivista di filosofia scientifica* che fu pubblicata dal 1881 al 1891. Essa ebbe lo scopo di sviluppare in Italia, in forma interdisciplinare, l’impostazione del pensiero positivista.

²² Morselli, 1887-88, pp. 7-8.

Morselli affermerà inoltre: “Io seguo questo schema, ritenendo col Griesinger, che non si può comprendere il cervello, se non vien studiato con lo stesso indirizzo del midollo spinale.”²³

Assumendo una metodologia basata sulla ricerca degli elementi neurali “di base”, Morselli indicò come punto di partenza un preciso elemento funzionale: l’atto riflesso.

L’atto riflesso, la più elementare delle funzioni del sistema nervoso è quello a cui si possono ridurre tutte le funzioni apparentemente complicate di questo. Può essere raffigurato da una fibra centripeta, che parte da una superficie sensiente A, giunge a un gruppo di cellule B unite per fibre di associazione ad un altro gruppo di cellule più grandi C, da cui parte una fibra centrifuga, la quale si porta ad una fibra muscolare D.²⁴



Con questo modello, lo psichiatra italiano si avvicinò allo studio del cervello: “Nello stesso modo nel cervello abbiamo i centri che presiedono all’atto psichico riflesso, e nei quali si ripercuotono gli stimoli centripeti per trasformarsi in impulsi centrifughi.”²⁵ Egli riconobbe che il cervello aveva (comunque) compiti più complessi, che andavano oltre alla semplice funzione automatica: esso operava come un “reostato”, “ritenendo gli impulsi”.

Questo potere che ha il cervello di ritenere allo stato latente gli impulsi giunti dalla periferia, questa resistenza che le cellule dei centri corticali oppongono a che quelli stimoli si esplichino immediatamente nell’atto riflesso costituisce il potere inibitorio del cervello, quello che i filosofi avevano chiamato volontà. La volontà non è che la inibizione che la corteccia cerebrale oppone alla trasformazione immediata della funzione in movimento.

²³ Morselli, Genova, p. 8.

²⁴ Ivi, p. 6-7.

²⁵ Ivi. 7.

La sensazione rimane là come un ricordo, le sensazioni immagazzinate e ricordate danno l'idea.²⁶

Morselli ricercava anche in campo psicologico una funzione "di base". Questa era all'epoca individuata nel concetto di sensazione²⁷. Come per Griesinger, anche per lo psichiatra italiano affrontare in modo compiuto lo studio della malattia mentale significava accompagnare lo studio anatomico a un preciso studio psicologico²⁸, effettuato sulla base di una psicologia scientifica che offriva ormai percorsi di ricerca in linea con le richieste di una scienza positivista²⁹.

Le conoscenze disponibili nel campo di una psichiatria proiettata sempre più verso l'anatomia risultavano fondate ed euristicamente significative³⁰. Morselli poté così sintetizzare una delle prime elaborazioni "spaziali" del cervello indicando una "localizzazione probabile dei centri motori e sensitivi nel cervello umano".³¹

²⁶ Ivi, pp. 7-8.

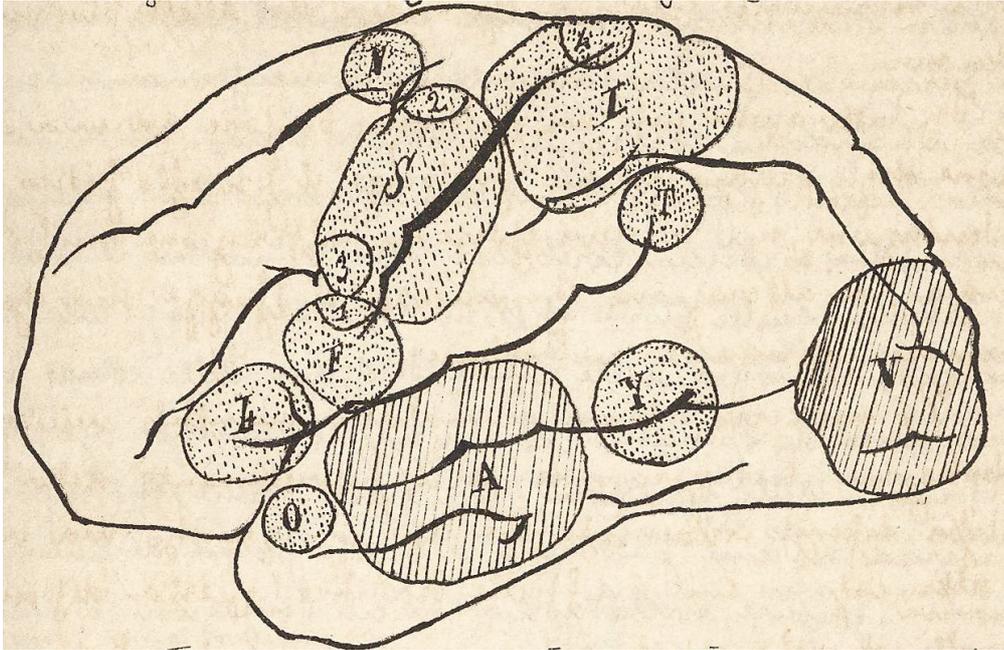
²⁷ Negli studi che costituirono l'inizio della psicologia scientifica, la "sensazione" fu riconosciuta come la 'chiave' per accedere allo studio della psiche. Questi erano iniziati con le ricerche di Ernst Heinrich Weber, docente di anatomia a Lipsia. Studiando la discriminazione percettiva egli arrivò a una fondamentale correlazione: il rapporto fra la soglia differenziale (incremento di una sensazione) e lo stimolo di riferimento è costante all'interno di ogni specifica forma di sensibilità (legge di Weber). Furono poi gli studi psicofisici di G. Fechner a sviluppare più compiutamente il concetto di sensazione. Egli dimostrò una caratteristica basilare della sensazione: la sua crescita avviene in progressione aritmetica in conseguenza dell'aumento dello stimolo che invece cresce in progressione geometrica. La sensazione risultò quindi una grandezza proporzionale al logaritmo dello stimolo che la produceva. Cfr. Zunini, 1972.

²⁸ Cfr. Morselli, 1894.

²⁹ La "nuova" psicologia, scientificamente fondata sul metodo sperimentale, si sviluppò in particolare nei laboratori delle università tedesche. Lo scopo centrale delle sue ricerche fu l'individuare funzionalità psichiche misurabili. Cfr. F. Meotti, 1971.

³⁰ "[...] la Psichiatria non ha forse cercato per mezzo di Meynert di riunire le leggi della patogenesi della pazzia alle leggi comuni dell'innervazione? Non ha egli cercato di determinare i rapporti che passano fra le funzioni mentali lese e le associazioni fra le cellule cerebrali?" Morselli, 1887-88, p. 14.

³¹ "Nella figura qui annessa noi abbiamo distribuito secondo gli ultimi studi le diverse zone della corteccia cerebrale in cui arrivano le impressioni dalla periferia, vengono percepite e immagazzinate, per essere poi a loro volta risvegliate da stimoli e trasformate in immagini ed idee". Morselli, Genova, p. 11-12.



Legenda: "Centri motori – L centro dei movimenti d'articolazione della parola.

F centro dei movimenti della faccia

1. Centri masticatori – N.

Centri dei movimenti della testa (nuca e collo) – Y.

Centro di movimenti degli occhi – O.

Centro di movimento delle orecchie – T.

Centro di movimento del tronco – S.

Centro di movimento delle membra superiori.

2. Centro dei movimenti della spalla. I. centro di movimento delle membra inferiori.

3. Centro di movimento del polso.

4. Centro di movimento delle dita grosse dei piedi.

Centri Sensitivi

V. Centri visivi.

A. Centri auditivi.

Nota bene: i centri motori son rappresentati dalle zone punteggiate, i centri sensitivi dalle zone tratteggiate."³²

Alla fine del XX secolo l'impianto conoscitivo nella neurologia e nella psichiatria era veramente esteso e complesso. Metodi di indagine sempre più precisi consegnavano alla

³² Morselli, Genova, pp. 11-12.

comunità scientifica dati anatomici, istologici, psicologici. Le conoscenze cominciavano a venire organizzate anche sul piano teorico. Un filone di ricerca relativo a queste sistematizzazioni, fu la rappresentazione spaziale delle funzioni psichiche nel cervello. Importanti furono gli studi di David Ferrier³³ e di Konstantin Brodmann³⁴. Alcuni risultati ottenuti erano rivolti a sviluppare una articolata concezione di complessità funzionale. Come auspicava Morselli:

Molti avrebbero la pretesa che la psichiatria indicasse il centimetro quadrato o il millimetro cubo di sostanza nervosa dove succede il delirio, l'idea fissa, il perversimento degli affetti: ma ciò è un assurdo [...] Ad ogni idea, ad ogni emozione prende parte tutto o quasi tutto l'organo cerebrale ed allora è evidente che prima di stabilire nel pazzo il punto preteso corrispondente alla lesione, sarebbe utile esaminare minutamente tutto il meccanismo, ma non il meccanismo arrestato dalla morte, bensì il meccanismo complicatissimo in istato di funzionalità attuale.³⁵

3. Medicina Nucleare

La richiesta anticipatrice di Morselli è oggi in larga misura concretizzata dal *neuroimaging*. Abbiamo già descritto come la medicina nucleare operi attraverso procedure complesse. Per cogliere l'importanza di questa attività conoscitiva ritengo importante focalizzare l'attenzione su alcuni dei processi che ne caratterizzano la realizzazione.

Sebbene l'aspetto terapeutico rivesta un ruolo importante, in questo lavoro saranno analizzati gli aspetti che riguardano la ricerca neurologica *in vivo*.

Le ricerche di *neuroimaging* sono oggi centrali nella gestione dei pazienti e permettono di studiare aspetti diversi – morfologico o funzionale – dell'organo interessato, attraverso la valutazione di immagini scintigrafiche³⁶.

Il *brain imaging*, oggetto di questo studio, è andato sviluppandosi e affinandosi nel corso degli ultimi venti anni permettendo oggi l'indagine delle reti neurali e dei grandi aggregati neuronali, 'regioni' che risultano attivarsi in relazione a compiti cognitivi (linguistici, motori, di apprendimento).

³³ Cfr. Ferrier, 1876.

³⁴ Brodmann fu direttore del laboratorio di anatomia nella clinica psichiatrica di Tubinga e successivamente della sezione istologica dell'Istituto Psichiatrico di Monaco. All'interno degli studi istologici i contributi specifici di Brodmann furono le ricerche sulla struttura cellulare della corteccia cerebrale e la realizzazione di tavole citoarchitettoniche ancora oggi utilizzate e continuamente aggiornate.

³⁵ Morselli, 1887-88, p. 14.

³⁶ Le immagini scintigrafiche si formano con la rilevazione delle emissioni di un radiofarmaco, la cui radiazione interagisce con un cristallo. L'energia ceduta viene poi trasformata in impulso elettrico, convertito a sua volta in segnali digitali analizzati da un computer.

4. Conoscenze e competenze

Queste tecniche richiedono però una valutazione complessa e articolata sul piano epistemologico.³⁷

Se come abbiamo visto la complessità della medicina nucleare è riscontrabile nel processo operativo e gestionale che porta alla produzione dei dati di *imaging*, dobbiamo riconoscere però che tale complessità si ritrova anche nelle implicazioni conoscitive che l'intero processo porta con sé.

Ritengo utile focalizzare l'attenzione su aspetti specifici sia della struttura conoscitiva che delle conoscenze e competenze proprie del medico, "vero" produttore – in questa ottica – della conoscenza trasportata dal processo di *imaging*. Fra questi ha un significato particolare, l'aspetto del *grado di invasività*, una variabile che si correla ai vari tipi di necessità della ricerca. Ad esempio in un rapporto rischio/beneficio a fronte di una patologia il rischio radiologico può essere accettato senza eccessive resistenze. In relazione poi alle caratteristiche di "informazione" veicolate dall'*imaging*, l'aspetto della *risoluzione spazio-temporale* riveste una particolare significatività soprattutto quando sia in relazione ai bisogni diagnostici.

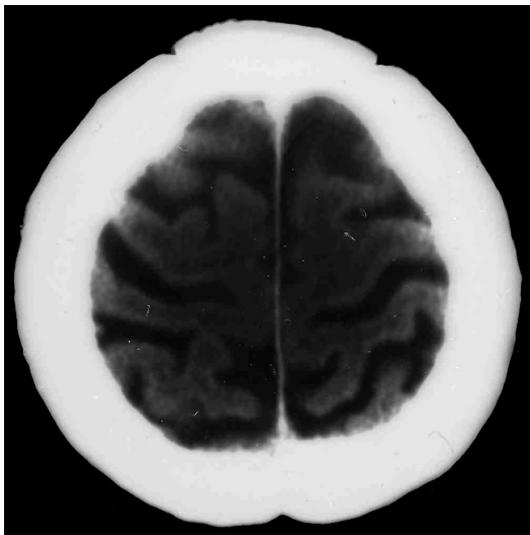
Infine risulta importante dal punto di vista epistemologico centrare l'attenzione sulla dimensione *correlazionale*. Essa riguarda la relazione tra un *dato* – ad esempio comportamentale – e l'*attività* in un determinato sistema neurale. Questo tipo di informazione – centrale nei processi di ricerca – è legata anche alla correttezza del processo. La correttezza emerge da un lato dalla qualità dell'output del processo tecnologico e dall'altro dalla precisione dell'interpretazione del referto. I parametri di rendimento legati alla "esperienza" del medico e alla "attendibilità" della strumentazione sono oggetto di continui studi. Risulta quindi centrale rendere più "affidabile" la strumentazione. Sappiamo infatti che c'è un legame diretto tra miglioramento della tecnologia e riduzione di errori interpretativo-valutativi

Gli aspetti sopra sinteticamente analizzati hanno delineato alcune delle problematiche interne al processo conoscitivo del *neuroimaging*. Da un lato abbiamo l'aspetto "tecnico" per la creazione di immagini-referto; dall'altro sappiamo che la lettura di tali immagini dipende anche dall'abilità del medico di rintracciare i dati salienti. Verrà ora analizzato il tipo di conoscenze disponibili (morfologiche e funzionali) non tanto in relazione agli aspetti tecnici e strumentali, quanto in relazione ai contributi conoscitivi che possono essere significativi sul piano euristico.

³⁷ Morabito, 2007, p. 131.

5. Conoscenze morfologiche

Uno tra i primi strumenti che permisero di indagare *in vivo* la morfologia degli organi interni del corpo umano fu la TAC. Le immagini scintigrafiche prodotte con una TAC hanno un alto grado di risoluzione spaziale e garantiscono quindi reperti significativi da un punto di vista diagnostico. Nella TAC il dato morfologico emerge in relazione ai diversi coefficienti di assorbimento della radiazione X da parte dei tessuti presenti nella scatola cranica.



Da TC CRANIO (esame diretto)

Nell'affrontare la discussione sulla valenza euristica del *neuroimaging* possiamo citare uno dei primi studi sul cervello fatti con la TAC, nel 1976³⁸. Questo venne realizzato su pazienti psichiatrici affetti da schizofrenia. Rispetto a un gruppo di controllo, i risultati tomografici evidenziarono differenze topografiche (i ventricoli cerebrali apparivano ingranditi). L'impiego di questa tecnica permise di concludere il percorso conoscitivo iniziato nell'Ottocento da Theodor Meynert. All'epoca gli strumenti di cui si avvaleva non gli avevano permesso di giungere ad una conclusione definitiva, perché non fornivano una visione strutturale di insieme per i cervelli analizzati³⁹. La tecnica di *neuroimaging*, con la possibilità di evidenziare l'insieme del cervello, offrì i dati che permisero di affrontare la correlazione fra sintomi e strutture neuronali.

³⁸ Nel 1976 lo studio fu condotto presso il *Clinical Research Centre* di Londra-Harrow dai ricercatori Eve Johnstone e Timothy Crow.

³⁹ Cfr. Shorter, 2000.

6. Conoscenze funzionali

Se le istanze conoscitive nei confronti del sistema nervoso centrale hanno avuto una presenza costante in medicina, diverse sono state le strade di acquisizione delle conoscenze ad esso relative. A partire dalla psichiatria biologica ottocentesca il percorso principe fu quello delle patologie e dell'analisi anatomica *post mortem*. Nel XX secolo si sono strutturati percorsi nuovi e diversi tra cui la possibilità dello studio della funzionalità *anche* non patologica *in vivo* e una acquisizione di dati più "completi" e significativi.

La qualità dei risultati di *imaging* ha portato ad ampliare e specificare una delle domande classiche nel campo della patologia mentale, relativa all'esistenza di correlati anatomici nella schizofrenia, facendo emergere con chiarezza come abbiamo visto, le possibilità euristiche di questo tipo di dati. Allo scopo di approfondire questa dimensione, qui di seguito viene analizzato uno studio – da considerare ormai *classico* – per la validazione di ipotesi funzionali nel linguaggio attraverso una delle tecniche di visualizzazione funzionale: la PET⁴⁰. Le prime ricerche sul linguaggio fatte da Broca⁴¹ e Wernicke⁴² hanno costituito la base per i successivi studi del linguaggio su basi neurali che hanno portato a classificare in modo sempre più articolato e specifico i relativi aspetti patologici⁴³.

In particolare Wernicke avanzò l'ipotesi che le funzioni cerebrali potessero essere trattate come strutture semplici e che le loro funzioni "complesse" derivassero da una loro connessione. Una specifica lesione avrebbe così danneggiato non tanto la funzione in sé

⁴⁰ Un settore di ricerca particolarmente significativo è costituito dai metodi di visualizzazione funzionale che permettono, grazie alle elevate caratteristiche di risoluzione spaziale e di contrasto, l'osservazione delle strutture profonde del cervello. La tomografia ad emissione di positroni – PET – viene usata per gli studi sulle funzioni cognitive e senso-motorie. Tramite PET si possono evidenziare le modificazioni locali del flusso sanguigno. Le immagini della attività cerebrale si ottengono "tracciando" modelli del flusso del sangue: quando in una parte del cervello aumenta l'attività, è lì che viene richiamato un maggior afflusso di sangue. La prova di questo emerse nel 1928 da studi compiuti su un paziente che dopo un'operazione chirurgica era rimasto senza calotta cranica, solo con il cuoio capelluto. Con lo stetoscopio fu possibile controllare che, mentre il soggetto era impegnato in un compito (lettura ad esempio), il flusso sanguigno aumentava. Quando il compito cessava, cessava anche il rumore collegato al flusso sanguigno. Il metodo di analisi della PET consiste nel contrassegnare il sangue (o il glucosio) con una "etichetta radioattiva" che consente di rilevare l'attività: il tracciante emette positroni che in interazione con gli elettroni del corpo, emettono raggi gamma rilevabili dalle apparecchiature. Un'altra tecnica di visualizzazione funzionale è la SPECT. Essa permette di misurare la perfusione cerebrale e il sistema recettoriale del Sistema Nervoso Centrale. I campi specifici di applicazione sono le malattie cerebrovascolari e quelle psichiatriche.

⁴¹ Broca, 1861, trovò una correlazione tra una lesione cerebrale (terzo posteriore della circonvoluzione frontale) e disturbi dell'articolazione del linguaggio, dimostrando così come una funzione intellettuale fosse legata ad una specifica area del cervello.

⁴² Wernicke, 1874, trovò una correlazione tra lesioni nel lobo temporale dell'emisfero sinistro e gravi deficienze a livello di comprensione del linguaggio (disturbi nella scelta corretta delle parole).

⁴³ Oggi la tassonomia delle afasie è ben più articolata rispetto ai due primi disturbi studiati e comprende: afasia di Broca (difficoltà a capire le frasi subordinate – implicante regole grammaticali); afasia di Wernicke (alterazione della capacità di comprensione e ripetizione); afasia di conduzione (alterazione della capacità di ripetizione e difetto di articolazione); afasia globale (perdita totale della comprensione); afasia transcorticale motoria (linguaggio non fluente); afasia transcorticale sensoriale (comprensione alterata).

quanto il sistema delle connessioni. Elaborò poi un modello di funzionamento del linguaggio.

Lo sviluppo delle conoscenze neurologiche portò successivamente a una concezione della funzionalità cerebrale più complessa rispetto alle prime ipotesi localizzazioniste. Nella concezione tradizionale si cercava infatti di ricondurre una funzione psichica ad un solo luogo, ad una sola struttura cerebrale.

Con i suoi studi sul linguaggio che si basavano sulle ipotesi linguistiche di Roman Jakobson, Aleksandr Luria sviluppò un nuovo modello – proponendo il concetto di “sistema funzionale”⁴⁴ – secondo il quale la funzione “emergeva” dalla interconnessione di aree corticali che svolgevano ognuna una funzione elementare.

“Sfruttando” le “nuove” possibilità di indagine offerte dalla PET, Marcus Raichle e Michael Posner⁴⁵ approfondirono questo “nuovo” concetto di sistema funzionale nel campo neurolinguistico.

Il quesito posto dai ricercatori aveva come punto di riferimento una ipotesi “storica” di base: lo schema proposto da Wernicke. In questo modello il processo linguistico attivato da parole *scritte* o *udite* era concepito sostanzialmente come seriale: dagli *input* sensoriali diversi (vista; udito) il messaggio passava ad un’area associativa, poi all’area di Wernicke (dove veniva associato al significato) e successivamente a quella di Broca (che determinava la produzione della parola). La struttura degli esperimenti sopra citati era tesa a visualizzare tramite PET i comportamenti linguistici analizzati da Wernicke: lettura, ascolto, elaborazione del significato. Raichle e Posner intendevano controllare la validità o meno dell’ipotesi funzionale di Wernicke. I risultati ottenuti confermarono solo in parte la primitiva ipotesi: gli *input* visivi se non integrati “in ingresso” dalla lettura o dalla loro pronuncia (canale uditivo), passavano “direttamente” all’area di Broca generando una risposta di eloquio. Gli *input* uditivi passavano invece all’area di Wernicke, dove venivano associati al significato e successivamente passavano all’area di Broca. I canali di *input* visivo e uditivo erano quindi diversi e afferivano a strutture cerebrali differenti. Raichle e Posner conclusero che le funzioni *leggere, ascoltare, riflettere sul significato* corrispondevano ad aree cerebrali diverse e potevano essere processate sia “in serie” che “in parallelo”⁴⁶. Questo risultato sperimentale confermò l’importanza del *neuroimaging* e permise di evidenziare la sua reale portata euristica. Possiamo dunque affermare che le ricerche di *imaging* su aspetti così fondamentali come le abilità cognitive concorrono a

⁴⁴ Cfr. Luria, 1967.

⁴⁵ Cfr. Posner, 1988.

⁴⁶ Cfr. Posner, 1988. Cfr. anche Kandell, 2007, p. 14.

definire non solo problemi altamente specifici ma anche concezioni sul cervello sempre più ampie e articolate⁴⁷ e che portano a considerare ogni funzione psichica come una complessa architettura neurale.

Come fa notare Eric Kandel:

È perciò altamente improbabile che le basi nervose di una qualsiasi funzione cognitiva come il pensiero, la memoria, la percezione e il linguaggio possano venir comprese concentrando l'attenzione soltanto su una certa regione cerebrale senza considerare i rapporti di quella regione con altre.⁴⁸

7. Discussione

Le riflessioni qui sviluppate ci conducono ad un inquadramento critico del *neuroimaging*. In primo luogo dobbiamo rilevare come non sia esaustivo parlare di questa metodologia conoscitiva riducendola solo ad aspetti "tecnologici". In relazione ai quesiti che erano stati posti all'inizio sulla presenza di momenti di continuità e di innovazione si può affermare che queste dimensioni risultano ambedue consistenti. In effetti l'analisi storica ci ha condotto a riconoscere nelle istanze conoscitive del *neuroimaging* problemi e domande comuni a discipline che già nel corso del XIX secolo avevano iniziato ad affrontare lo studio del cervello su basi biologiche. Possiamo pertanto riconoscere al *neuroimaging* un diritto di appartenenza ad una tradizione di ricerca che oggi confluisce, in termini di singole discipline e di metodologie di ricerca, nell'ampio sistema delle neuroscienze. Al contempo, si possono riconoscere momenti specifici di innovazione. Un primo momento è relativo alla "potenza" del processo tecnologico. Questa deriva sia dalle caratteristiche tecniche delle macchine a disposizione sia dalla tecnologia di processo basata sui rapporti interdisciplinari fra operatori con specialità diverse e sulle procedure di controllo. Un secondo aspetto riguarda quella che possiamo definire la potenza euristica dei referti di *neuroimaging*, come abbiamo visto nelle ricerche citate.

Non ritengo comunque che la domanda epistemologica sul *neuroimaging* possa esaurirsi in relazione alle due dimensioni di "continuità" e "innovazione". L'analisi sopra svolta ha condotto infatti a considerare epistemologicamente significativa la dinamica di correlazione fra gli interrogativi che le istanze di ricerca pongono o possono porre, e la tipologia di "risposte" tecnologico/processuali offerte attualmente dal *neuroimaging*.

⁴⁷ Lo schema oggi proposto si sviluppa su tre sistemi ben definiti: sistema di implementazione; sistema di mediazione; sistema concettuale. Per un approfondimento Cfr. Kandel, 2007.

⁴⁸ Kandel, 2007, p. 1162

Emerge così un quadro conoscitivo costituito da un impianto sistemico ampio al quale concorrono anche i due poli sopra evidenziati, ma che risulta essere ancora più “ricco” in termini di possibilità conoscitive.

Molti autori riconoscono che i percorsi conoscitivi offerti dalle tecniche di visualizzazione costituiscono praticamente un “ponte” fra discipline diverse che affrontano problematiche comuni ma che hanno statuti disciplinari diversi. Se si intende il concetto di ponte non tanto come momento legato all’utilizzo di una tecnologia condivisa, ma come un vero e proprio fattore conoscitivo, possiamo cogliere con più precisione la dimensione epistemologica del *neuroimaging*. Questo viene quindi assunto in tutte le sue dimensioni conoscitive – qui richiamate in termini di conoscenze e competenze – e inteso come parte costitutiva delle dinamiche concrete del lavoro di ricerca neuroscientifico.

Matteo Borri

Bibliografia

- Aglioti Salvatore M., Franco Fabbro (2006), *Neuropsicologia del linguaggio*, Bologna, Il Mulino.
- Bechtel William, Abrahamsen Adele, Graham George (2004), *Menti, cervelli, calcolatori. Storia della scienza cognitiva*, Roma-Bari, Laterza.
- Borri Matteo (2001), *Il gioco come "momento ermeneutico"*, Bari, Laterza.
- Boschi Filippo (1977), *Psicologia della lettura. Studi e ricerche sulle capacità del lettore abile e maturo*, Firenze, Giunti - Barbera.
- Boschi Filippo, Aprile Luigi, Scibetta Ida (1992), *Le parole e la mente. Studi e ricerche sullo sviluppo della competenza lessicale nel ragazzo*, Firenze, Giunti.
- Brodmann Konstantin (1909), *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf grund des Zeelenbaues*, Leipzig, Johannes Ambrosius Barth.
- Camaioni Luigi (2004), *La teoria della mente. Origini, sviluppo e patologia*, Roma-Bari, Laterza.
- Conforti Rosario (a cura di) (2006), *La psicoanalisi tra scienze umane e neuroscienze*, Salerno, Rubbettino.
- Ellenberger Henri F. (1976), *La scoperta dell'inconscio*, Torino, Bollati Boringhieri.
- Ferrier David (1876), *The Function of the Brain*, London, Smith Elder.

- Feyerabend Paul K. (2005), *Contro il metodo. Abbozzo di una teoria anarchica della conoscenza*, Milano, Feltrinelli.
- Foucault Michel (1998), *Nascita della clinica. Una archeologia dello sguardo medico*, Torino, Einaudi.
- Gall Franz Joseph (1822-23), *Sur les fonctions du cerveau et sur celles de chacune de ses parties*, Paris.
- Gazzaniga Michael S., Richard B. Ivry , George Mangun (2005), *Neuroscienze cognitive*, Bologna, Zanichelli.
- Grmek Mirko D. (1998), *Storia del pensiero medico occidentale. 3. Dall'età romantica alla medicina moderna*, Roma-Bari, Laterza.
- Groppo Mario, Antonietti Alessandro (1995), *Nuove teorie della mente*, Milano, Vita e Pensiero.
- Kandel Eric R., Schwart James H., Jessel Thomas M. (2007), *Principi di neuroscienze*, Milano, Casa Editrice Ambrosiana.
- Keel Othmar (2007), *La nascita della clinica moderna in Europa 1750-1815. Politiche, istituzioni e dottrine*, Firenze, Polistampa.
- Luciani Luigi, Seppilli Giuseppe (2002), *Le localizzazioni funzionali del cervello*, Firenze, Giunti.
- Luria Aleksandr R. (1967), *Le funzioni corticali superiori nell'uomo*, Firenze, Giunti.
- Mado Proverbio Alice, Zani Alberto (2000), *Psicofisiologia Cognitiva*, Roma, Carocci.
- Maffioli Lorenzo, Mazzuca Nicola, Bombardieri Emilio (2006), *Il libro bianco della Medicina Nucleare in Italia*, Associazione Italiana di Medicina Nucleare ed Imaging Molecolare.
- Meotti Franca (1971). "La nascita della psicologia scientifica", in Geymonat Ludovico (a cura di) *Storia del pensiero filosofico e scientifico V*, Milano, Garzanti.
- Meyer Johann Christoph Andreas (1779). *Anatomisch-physiologische abhandlung vom gehirn, Rückmark, und Ursprung der Nerven*. 4to. Berlin-Leipzig: Decker.
- Morabito Carmela (1996), *La cartografia del cervello. Il problema delle localizzazioni cerebrali nell'opera di David Ferrier, fra fisiologia, psicologia e filosofia*, Milano, Franco Angeli.
- Morabito Carmela (2007), *La mente nel cervello. Un'introduzione storica alla neuropsicologia cognitiva*, Roma-Bari, Laterza.
- Morselli Enrico, *Lezioni di psichiatria tenute per cura d'alcuni studenti*, Genova, Manoscritto.
- Morselli Enrico (1887-88), *Lezioni dettate dal Chiar. Prof. E. Morselli raccolte per cura degli*

studenti U. Filippa e C. Mussatti, Lit. F.lli Bertero.

- Morselli Enrico (1894), *Esame psicologico degli alienati*, Milano, Francesco Vallardi.
- Posner Michael I. , et al. (1988). "Localization of Cognitive Operation in the Human Brain." Science **240**: 1627-1631.
- Robins Robert H. (1997), *La linguistica moderna*, Bologna, Il Mulino.
- Shallice Tim (1990), *Neuropsicologia e struttura della mente*, Bologna, Il Mulino.
- Shorter Edward (2000), *Storia della Psichiatria*, Milano-Parigi-Barcellona, Masson.
- Telfener Umberta, Casadio Luca (a cura di) (2003), *Sistemica. Voci e percorsi nella complessità*, Torino, Bolalti Boringhieri.
- Thompson Richard F. (1997), *Il cervello. Introduzione alla neuroscienze*, Bologna, Zanichelli.
- Umiltà Carlo (1999), *Manuale di neuroscienze*, Bologna, Il Mulino.
- Young Robert M. (1970), *Mind, Brain and Adaptation in the Nineteenth Century*, Oxford, Clarendon Press.
- Zunini Giorgio (1972), *Psicologia*, Brescia, Morcelliana.

Recensione

NEUROETICA. LA MORALE PRIMA DELLA MORALE

Raffello Cortina, Milano 2008

Laura Boella



Il testo di Laura Boella affronta in maniera bilanciata la definizione dei confini e dell'oggetto di ricerca di una nuova disciplina che prende il nome di neuroetica. Due paiono i cardini di questo cammino: in primo luogo questi nuovi studi vengono analizzati a vari livelli e ciò consente al lettore di avvicinarsi ai temi trattati attraverso riflessioni che intersecano vaste aree del sapere, dall'ambito clinico alla riflessione filosofica, dalla psichiatria alla fisica sanitaria.

In secondo luogo risalta la scelta formale del volume, diviso in due parti: la prima relativa ad un inquadramento storico e concettuale delle neuroscienze, la seconda che affronta intersezioni specifiche tra neuroscienze ed etica.

Il volume, nella prima parte, è dedicato ai problemi che emergono dalle tecnologie in campo biomedico.

La scienza sta affinando la sua sensibilità per la persona nella sua totalità, con una storia e caratteristiche individuali uniche, uno stile, delle scelte e un progetto di vita. [...] Anche lo studio del sistema nervoso centrale, effettuato a livello di biochimica e biologia molecolare, sta portando i ricercatori ad affacciarsi su territori che un tempo erano di esclusiva competenza della filosofia, della morale o della religione. (XI)

In questo ampio campo di conoscenze, Boella indica come particolarmente rilevante il campo neuroscientifico che, se da un lato è strutturato come un settore altamente specializzato, dall'altro ha catturato l'attenzione del grande pubblico.

Il fenomeno decisivo nel campo specifico delle neuroscienze è infatti l'inserzione ormai diretta degli studi sul cervello nella vita e nei comportamenti delle persone. (XV)

Nell'analisi delle neuroscienze possiamo evidenziare l'attenzione posta sul tema del *neuroimaging*.

Affrontando i temi dei correlati neurali della personalità, l'autrice esamina criticamente questo tipo di conoscenza in base a temi etici quali il libero arbitrio e la responsabilità. In seguito essa esamina anche i vari punti critici, presenti in letteratura, come ad esempio i problemi legati alla risoluzione spazio temporale o le implicazioni del passaggio dall'analisi di una struttura all'analisi di una funzione (cfr. p. 29).

Si arriva così a evidenziare uno specifico rischio interpretativo laddove la ricezione dei dati di *imaging* è più orientata al "dove" avviene una certa attivazione cerebrale – quasi una "frenologia *hightech*" – piuttosto che a sviluppare la ricerca verso l'intero processo di elaborazione dell'informazione. Passando poi da considerazioni centrate sul piano scientifico a riflessioni proiettate verso l'etica, l'autrice coglie alcuni punti importanti. Un primo punto critico riguarda la significatività della conoscenza del *neuroimaging* che risulta basata sì sull'opera dell'osservatore, ma anche fondata su un'interpretazione statistica dei dati (cfr. p. 32).

Altro punto delicato è la specificità dell'osservazione in quanto essa si riferisce ad un preciso esperimento e non a situazioni di vita reale; in più anche quel dato non è "assoluto" ma emerge dal "rapporto" fra due situazioni (metodo della sottrazione).

Infine l'autrice affronta un altro aspetto della discussione etica: la responsabilità verso i soggetti che si prestano ad esperimenti per il possibile presentarsi di "risultati non previsti" (cfr. p. 33). La visione critica dei vari fattori che attualmente caratterizzano l'impiego del *neuroimaging* porta a considerare questo metodo di indagine e i risultati conoscitivi che da esso emergono come oggetti di un impegno critico forte e costante.

La trasmissione della conoscenza al pubblico implica che gli stessi ricercatori siano consapevoli delle sfide epistemologiche e filosofiche insite nel fatto che le neuroscienze cognitive stanno rapidamente affrontando fenomeni che appartengono all'ambito sociale e morale. (p. 38).

Nella seconda parte del libro viene invece sviluppata un'intensa riflessione sui collegamenti possibili fra le ricerche neuroscientifiche, che riguardano interrogativi complessi – ad esempio quello delle emozioni, trattato partendo dagli studi di Damasio – , e le intersezioni possibili con le riflessioni etiche.

Nell'ultimo capitolo un contributo particolarmente significativo alla riflessione neuroetica viene dall'analisi delle dinamiche dei rapporti fra conoscenze diverse in relazione allo specifico tema dell'empatia, tema già ampiamente trattato dalla stessa Boella in altri lavori.

[...] il percorso degli studi sperimentali e della riflessione filosofica sull'empatia appare un esempio – particolarmente avanzato – della ricerca, in cui è impegnata la neuroetica, di ponti e passaggi fra strutture comuni, generali, del funzionamento organico e contenuti e qualità individuali del comportamento, fra funzionamenti automatici e attività intenzionali e coscienti. (p. 104)

In sintesi l'autrice, riconoscendo il valore del grande sviluppo delle neuroscienze, si dichiara contraria ad un percorso che riduca la "complessità dell'esperienza morale a meccanismi cerebrali" (p. 106).

Matteo Borri

Indice

Premessa

Introduzione

PARTE PRIMA

LA NEUROETICA

1. *Brave neuro world*

2. La neuroetica entra nel "mondo reale;

La neuroetica "tra" neurobiologia e filosofia

L'etica delle neuroscienze;

Neuroscienze e problemi del sé, della responsabilità e della condotta umana ;

3. Che cosa fanno "vedere" le immagini della fMRI o della PET?;

Le tecniche di visualizzazione cerebrale; Derivare conclusioni *bold* dal segnale BOLD?

Neuroimaging e neuroetica;

PARTE SECONDA

LA MORALE PRIMA DELLA MORALE

4. La morale prima della morale;

Le precondizioni della capacità morale;

Moral sense: un problema attuale;

5. Esercizi morali per il cervello;

Cervello ed emozioni;

Intuizioni cerebrali;

Utilitaristi "irrazionali";

Esiste una grammatica morale universale?;

Libertà del volere, responsabilità: le neuroscienze cambiano tutto o niente?;

6. La riscoperta dell'empatia; L'empatia nasce nel cervello?;

Storia di un enigma; L'empatia oggi: un problema;

Le basi neurofisiologiche dell'empatia oltre i neuroni specchio;

Gli esperimenti sull'empatia del dolore;

CONCLUSIONI;

Bibliografia.

Recensione

IL VOLTO E LA PAROLA. PSICOLOGIA DELL'APPARENZA

Quodlibet, Macerata, 2007

Felice Cimatti



Nel saggio *Il volto e la parola. Psicologia dell'apparenza*, Felice Cimatti si propone un obiettivo preciso: immaginare una scienza (non esplicativa) che sfugga alla contrapposizione tra l'interno dell'uomo (pensieri, sentimenti, desideri) e l'esterno (il linguaggio, i gesti, il comportamento). La proposta dell'autore è dunque quella di superare la distinzione interno-esterno così carica, a suo dire, di pregiudizi, per fondare una psicologia del "soggetto superficiale".

Questa espressione sembrerebbe alludere ad un altro soggetto che starebbe sotto di esso, in profondità. In realtà, come una figura dipinta è superficiale, in piena vista, così il soggetto è del tutto implicato in ciò che appare. Quello che interessa è la "superficie" del soggetto, il visibile, perché in questo apparire c'è tutta la vita dell'individuo. Per spiegare questo concetto Cimatti usa l'immagine del nastro di Möbius, un nastro che mette in vista tutto quello che c'è e ciò che non c'è: manca un "dietro", ma senza "dietro" evidentemente non può esserci un "davanti". L'idea di questo libro è che la soggettività umana sia qualcosa del genere: essa esiste, per gli altri e per sé, se è manifesta, se è pubblica; altrimenti non c'è nessuna soggettività, perché non c'è nulla di misterioso nell'"io". La "superficie" è il "luogo" in cui si svolge la vita umana, un luogo non individuabile nello spazio, ma che in realtà c'è: si tratta della lingua che parliamo comunemente ed in cui pensiamo. Come viene sottolineato nel titolo del saggio, il volto appare nella parola, perché è nella parola che si manifestano i fenomeni specifici della vita umana: il volto parla.

In un dialogo costante con Wittgenstein, Cimatti sviluppa questa idea secondo cui il "soggetto superficiale" vive nella lingua, in particolare nel "gioco linguistico" che, a sua

volta, non ha lati interni, è reale solo nella sua applicazione: il gioco c'è soltanto se qualcuno lo gioca.

Il soggetto non è separabile dal gioco linguistico in cui partecipa, quello in cui chi parla dice "io" ad un "tu". L'individuo, non distinguendosi dalle pratiche della lingua in cui si manifesta come soggetto, non ha un'essenza, proprio come l'essenza dello stormo è tutta nel fatto che quello stormo si compone in una certa temporanea forma: l'apparenza coincide con la sostanza.

L'autore si sofferma anche su un altro aspetto della questione: la lingua è un'entità naturale, biologica in senso forte, perché rappresenta l'ambiente in cui vive, ma allo stesso tempo non può essere spiegata come si spiegano gli altri oggetti biologici. Essa infatti non ha una propria materia, come invece un animale o un vegetale, perché è la "superficie senza spessore" dove pensiero e suono si incontrano in un segno per diventare significato e significante. Questa combinazione produce quella forma in cui si manifestano i giochi linguistici e, dal momento che il soggetto umano è tale perché partecipa a questi giochi, sarà allora un soggetto superficiale com'è superficiale l'ambiente cioè la lingua in cui soltanto può vivere. Cimatti precisa che questo individuo ha ovviamente anche un corpo, fatto di carne e sangue, ma è solo in virtù della logica che questo corpo può dare ragioni del proprio comportamento, e non semplicemente reagire alle sollecitazioni esterne che lo colpiscono.

È importante sottolineare che, secondo l'autore, questa visione della soggettività ha delle conseguenze per quanto riguarda la psicoanalisi: l'analisi è interminabile perché il soggetto non è definito come una sostanza stabile, ma si sviluppa in una continua ridefinizione di sé, articolandosi nella lingua e nella storia.

Nell'ultima parte del saggio, Felice Cimatti si concentra sul problema dell'altro, del rapporto fra individui, sostenendo che una soggettività radicalmente singolare non ha senso perché essa si costruisce nello spazio pubblico delle relazioni con gli altri: è solo perché ci sono gli altri che può esserci l'"io". In quest'ottica l'etica è immanente al gioco linguistico, è nell'altro con cui dialoghiamo, se accettiamo la possibilità di parlare con lui e di ascoltarlo.

Stefano Liccioli

Indice

Introduzione

Il filosofo e le forme

Chi studia un oggetto che è *soltanto* apparenza?

Nel «gioco linguistico»

Wittgenstein «seguace di Freud»

Rinunciare all'essenza

La natura della lingua

2.1 Le lingue (e i parlanti) come *epifenomeni*

2.2 Misteri per gli umani, misteri per i ratti

2.3 «Si gioca questo gioco linguistico»

2.4 Cos'è una lingua?

Chi parla?

3.1 La «langue» come fenomeno originario

3.2 L'*ambiente naturale* del soggetto umano

3.3 La forma della «massa amorfa»

3.4 L'*io* che dice "io"

Storia naturale dell'interiorità

4.1 Cos'è un "dolore"?

4.2 Dall'*essere* all'*avere* un dolore

4.3 Biolinguistica di un'emozione, "la gelosia"

4.4 Un corpo superficiale in una natura artificiale

L'impensabile: corpo ed autoscienza

5.1 Quando finisce l'analisi?

5.2 Il «fattore quantitativo»

5.3 La «roccia basilare» e la fine dell'analisi

5.4 Senso e sensi possibili

Il gioco e la libertà

6.1 C'è un gioco "giusto"?

6.2 Regole e «regolarità»

6.3 Il volto dell'*altro*

6.4 Per una vita «felice»

Riferimenti bibliografici

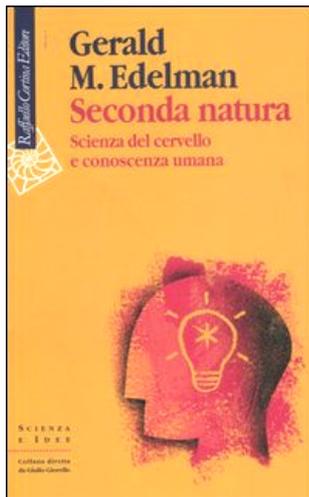
Recensione

SECONDA NATURA

SCIENZA DEL CERVELLO E CONOSCENZA UMANA

Raffaello Cortina, Milano, 2007

Gerald M. Edelman



Seconda Natura rappresenta l'ennesima versione del tentativo di sciogliere i numerosi intrighi che ruotano attorno alla comprensione scientifica della coscienza. Lo scopo del nuovo lavoro di Edelman è quello di rendere comprensibile la natura del rapporto che intercorre tra i fenomeni che contraddistinguono la coscienza umana, intesa qui come lo spazio cognitivo nel quale si collocano l'insieme delle percezioni, dei ricordi e degli atteggiamenti individuali di un soggetto, (quella che l'autore chiama la nostra *seconda natura*), e la realtà che ci circonda.

Per Edelman la possibilità di raggiungere una reale comprensione scientifica della coscienza passa attraverso la definizione di un quadro concettuale unitario, all'interno del quale tradizioni di pensiero per consuetudine separate vengono definitivamente a ricongiungersi. Individuando in Vico e Dilthey da una parte e in Galileo e Cartesio dall'altra i capiscuola di quelle correnti – l'una umanistica, l'altra scientifica – che fino a oggi hanno contribuito a propagare l'idea di un'insanabile separazione tra scienze dell'uomo e scienze della natura, Edelman pone le basi per una riunificazione dei due domini proponendo una lettura *riduzionista* del rapporto mente corpo. Al centro di tutti i lavori di Edelman si colloca il tentativo di tracciare un percorso teorico e sperimentale indirizzato alla comprensione scientifica della coscienza, anche "a dispetto" del suo carattere indissolubilmente soggettivo (pag. 6).

Riportando la coscienza nell'alveo delle scienze naturali, Edelman intende ricomporre quello iato esistente tra saperi sperimentali e saperi umanistici che, ormai da quattro secoli, condiziona il nostro panorama culturale.

Due sono secondo Edelman i principali aspetti della coscienza che una teoria naturale deve poter *spiegare*: la *varietà* dei fenomeni che contraddistinguono la sfera del pensiero

umano e la *regolarità* delle correlazioni cerebrali che essa manifesta. Scartata l'ipotesi rappresentata da una concezione computazionale della mente, la proposta di Edelman si contraddistingue per l'attenzione dimostrata nei confronti di una prospettiva *evoluzionistica*. L'originaria idea darwiniana di selezione naturale viene, infatti, a ricoprire un ruolo esplicativo di primo piano, risultando indispensabile per la comprensione di fenomeni complessi come quello rappresentato dalla comparsa della coscienza umana.

Condizione preliminare, affinché si possa procedere alla definizione di un'adeguata concezione dei rapporti mente corpo, è rappresentata secondo Edelman dalla messa a punto di un modello epistemologico in grado di far emergere il ruolo che in essa rivestono fondamentali aspetti neuro-biologici. Ciò si configura come il tentativo di condurre a compimento quell'idea di epistemologia naturalizzata proposta da Quine durante la prima metà del novecento e connotata poi da Edelman attraverso una marcata impronta biologica, basata sullo studio dei processi neuronali posti alla base della cognizione umana (pag. 51). Secondo Edelman, un'epistemologia siffatta deve fondarsi sulla teoria dell'evoluzione, ovvero sullo studio dei processi di selezione naturale che hanno condotto alla comparsa di quei meccanismi neuronali preposti all'acquisizione, elaborazione e conservazione della conoscenza.

Per Edelman le capacità cognitive umane e con esse le proprietà ascrivibili alla coscienza sono riconducibili al risultato di un lungo processo evolutivo riguardante gli aspetti morfologici e funzionali del cervello, quest'ultimo inteso come *sistema* in grado di selezionare al suo interno la formazione di strutture neurali adatte alle diverse condizioni ambientali. Seppur introdotto con maggiori dettagli in un'opera precedente (Edelman, 1995), il *darwinismo neurale* di Edelman sfrutta i concetti fondamentali del pensiero popolazionista al fine di fornire una visione evolutiva delle proprietà emergenti a partire dalla complessa organizzazione cerebrale (Edelman è in primo luogo un immunologo il cui tentativo consiste nell'applicare concetti chiave della propria disciplina al contesto delle scienze cognitive). Al centro di tale visione si colloca l'analisi dei fenomeni di "degenerazione", ovvero le sorprendenti capacità adattative legate alla presenza di strutture morfologicamente diverse, ma dotate di funzioni biologiche analoghe. Secondo l'autore proprio la diversificazione delle caratteristiche strutturali conferirebbe una maggiore adattabilità alle diverse condizioni ambientali, permettendo nel caso specifico del sistema nervoso una capacità di elaborazione adeguata rispetto alle variabili informazioni provenienti dal mondo esterno. Il processo di selezione di reti specifiche di

gruppi neurali si trova dunque, secondo Edelman, alla base dei processi cognitivi che regolano il comportamento di ogni animale dotato di sistema nervoso complesso, permettendo così la comparsa delle categorie e dei sistemi di valore che disciplinano la facoltà di giudizio. Tra le strutture neurali che l'evoluzione ha selezionato nell'uomo, particolare importanza è attribuita da Edelman a determinate specie di circuiti – definiti *rientranti* – rappresentati ad esempio dalla via talamo-corticale. Tali strutture rappresenterebbero la risposta naturale utilizzata dal cervello per consentire la coordinazione e la sincronizzazione tra aree funzionali diverse, permettendo inoltre la comparsa di raffinate facoltà discriminatorie.

Proprio l'emergenza di tali funzioni discriminatorie, legate alla presenza di circuiti rientranti, permetterebbe la comparsa delle peculiari caratteristiche qualitative che contraddistinguono la coscienza. In altre parole, i *qualia* non sarebbero altro che il riflesso dei molteplici stati di interazione in cui si possono trovare aree diverse del cervello, connesse tra di loro proprio per mezzo di circuiti di tipo rientrante (cfr. pp. 36 e 83).

Nel tentativo di prevenire le comuni reazioni avverse a una concezione epifenomenica della coscienza, Edelman cerca di salvaguardarne il ruolo evolutivo attribuendo a questa una funzione informativa (p. 90). Il breve accenno connesso alla definizione di tale funzione non chiarisce però chi sia il destinatario delle informazioni e nemmeno perché siano coinvolti stati qualitativi coscienti e non solo inconsapevoli.

Seconda Natura non rappresenta certo una delle più complete opere di Edelman, sebbene il presente lavoro possa essere considerato un eccellente compendio degli importanti contributi forniti dal premio Nobel in anni passati sull'individuazione dei meccanismi cerebrali coinvolti dagli stati di coscienza umani, nonché nel tentativo di applicare alle neuroscienze cognitive il modello interpretativo evuzionistico.

Lo sforzo di rendere evidente il ruolo naturale della coscienza riesce ad Edelman solo in parte. Se da un lato il collegamento tra le capacità discriminatorie emerse nel cervello e le capacità discriminatorie proprie degli stati di coscienza apre la strada alla formulazione di varie ipotesi attorno al ruolo evolutivo di quest'ultima, tale ruolo appare ancora poco chiaro. Ciò lascia in sospeso il tentativo di fornire un'adeguata comprensione dei rapporti che intercorrono tra gli stati cerebrali e quelle manifestazioni del tutto peculiari e personali della coscienza.

Indice

Prefazione

Ringraziamenti

Introduzione

1. L'arco Galileiano

2. Coscienza, corpo e cervello

3. SeleZIONISMO. Un requisito indispensabile per la coscienza

4. Dall'attività cerebrale alla coscienza

5. L'epistemologia e i suoi malcontenti

6. Un approccio basato sul cervello

7. Forme di conoscenza. Il divorzio tra la scienza e le discipline umanistiche

8. Sanare la frattura

9. Causalità, illusioni e valori

10. Creatività. Il gioco tra specificità e varietà

11. Stati anormali

12. Congegni basati sul cervello. Verso un artefatto cosciente

13. Seconda natura. La trasformazione della conoscenza

Indice analitico

Recensione
MENTI MORALI,
LE ORIGINI NATURALI DEL BENE E DEL MALE

Il Saggiatore, Milano 2007

Marc D. Hauser



La nostra tradizione filosofica ha sempre trattato la morale come una 'scienza dello spirito' (o forse è meglio dire come la scienza dello spirito per eccellenza), come un oggetto di studio estraneo alle metodologie scientifiche.

In *Menti Morali*, Hauser tenta di rompere con questa tradizione esponendo i primi risultati ottenuti, attraverso un possibile percorso di ricerca teso allo sviluppo di modelli cognitivi in grado di fornire un resoconto naturalista della nostra facoltà morale. Questa 'scienza della morale' non punta a definire quale sia il corretto contenuto dei giudizi

morali (X è giusto, Y è sbagliato) ma ad indagare le strutture cognitive che ci permettono di elaborare tali giudizi e di deliberare sulla base di questi. Una scienza che ha per oggetto la nostra facoltà morale deve in particolare rendere conto delle seguenti questioni:

- che cos'è: quali dinamiche fra le funzioni cerebrali portano all'emergere dei nostri giudizi morali;
- come si sviluppa: in che maniera il bambino apprende un sistema di valori morali, quali strutture permettono e vincolano tale apprendimento e come l'esperienza determini tale sviluppo;
- come si è evoluta: quali sono state le spinte adattive che hanno permesso l'evoluzione delle funzioni cognitive che permettono i nostri giudizi morali.

Le ricerche presentate da Hauser si inseriscono dichiaratamente all'interno di quel settore delle Scienze Cognitive, di impostazione radicalmente chomskiana, interessato all'individuazione di strutture innate che permettono e vincolano lo sviluppo delle funzioni cognitive manifeste. Le sorti del progetto di ricerca delineato da Hauser pertanto sono prima di tutto vincolate alle sorti di questa impostazione, mirata allo studio della cognizione e sul cui valore si è già molto dibattuto.

Secondo l'impostazione presentata, alla base della forte variazione interculturale nelle norme morali, esiste una grammatica morale universale (cioè un insieme di principi generali ma astratti, privi di un contenuto necessario) che vincola le possibilità di variazione stessa e che permette a ogni persona – formatasi in uno specifico ambiente – di sviluppare e adottare un peculiare sistema di valutazione morale. Quando formuliamo un giudizio morale, lo facciamo spesso inconsciamente, utilizzando un sistema di principi che opera inconsciamente e che ci è inaccessibile: la morale affonda le sue radici nella nostra biologia, ed è questa a stabilire la nostra gamma di atteggiamenti possibili.

Il libro è organizzato in tre parti: *1. Dichiarazioni universali, 2. Senso innato, 3. Codice in evoluzione.*

La prima parte evidenzia come il nostro comportamento morale sia legato a un senso di equità e giusto scambio – variabile da cultura a cultura ma sempre presente – e come la risoluzione di conflitti morali avvenga per mezzo di scelte intuitive, non immediatamente giustificabili: possiamo giustificare le nostre scelte solo a posteriori, riferendoci a processi consci sia emotivi che razionali. Gli esempi portati mostrano come i modelli classici non possano rendere conto del nostro comportamento. Sia che ci riferiamo a principi utilitaristi, a principi razionali o alla sola emotività, è necessario ammettere l'esistenza di una serie di regole inconscie che disciplinano i nostri giudizi morali.

Nella seconda parte sono presentate varie funzioni cognitive che giocano un ruolo determinante nella formulazione dei giudizi morali: la facoltà di identificare quali corpi siano agenti e quali eventi siano azioni intenzionali; la capacità di attribuire intenzioni in maniera efficiente sulla base del solo comportamento manifesto; l'empatia; l'autoriflessione; l'immaginazione di possibili scenari futuri...

Funzioni come queste, che secondo l'impostazione chomskiana si sviluppano in conformità a strutture innate, ci pongono in condizione di valutare responsabilità (riconoscendo non solo che un agente ha compiuto un'azione la quale ha avuto certi effetti positivi/negativi, ma anche se ha fatto ciò intenzionalmente o meno), di crearci aspettative sul comportamento degli altri o di ritardare la gratificazione in vista di maggior vantaggi futuri; tutti ingredienti necessari per lo sviluppo di un senso morale.

Nella terza parte vengono invece prese in considerazione quali tra le capacità umane – identificate sopra come fondamentali per lo sviluppo della facoltà morale – siano presenti anche negli animali: emozioni, capacità di attribuzione di stati intenzionali agli altri agenti,

autoriflessione e pazienza sono, in gradi diversi, presenti in molti animali evoluti, anche se solo gli esseri umani sembra abbiano questo 'insieme' completo e ben sviluppato. L'autore prende poi in esame quali possibili pressioni selettive possano aver portato all'evoluzione della nostra struttura morale, ponendo il problema se tali pressioni siano 'attualmente' sempre presenti o meno e se il nostro 'senso morale' abbia ancora un'efficacia adattiva.

Il volume, per quanto si rivolga ad un ambiente accademico, conserva un valore divulgativo, all'interno di una disciplina che sta muovendo i primi passi. Presenta grande chiarezza espositiva anche se si sente la mancanza di una problematizzazione dell'oggetto di studio, soprattutto da un punto di vista fondativo. La possibilità dell'utilizzo di una impostazione chomskiana per lo studio di una disciplina come la morale viene assunta acriticamente, e il concetto di 'morale', nato e sviluppatosi finora in ambiti non scientifici, viene inserito in un contesto scientifico, senza che venga tematizzata alcuna necessità di una sua ridefinizione. L'impressione che si ottiene è che le ricerche siano un po' 'dispersive', sia riguardo l'oggetto di ricerca e le nozioni teoriche di riferimento, sia per i metodi di ricerca utilizzati. Le prove empiriche chiamate a sostegno spaziano dai test psicologici e dai giochi di ruolo (i mezzi più frequentemente citati) all'*imaging* cerebrale, passando dalle osservazioni di antropologi ed etologi agli studi sugli ormoni. Non viene inoltre ben delineato se e come i risultati provenienti da aree tanto diverse possano coordinarsi.

Oltre a possibili problemi relativi all'integrazione delle varie discipline messe in campo, alcune argomentazioni appaiono meno solide di quanto si vorrebbe. Il valore scientifico dell'identificazione di alcuni atteggiamenti intenzionali negli animali – ad esempio – ottenuto interpretando esperimenti sul loro comportamento, può apparire dubbio; almeno nella forma in cui è stato presentato.

Una pecca di questa nascente disciplina sembra essere il non aver ancora raggiunto risultati scientifici, ma essere ancora ai primi passi di una disciplina tutta in divenire.

Nonostante questi dubbi, bisogna ammettere che Hauser non presenta le sue argomentazioni come risultati scientifici ben consolidati, ma sembrerebbe ben conscio del valore poco più che ipotetico ed euristico di queste prime proposte da una disciplina tutta in divenire.

Un discorso a parte va fatto per il valore del progetto di ricerca: le scienze cognitive in genere stanno avendo un impatto interdisciplinare forte, andando ad influenzare e ad alimentare dibattiti nei campi più disparati. Se questo approccio rivoluzionario alla morale

si evolverà in una disciplina solida e dai risultati interessanti, l'impatto culturale sarà potenzialmente enorme, andando ad influenzare anche ambiti tradizionalmente poco ispirati dalle dinamiche scientifiche.

Giovanni Casini

Indice

Sommario: Prologo

Voci virtuose

Cap.1 - Che c'è di sbagliato?

PARTE PRIMA - Dichiarazioni universali

Cap.2 - Giustizia per tutti

Cap.3 - Grammatiche della violenza

SECONDA PARTE - Senso innato

Cap.4 – L'organo morale

Cap.5 – Istinti leciti

TERZA PARTE – Codice in evoluzione

Cap.6 – Radici del giusto

Cap.7 – Principi primi

Epilogo – L'impulso giusto

Recensione
MIND TIME.
IL FATTORE TEMPORALE NELLA COSCIENZA.

Raffello Cortina, Milano 2007

Benjamin Libet



Il libro di Benjamin Libet ruota intorno alla problematica della relazione fra cervello ed esperienza conscia. La questione che Libet si pone è relativa a come le attività fisiche delle cellule nervose del cervello possano produrre fenomeni non fisici, cioè le esperienze soggettive coscienze e quale sia la relazione tra il *fisico*, il cervello, il *mentale*, e la consapevolezza sensoriale del mondo esterno.

La prospettiva dalla quale Libet affronta queste problematiche è quella sperimentale. I suoi studi sono focalizzati sulle relazioni temporali fra gli eventi neurali e l'esperienza cosciente. Il libro è incentrato sugli esperimenti dell'autore che portano alla scoperta che la decisione di agire avviene *prima* di quella relativa al compiere la determinata azione. Ciò porta con sé implicazioni importanti anche per la questione del libero arbitrio. Il nostro cervello sembra 'sapere' che vogliamo compiere una certa azione con un anticipo di circa mezzo secondo. La consapevolezza è dunque in ritardo su una parte del nostro cervello e quindi la convinzione di aver deciso liberamente e autonomamente di compiere una certa azione sembra configurarsi come un'illusione. Così Libet spiega il fattore temporale della coscienza:

Se si batte un dito sul tavolo si ha la percezione dell'evento come se accadesse in "tempo reale". Soggettivamente, cioè, si sente il tocco nello stesso momento in cui il dito entra in contatto con il tavolo. Ma i nostri risultati sperimentali conducono con ogni evidenza a una scoperta sorprendente, direttamente contrapposta alla nostra personale intuizione e percezione: il cervello ha bisogno di un periodo relativamente lungo – fino a circa mezzo secondo – per attivarsi in modo appropriato e indurre la consapevolezza dell'evento! L'esperienza cosciente o la consapevolezza del dito che tocca il tavolo appare così solo dopo che le attività cerebrali si sono attivate per produrne la consapevolezza¹.

Secondo Libet la grande maggioranza degli impulsi sensoriali rimane inconscia perché non sviluppa nel cervello una durata sufficientemente lunga di attività neurali appropriate. Dunque il cervello sembra iniziare i propri processi circa 550 msec prima dell'atto 'liberamente volontario'. Questi esperimenti hanno una risonanza fondamentale per le implicazioni che portano alla questione del libero arbitrio. Gli studi di Libet mostrano infatti che «l'avvio della preparazione che culmina in un movimento liberamente volontario nasce nel cervello in modo inconscio, e che precede la consapevolezza cosciente della volontà e dell'intenzione di "agire adesso" di circa 400 msec o più»². Secondo Libet le iniziative non coscienti preposte alle azioni volontarie devono essere interpretate come un «borbottio inconscio del cervello»³. La volontà cosciente cioè seleziona quali iniziative possono continuare fino a produrre un'azione e quali devono invece essere vietate o bloccate in modo da non manifestarsi. Libet critica fortemente il determinismo e l'idea che la natura deterministica del mondo fisicamente osservabile possa spiegare le funzioni e gli eventi coscienti soggettivi. L'idea di Libet è che l'esistenza del libero arbitrio sia un'opinione migliore rispetto alla sua negazione in base alla teoria deterministica delle leggi naturali. Libet conclude il capitolo dedicato alla questione del libero arbitrio citando Singer:

Il più grande dono che l'umanità abbia ricevuto è la libera scelta. È vero che abbiamo dei limiti per farne uso. Ma quella piccola libertà di scelta che abbiamo è un dono così grande e potenzialmente così ricco che la vita vale la pena di essere vissuta proprio per esso⁴.

Chiara Erbosì

Note

1. B. Libet, *Mind Time. Il fattore temporale nella coscienza*, Raffaello Cortina Editore, Milano, 2007, p. 37.
2. Ivi, p. 145.
3. Ivi, p. 153.
4. Ivi, p. 160.

Indice

Prefazione all'edizione italiana

(Edoardo Boncinelli)

Prefazione *(Stephen M. Kosslyn)*

Introduzione. Come sono arrivato a scrivere questo libro

1. Il problema
2. Il ritardo nella nostra consapevolezza sensoriale cosciente
3. Funzioni mentali cosce e inconsce
4. L'intenzione di agire: esiste il libero arbitrio
5. Il campo mentale cosciente

6. Che cosa significa tutto questo?

Bibliografia

Indice Analitico

Recensione
**IMMAGINI DELLA MENTE.
NEUROSCIENZE, ARTE, FILOSOFIA.**

Raffaello Cortina Editore, Milano 2007

G. Lucignani e A. Pinotti (cura)



Questo volume raccoglie i contributi di scienziati, filosofi e storici dell'arte, nel tentativo di articolare un discorso utilmente critico riguardo al rapporto che le discipline di loro competenza possono intrattenere con la questione estetica. Già nell'introduzione dei curatori – Giovanni Lucignani e Andrea Pinotti –, viene sottolineata una doppia accezione del termine 'estetica' a cui è necessario volgere una costante attenzione; semplificando, si può parlare di estetica come “indagine intorno alla conoscenza sensibile”¹ oppure come teoria del gusto e del bello. Il quesito centrale affrontato da tutti i saggi è se le più recenti conquiste delle neuroscienze possano approdare ad una conoscenza scientificamente verificabile dei processi che si attivano nel cervello dell'uomo durante un'esperienza estetica.

Ogni autore propone la propria teoria avvalendosi dei mezzi a sua disposizione, non di rado integrando conoscenze tecniche e riflessioni speculative. La costitutiva frammentarietà del testo pone in evidenza quello che, in ultima analisi, risulta essere lo scopo di una simile antologia: al di là del consenso (o del dissenso) che si può accordare a ciascuna ipotesi, si ha la possibilità di entrare in contatto con la complessità tanto articolata di un dialogo interdisciplinare intorno ad uno specifico argomento.

Il volume, diviso in tre parti, offre inoltre al lettore, che non si può presupporre onnisciente, anche una serie di cenni – in verità non sempre di agilissima lettura- riguardo allo stato di avanzamento delle più moderne tecniche neuroscientifiche, cercando di porlo nella condizione di poter comprendere il contenuto di pagine che rimarrebbero altrimenti oscure.

1G. Lucignani A. Pinotti, *Immagini della mente*, 2007, p. XVI.

Non va senz'altro tralasciato uno degli aspetti, non solo tematicamente rilevante, ma anche portatore di un sicuro fascino, con cui più spesso gli autori si confrontano, ossia il rapporto con l'opera d'arte. Da questo incontro emergono le questioni più controverse: su questo terreno si acutizzano e rendono più visibili sia i punti di contrasto tra le diverse prospettive di analisi, sia le argomentazioni interne ad uno specifico ambito.

Il testo, complessivamente, offre una buona panoramica di studi e aiuta a comprendere il significato della parola 'neuroestetica'; Alfredo Civita in conclusione del suo saggio *Estetica, filosofia e neuroscienze*, esprime, in merito a questo tipo di ricerca, un giusto dubbio che dovrebbe essere assunto nella forma di un atteggiamento cauto nell'accostamento a simili argomenti:

Il rischio che vediamo profilarsi nell'ambito della neuroestetica, e in generale delle neuroscienze, è di partorire progetti di ricerca che pur essendo non razionali, almeno allo stato attuale delle conoscenze, siano tuttavia in grado di guadagnarsi l'approvazione dell'opinione pubblica non in virtù del loro contenuto scientifico, che è appunto non razionale, ma grazie alle apparenze, ossia grazie all'intenso potere seduttivo che la scienza e la tecnologia del nostro tempo sono in grado di esercitare.²

Scilla Bellucci

Indice

Autori

Introduzione (Giovanni Lucignani, Andrea Pinotti)

PARTE PRIMA - ARTE E CERVELLO

² Alfredo Civita, *Estetica filosofia e neuroscienze*, in Op. cit. , p. 185.

1. Spettatori attivi (Edoardo Boncinelli, Giulio Giorello)
2. Empatia, movimento ed emozione (David Freedberg)
3. I diversi sentieri della memoria e l'arte visiva
(Lamberto Maffei)
4. Neurologia dell'ambiguità (Semir Zeki)
5. Arte e immagine corporea:
considerazioni neurologiche (Stefano F. Cappa)

PARTE SECONDA - FILOSOFIA E NEUROSCIENZE

6. Immagine e pensiero (Elio Franzini)
7. Estetica, filosofia e neuroscienze (Alfredo Civita)

PARTE TERZA - PROBLEMI DI NEUROESTETICA

8. Rappresentazione dello schema corporeo. (Gabriella Bottini, Martina Gandola)
9. Negligenza spaziale. (Anna Berti, Alessia Folegatti)
10. Musica e neuropsicologia. Punti di incontro (Marco D. Poli, Danilo Spada)
11. Effetti culturali nella lettura in soggetti normali e nella dislessia. (Eraldo Paulesu, Uta Frith)

Recensione

**PSICOLOGI PER NATURA INTRODUZIONE AI
MECCANISMI COGNITIVI DELLA PSICOLOGIA INGENUA**

Carocci, Roma, 2007

Cristina Meini



Il *senso comune* è un metro di paragone di fondamentale importanza, non solo nella vita di tutti i giorni, ma anche in ambiti di ricerca come quello psicologico o filosofico. La propensione a interpretare i propri comportamenti e quelli altrui ricorrendo all'attribuzione di *stati mentali* rappresenta un fenomeno sul quale s'interrogano psicologi e filosofi, oltre ad antropologi e neuroscienziati. Il libro di Cristina Meini propone una ricostruzione del recente dibattito che ruota attorno al tema della *psicologia ingenua*, rilevandone il carattere interdisciplinare e mostrandone le principali questioni che rendono appassionante questo settore di ricerca.

Al fine di illustrare i temi attorno a cui ruota il libro, è opportuno prendere le mosse dalla definizione di ciò che s'intende per *psicologia ingenua*. Comprendere i comportamenti propri e altrui attraverso l'attribuzione di stati mentali è una capacità che si colloca alla base del sistema sociale in cui viviamo. Attribuire a noi stessi e a coloro che ci circondano stati percettivi, desideri, credenze, significa svolgere un'operazione *meta-rappresentazionale*, fornire cioè un'interpretazione di come le persone, compresi noi stessi, formino una propria rappresentazione del mondo.

Non tutti gli esseri umani possiedono però le stesse capacità interpretative, né tutte le fasi dello sviluppo sono contraddistinte dal possesso di un'identica facoltà meta-rappresentazionale. Uno dei temi centrali nel dibattito concernente la psicologia ingenua riguarda il tentativo di mettere a punto una teoria generale della mente, in grado di comprendere le discrepanze osservabili tra soggetti. Dalla ricerca attorno ai processi cognitivi implicati nelle dinamiche sociali, alla discussione riguardo la natura degli stati mentali, per giungere fino alla frontiera delle neuroscienze, molte sono le

competenze e gli interessi che vengono a sovrapporsi quando si parla di psicologia ingenua.

Per comprendere il ruolo epistemologico rivestito dalle indagini attorno alla psicologia ingenua, occorre prendere le mosse dagli esiti e dalle questioni che si legano al classico test delle *credenze false*. Di tale esperimento ne esistono diverse versioni, in quella messa a punto da Baron Cohen, Leslie e Frith, a un bambino viene presentata una scena nella quale i protagonisti sono delle bambole: la prima posa una biglia in un luogo mentre la seconda sposta la biglia all'insaputa della prima. Il test si considera superato se il bambino comprende che la prima bambola crede erroneamente che la biglia si trovi nel luogo in cui l'ha lasciata. Il bambino cioè deve essere in grado di capire che, in funzione della situazione che si è creata, la prima bambola si è formata una rappresentazione *falsa* della realtà. Generalmente i risultati confermano che i bambini di età inferiore a tre anni tendono a fornire una risposta errata. Nell'età successiva ai tre anni, una parte di loro inizia a rispondere correttamente. I bambini affetti da sindrome da autismo continuano a riscontrare difficoltà nel superamento del test anche in età più avanzata.

Da cosa dipende il corretto conferimento di uno stato mentale? Quali processi biologici sono coinvolti nella discriminazione e nell'attribuzione di stati mentali a se stessi e ad altri? Nel corso degli anni, attorno ai risultati del test delle *credenze false* e alle sue versioni analoghe, si sono radunate diverse prospettive interpretative, facendo della psicologia ingenua un crocevia di fondamentale importanza per il più generale dibattito sui rapporti tra mente e corpo. In particolare, due sono le principali correnti che si affrontano in questo contesto, la così detta *teoria della teoria del mentale (TT)* e la *teoria della simulazione (ST)*. Nel suo libro Cristina Meini mostra i tratti salienti di entrambe le prospettive, avendo premura di metterne in risalto le rispettive articolazioni interne.

Alla base dello sviluppo delle concezioni raccolte sotto l'etichetta *TT* si colloca l'assunto per cui "la psicologia ingenua si fonda sul possesso di una conoscenza teorica, ovvero un sapere organizzato relativo alla mente nostra e altrui" (p. 30). In funzione di tale condizione teorica, il riferimento a particolari enti non osservativi come gli stati mentali, acquista un peculiare ruolo esplicativo rispetto all'osservazione e alla pianificazione di atteggiamenti comportamentali. Analogamente a quanto avviene all'interno delle teorie scientifiche stati mentali come la credenza, il desiderio, o altre attribuzioni intenzionali assolvono uno specifico ruolo intra-teorico e

delineano relazioni causali tra stimoli sensoriali e stati mentali; fra stati mentali e stati comportamentali; tra stati mentali differenti.

Per comprendere l'analogia che c'è tra psicologia ingenua e teorie scientifiche nel contesto di *TT*, l'autrice sottolinea le posizioni sostenute da A. Gopnik e J. Perner. Per entrambi, lo sviluppo di una psicologia ingenua si lega al progressivo aumento di conoscenze cui si accompagna la normale crescita di un bambino. Stando a questa interpretazione, proprio come uno scienziato modifica o sostituisce le ipotesi che costituiscono la sua teoria con altre sempre più raffinate ed esatte, così il bambino sostituirebbe col tempo le assunzioni che costituiscono la sua personale teoria psicologica ingenua, seguendo un percorso basato sulla rivedibilità in funzione dei successi e dei fallimenti nei quali essa incorre.

Restando sempre all'interno del contesto contrassegnato dall'approccio *TT*, a una concezione costruttivista come quella appena delineata, se ne affianca una contraddistinta dalla vicinanza col modularismo di J. Fodor. All'interno di tale prospettiva, la psicologia ingenua viene delineandosi come sapere 'implicito', immagazzinato all'interno di speciali moduli cognitivi che si attivano gradualmente nel corso dello sviluppo. Troverebbero in questo modo spiegazione i dati riguardanti il mancato superamento del test della falsa credenza, sia nei bambini di età inferiore ai quattro anni, sia nei soggetti affetti da sindrome da autismo. I primi risulterebbero, infatti, dotati di un sistema cognitivo prematuro, mentre i secondi risulterebbero deficitari rispetto al possesso di una struttura cognitiva adibita allo sviluppo di una normale teoria degli stati mentali.

In generale, una prospettiva come quella che si delinea nel contesto *TT* pone interessanti questioni rispetto al ruolo evolutivo svolto dalla presenza, altamente diffusa, di una teoria ingenua degli stati mentali. Che cosa ha permesso la selezione di un particolare modulo cognitivo basato sul ricorso esplicativo a stati mentali? Quali vantaggi adattativi può aver apportato? Domande, queste, a partire dalle quali è possibile procedere verso un livello d'analisi naturalistico.

Alternativo al modello *TT* è l'approccio interpretativo noto come "teoria della simulazione" (*TS*). In linea generale, secondo *TS*, la psicologia ingenua non coinciderebbe con una qualche forma di sapere teorico, ma con la tendenza naturale a immaginare e simulare i comportamenti altrui (p. 63). In altre parole, il ricorso a una qualche capacità interpretativa rispetto ai comportamenti osservati si baserebbe principalmente sulle capacità immaginative del soggetto, indipendentemente dalla

formulazione d'ipotesi teoriche riguardo lo stato della propria mente, o di quella altrui.

Sviluppata negli ultimi decenni del secolo scorso, la teoria della simulazione si richiama al dibattito relativo alla comprensione delle *funzionalità empatiche*, rimandando per questo ad antecedenti filosofici illustri, tra i quali Vico, Kant e Dilthey. Nel contesto di TS risulta particolarmente importante lo sviluppo di capacità di elaborazione cognitiva non consapevoli (off-line), attraverso le quali dar luogo a processi di simulazione che mettano in grado di valutare prospettive diverse, senza per questo perdere la centralità del riferimento al proprio punto di vista "in prima persona". Nell'illustrare le prerogative di TS, Cristina Meini, distingue tra una versione moderata e una versione radicale, la prima riconducibile ai lavori di Alvin Goldman, la seconda presente invece nei lavori di Robert Gordon.

Nella versione moderata di Goldman, il ricorso a processi simulativi diviene parte di un più generale ragionamento abduttivo, all'interno del quale emerge il ruolo svolto dalle proprietà qualitative legate all'*effetto* che può fare l'immaginarsi in un certo tipo di stato. A partire dalla formazione di un codice introspettivo personale, il soggetto procede verso la comprensione degli altrui stati mentali attribuendo genuini concetti intenzionali prima a se stesso, per trasferirli in un secondo momento alla persona interpretata.

Mentre nella concezione di Goldman permane lo spazio per un contributo teorico alla formazione dell'atteggiamento psicologico ingenuo, collocabile nella definizione del codice introspettivo e nella sua applicazione a casi in terza persona, nella variante radicale di TS data da Gordon tale spazio viene escluso. Secondo questa versione le attribuzioni mentali non passerebbero attraverso un percorso immaginativo-introspettivo, ma risulterebbero piuttosto una proiezione *diretta* e in *prima persona* nei panni del soggetto da interpretare, estromettendo in questo modo il carattere teorico-inferenziale legato al processo introspettivo definito da Goldman.

Una possibile soluzione al contrasto tra TT e TS sembra poter provenire dalla ricerca neurobiologica. Il libro di Cristina Meini si conclude con una rassegna di alcune delle più significative prospettive di ricerca neurologiche coinvolte nel dibattito concernente la natura della psicologia ingenua. Conoscere i meccanismi fisiologici alla base dei particolari atteggiamenti cognitivi umani può, infatti, risultare determinante

se si vuole comprendere quale percorso naturale seguiamo nell'interpretare e nel pianificare i comportamenti nostri e altrui.

Tra i diversi meccanismi biologici coinvolti, particolare importanza riveste la recente scoperta di neuroni specchio anche nell'essere umano. Secondo alcune interpretazioni, la presenza di sistemi specchio in specifiche aree del cervello, attivi sia quando si compie un'azione sia quando la si osserva, permetterebbe di cogliere immediatamente il significato di un comportamento, rievocandone in maniera soggettiva, ad esempio, le caratteristiche intenzionali. Al di là delle evidenti funzionalità simulative possedute dai neuroni specchio, Cristina Meini sottolinea la presenza d'interpretazioni teoriche alternative, secondo le quali i sistemi specchio non sarebbero altro che strumenti impiegati per calcolare inferenze sulle azioni future, e non meccanismi adibiti alla comprensione *immediata*.

Il dibattito relativo alla natura della psicologia ingenua resta pertanto aperto. Comprendere i processi coinvolti nell'interpretazione dei comportamenti risulta un obiettivo di primaria importanza nel contesto delle ricerche indirizzate all'analisi dei rapporti tra mente corpo. Questioni legate alla definizione di una posizione epistemologica per enti come gli stati mentali, risultano infatti strettamente connesse alla comprensione delle modalità con cui ingenuamente attribuiamo credenze, sensazioni, intenzioni a noi stessi e alle persone che ci circondano.

Silvano Zipoli Caiani

Indice

Introduzione

1. La psicologia ingenua: che cos'è, come si sviluppa
 - Un po' di terminologia
 - Tempi di sviluppo: il test delle credenze false
 - Prima delle credenze false
 - Psicologia ingenua e pragmatica
 - Quando la psicologia ingenua non c'è
2. La teoria della teoria
 - Teoria: che cos'è?

- Teorie e funzionalismo
- Tipi di teorie
- La teoria della teoria del piccolo scienziato
- La teoria della teoria modularistica
- Il modulo della psicologia ingenua
- Perché abbiamo un sistema cognitivo specializzato per leggere la mente?

3. Le teorie della simulazione

- I meccanismi cognitivi della simulazione
- Attenzione condivisa, credenze false, autismo
- Teoria della simulazione o teorie della simulazione?
- La teoria della simulazione moderata
- La teoria della simulazione radicale
- Conclusioni

4. La dimensione interpersonale della soggettività

- Liotti: la coscienza
- Siegel: la mente razionale

5. Le basi della psicologia ingenua nel cervello

- I neuroni specchio
- Altri sistemi specchio: l'analisi delle emozioni e delle sensazioni
- Neuroni specchi o e teoria della simulazione
- Aree puramente percettive
- Altri sistemi neurali coinvolti nell'interpretazione di azioni complesse
- Conclusioni

Bibliografia

Recensione

LA MENTE ETICA

Codice, Torino 2006

Michael Gazzaniga



In questo libro l'autore analizza il rapporto tra etica e neuroscienze. In particolare l'indagine è rivolta alla bioetica e alle questioni morali degli ultimi anni derivate dai progressi della medicina e della ricerca scientifica. L'autore si sofferma su problemi che definisce di *neuroetica*, cioè su quelle problematiche della bioetica che riguardano il rapporto tra il cervello e il sistema nervoso centrale. Il primo tema che Gazzaniga affronta è l'indagine sullo *status* etico dell'embrione. Stabilire infatti l'inizio dello *status* etico ha implicazioni fondamentali sull'aborto, sulla fecondazione in vitro, sulla clonazione terapeutica e sulla ricerca sulle cellule staminali. I neuroscienziati che studiano il cervello – l'organo al quale dobbiamo la vita cosciente – sembrerebbero in grado di determinare lo *status* etico di un embrione o di un feto in base alla presenza o all'assenza del materiale biologico che alimenta la vita mentale. Ma se la neurobiologia è chiara, la neuroetica si trova in grande difficoltà quando cerca di tradurre fatti scientifici in risposte a questioni etiche e morali. Gazzaniga ripercorre con precisione il lungo e delicato cammino del cervello, dallo sviluppo fetale fino alla vita cosciente, arrivando a sostenere che «assegnare un identico status etico a una palla di cellule di quattordici giorni e a un bambino prematuro è una forzatura concettuale. Sostenere che essi siano identici è un puro atto di credenza personale»¹. Esiste una corrente di pensiero secondo cui la vita inizia al concepimento. «Secondo la tesi della continuità, un ovulo fecondato procederà fino a diventare persona. Dunque esso merita i diritti di un individuo perché indubbiamente qui ha inizio la sua vita. Se non siamo propensi a suddividere gli eventi successivi dello sviluppo, allora questa tesi diventa inevitabile: o ci crediamo o non ci crediamo»². Secondo questa teoria, al feto e all'embrione, in virtù del fatto che potrebbero diventare adulti, va sempre garantito uno status etico uguale a quello di un essere umano "già nato". Ma – sostiene Gazzaniga – «un ovulo e uno spermatozoo non sono un essere umano. Nemmeno lo è un embrione fecondato: richiede un utero e almeno sei mesi di gestazione e di sviluppo, di crescita e di formazione dei neuroni e di

duplicazione cellulare per diventarlo»³. Molti esperti di bioetica sostengono, dunque, che bisogna attribuire all'embrione uno status etico intermedio. Secondo questi autori si deve prendere in considerazione una serie di marcatori dello sviluppo come punti che delimitano l'inizio della vita:

Come ha evidenziato il President's Council on Bioethics nella pubblicazione *Monitoring Stem Cell Research*, un primo marcatore si manifesta al quattordicesimo giorno. Secondo taluni esso demarca il punto dopo il quale diventa impossibile la divisione gemellare, in cui l'individualità dello zigote è dunque consolidata. Un secondo marcatore è la formazione del sistema nervoso. A riguardo, vengono però sostenute date differenti, comprese tra la formazione della stria primitiva (quattordicesimo giorno) e la capacità di reagire a stimoli nocivi (ventitreesima settimana)⁴.

Secondo l'autore, un ovulo fecondato o un grumo di cellule prive di cervello non meritano lo status etico che attribuiamo ad un neonato o ad un adulto. «Il mero possesso del materiale genetico, che diventerà un essere umano, non ne produce uno»⁵.

Una descrizione puramente genetica della nostra specie non descrive un essere umano, il quale rappresenta un ben altro livello di organizzazione, diverso da un semplice embrione, come quest'ultimo lo è da un ovulo o da uno spermatozoo. È la dinamica tra geni e ambiente che crea un essere umano. In realtà, quasi tutti noi siamo disposti ad attribuire uno status speciale a un'entità che si sviluppa molto prima della sua nascita, ma certamente non prima che in quell'entità sia comparso un cervello⁶.

Un'altra questione affrontata da Gazzaniga nel libro è relativa al problema di come definire la 'fine della coscienza'. «Un simile indicatore sarebbe utile per stabilire quando interrompere il sostegno delle funzioni vitali, quando onorare le volontà da vivo e persino come crearle, e, più provocatoriamente, quando attuare l'eutanasia»⁷.

Sempre nell'ottica delle neuroscienze, l'autore affronta il problema del libero arbitrio:

Il cervello, un'entità fisica soggetta alle regole del mondo fisico, determina la mente. Ma il mondo fisico è determinato e dunque lo sarà anche il nostro cervello. Se quest'ultimo è determinato, ed è l'organo necessario e sufficiente per la mente, allora vanno ancora sciolti questi interrogativi: anche i pensieri che scaturiscono dalla mente sono determinati? E il libero arbitrio, che a noi sembra di vivere in prima persona, è una pia illusione? E, nel caso in cui lo sia, dovremo forse rivedere il concetto di responsabilità personale delle nostre azioni?⁸

Gazzaniga ritiene che non si debba affatto rinunciare al concetto di responsabilità personale, distinguendo tra *cervello*, *mente* e *persona* e concludendo che noi siamo liberi

e responsabili delle nostre azioni, mentre il cervello non lo è. Anche se il cervello è un organo automatico, governato da regole e determinato, le persone sono agenti responsabili liberi di prendere le proprie, personali, decisioni.

La responsabilità personale è un concetto pubblico, che esiste nel gruppo e non nell'individuo: se voi foste soli al mondo, il concetto di responsabilità personale non esisterebbe. La responsabilità è un concetto vostro verso le azioni altrui e che gli altri hanno verso le vostre. Il cervello è determinato; le persone (più di un singolo essere umano) seguono delle regole vivendo insieme e il concetto di libertà d'azione scaturisce da quell'interazione⁹.

La nostra libertà si fonda dunque in ultima istanza sull'interazione col mondo sociale. Le neuroscienze possono spiegare poco sulla responsabilità personale perché essa è una costruzione umana; non è attribuita ai cervelli ma alle persone; esiste nelle regole di una società e non nella struttura del cervello.

Chiara Erbosi

Note

- 1) Gazzaniga, M. S., *La mente etica*, Edizioni Codice, Torino, 2006, p. 9.
- 2) Ivi, p. 10.
- 3) Ivi, p. 12.
- 4) Ivi, p. 16.
- 5) Ivi, p. 18.
- 6) *Ibidem*
- 7) Ivi, p. 27.
- 8) Ivi, p. 86.
- 9) Ivi, p. 87.

Indice

Ringraziamenti

Prefazione

Parte I. Neuroetica del ciclo vitale

Capitolo 1

Quale status etico per l'embrione?

Capitolo 2

L'invecchiamento del cervello

Parte II. Il potenziamento del cervello

Capitolo 3

Alchimie genetiche per migliorare il cervello

Capitolo 4

Allenare il cervello

Capitolo 5

Plasmare un cervello intelligente

Parte III. Libero arbitrio, responsabilità personale e diritto

Capitolo 6

Me l'ha ordinato il cervello!

Capitolo 7

Pensieri antisociali e diritto alla privacy

Capitolo 8

Il cervello non è un registratore

Parte IV. La natura delle credenze morali e il concetto di etica universale

Capitolo 9

Le credenze nel cervello

Capitolo 10

Verso un'etica universale

Indice Analitico

Recensione

LA MENTE SOCIALE

Laterza, Roma-Bari, 2005

Massimo Marraffa e Cristina Meini



Poco più di due anni fa veniva pubblicato per i tipi di Laterza *La mente sociale* di Massimo Marraffa e Cristina Meini. Il libro analizza e accetta una variante moderatamente massiva della teoria modulare della mente, e la assume per spiegare la comunicazione umana.

L'opera è divisa in due parti: la prima, di carattere introduttivo, presenta i concetti chiave del funzionalismo, spiegando inoltre la tesi della modularità della mente e proponendone un'interpretazione liberalizzata. La seconda presenta risultati sperimentali favorevoli all'ipotesi che la capacità di interpretare il comportamento di un agente in termini di stati mentali sia resa possibile da uno specifico modulo cognitivo. In essa si analizzano i rapporti tra tale modulo della psicologia ingenua e un altro modulo, deputato alla comprensione del linguaggio, in base al quale è possibile rileggere in chiave cognitiva gli studi di pragmatica della comunicazione iniziati da Grice.

Scendendo a un maggior dettaglio, il primo capitolo presenta l'approccio alla mente del funzionalismo computazionale: la mente è l'insieme dei processi di elaborazione dell'informazione, intermedi tra la stimolazione percettiva e la risposta motoria, identificabili con le computazioni di una macchina di Turing. La tesi *modularista* dell'architettura cognitiva aggiunge a questa visione l'idea che tali processi siano organizzati in sottosistemi funzionalmente distinti. La complessità delle funzioni cognitive superiori può allora essere ricondotta all'interazione di processi semplici che manipolano simboli di natura puramente sintattica.

Il secondo capitolo è dedicato a un'analisi della nozione di modulo. Partendo dalle sue ricerche di linguistica generativa, Chomsky identifica i moduli con la conoscenza teorica specifica di un determinato dominio da essi implementata (moduli *epistemic*). Secondo Fodor i moduli presentano invece anche delle caratteristiche computazionali precise, tra le quali l'*accessibilità limitata* (i sistemi cognitivi centrali possono accedere solo alle

rappresentazioni in uscita, non a quelle intermedie), e l'*incapsulamento informativo* (le operazioni di un modulo si avvalgono unicamente di informazioni altamente specifiche contenute al suo interno). Essi sono inoltre selezionati geneticamente e localizzati in circuiti neurali determinati. Tale fatto li rende danneggiabili in modo selettivo, e quindi studiabili attraverso un'analisi del decadimento delle funzioni cognitive in seguito a lesioni cerebrali. Gli studi di neuropsicologia non adottano tuttavia la concezione fodoriana dei moduli, rifiutandone in particolare l'*incapsulamento informativo*. Gli autori tentano quindi di rivedere tale nozione, per mostrare come i risultati della neuropsicologia possano integrarsi in una versione "liberalizzata" della teoria modulare.

Nel terzo capitolo vengono analizzate diverse varianti della tesi della natura modulare della mente. Sulla base di una visione olistica della cognizione centrale, Fodor restringe la modularità ai sistemi percettivi e motori, riducendo così la psicologia computazionale a una parte della più ampia psicologia cognitiva. Altri autori (Sperber 1994, Pinker 1997) estendono invece la modularità anche alla cognizione centrale. Secondo un'ipotesi biologico-evoluzionista, i moduli che compongono la mente si sarebbero sviluppati nel corso dell'evoluzione per risolvere problemi differenti che si ponevano ai nostri antenati del Pleistocene. La cognizione centrale può quindi essere vista come il risultato dell'integrazione di moduli diversi. Se fossero completamente incapsulati, tale integrazione non sarebbe possibile: Marraffa e Meini si pongono così il problema di come la tesi dell'*incapsulamento informativo* possa essere allentata. Introducono per questo la *teoria della pertinenza* di Sperber e Wilson (1986/1995). In base a tale teoria, la mente è costantemente alla ricerca di nuove informazioni sull'ambiente. Dal momento che la loro elaborazione richiede risorse cognitive, i processi di pensiero si affidano normalmente a euristiche sommarie, veloci e poco dispendiose. I sistemi sono quindi incapsulati nel senso che il costo elevato da pagare per l'acquisizione alcune informazioni le esclude di fatto dal processo di pensiero, nonostante, in linea di principio, esse siano comunque accessibili. La limitatezza delle risorse cognitive costringe quindi il pensiero alla creatività, caratteristica esclusa dalla caratterizzazione fodoriana dei moduli.

Il quarto capitolo passa quindi ad analizzare la capacità cognitiva di attribuire stati mentali ad altri agenti. Una serie di risultati interessanti – in particolare il gruppo dei cosiddetti *test della falsa credenza* (Baron-Cohen, Leslie, Frith 1985) - fa ritenere a molti studiosi che tale capacità sia frutto del funzionamento di un modulo mentale, accrescitivo delle conoscenze teoriche che, nel complesso, definiscono la psicologia del senso comune

(*ipotesi innatista della teoria della mente*). Le ipotesi sulla specifica struttura e funzionamento di tale modulo sono diverse, e gli autori dedicano numerose pagine a una loro rassegna e valutazione; al di là delle differenze tecniche, essi ritengono che i risultati sperimentali costituiscano un'evidenza importante a favore della teoria della teoria.

Il quinto capitolo sottolinea infine il valore della facoltà metapsicologica per altre due abilità cognitive: la comunicazione e l'acquisizione del linguaggio. Riguardo alla prima, Marraffa e Meini sostengono l'interesse di una lettura psicologizzante della pragmatica griceana, e presentano la proposta di Sperber e Wilson. Tali autori vedono nella comprensione del significato il risultato del lavoro di un modulo cognitivo apposito, che opererebbe massimizzando il rapporto tra sforzo cognitivo e informazioni utili attese dalla comprensione di un enunciato. Viene altresì sostenuto che tale modulo è sempre attivo e utilizza informazioni che vengono elaborate dal modulo della teoria della mente. Il dibattito sull'esistenza del modulo della comunicazione è comunque aperto, e Marraffa e Meini analizzano come i bambini apprendano le parole per portare argomenti a favore. Molti (per esempio Bloom, 2000) sostengono al riguardo l'impossibilità di risolvere la conoscenza delle parole nella semplice associazione tra la percezione dell'oggetto e l'ascolto del nome corrispondente. Numerose evidenze sperimentali provano invece che la fissazione del significato dei nomi si appoggia sul riconoscimento dell'oggetto d'attenzione dell'adulto, attraverso la condivisione del suo sguardo. La capacità di interagire con l'altro in quanto individuo psicologicamente connotato sembra quindi imprescindibile per il corretto apprendimento del linguaggio; questo depone a favore della connessione tra capacità metapsicologiche e linguistiche, sebbene non costituisca una prova definitiva dell'esistenza del modulo della comunicazione.

In conclusione gli autori sostengono che la loro analisi della modularità permette di "difendere una nozione di modulo non vacua"¹ anche a livello della cognizione centrale; a conferma di questa ipotesi ricordano gli studi che hanno mostrato che "anche a livello concettuale esistono meccanismi specializzati, che godono delle proprietà caratteristiche dei moduli"², tra i quali il modulo della psicologia ingenua. Una valutazione dell'opera impone tuttavia di domandarsi fino a che punto la loro argomentazione sia stringente. A tal riguardo noto lo scarso valore che è stato dato nel libro a proposte contrarie alla teoria della teoria, come l'ipotesi simulazionista (Gordon 1986, Goldman 1989). La citazione di risultati sperimentali è accompagnata, tra l'altro, da interpretazioni fuorvianti: è il caso

¹ Pag. 178.

² *Ibid.*

dell'esperimento della fotografia datata³, che viene utilizzato come prova parziale dell'esistenza di una capacità di comprendere le rappresentazioni pubbliche indipendente dalla capacità di comprendere rappresentazioni private (come le credenze). Il risultato è valido, ma ben poco significativo senza la presenza di un risultato opposto, basato su un campione di soggetti capaci di comprendere rappresentazioni private, ma non rappresentazioni di tipo pubblico.

In secondo luogo, la spiegazione evoluzionista della selezione del modulo della teoria della mente può essere accattivante, ma non deve essere sopravvalutata. I vantaggi evolutivi di una determinata abilità cognitiva non hanno valore probatorio: sono delle ipotesi che poggiano su una ricostruzione mitica della vita dei nostri antenati e come tali andrebbero considerate. Inoltre che valore ha sostenere la natura modulare della psicologia ingenua data una definizione estremamente generale della modularità? La definizione fodoriana faceva della natura sintattica dei dati uno dei suoi punti di forza; parlare di un modulo della psicologia ingenua che acquisisce dati in forma semantica richiede invece di mostrare in che modo l'informazione sensoriale può acquisire tale formato. Questa domanda rimane inesa nel testo. Il libro non porta quindi giustificazioni sufficienti ad avvalorare l'esistenza di un modulo della mente.

Il problema dell'esistenza di un modulo corrispondente a una capacità cognitiva è ancora più grande nei confronti del modulo della comunicazione. Nell'introduzione viene affermato che le "inferenze pragmatiche sono il prodotto di un modulo [...] su cui poggia la più generale capacità di fare psicologia ingenua"⁴. Poche prove vengono date tuttavia a sostegno della tesi e, se nelle conclusioni gli autori si augurano di aver reso "l'idea di quanto l'area di ricerca di cui ci stiamo occupando sia promettente"⁵, questo è diverso dall'aver mostrato evidenze significative, che non emergono dalla letteratura sperimentale. Il libro costituisce quindi una introduzione alle teorie della modularità della mente e all'ipotesi della natura teorica della psicologia ingenua. Allo stesso modo delle teorie di cui si fa sostenitore, tuttavia, non riesce a fugare i dubbi di quanti non sono disposti ad abbracciarle senza prima averle vagliate attentamente.

Marco Fenici

³ Pag. 124.

⁴ Pag. VI.

⁵ Pag. 180.

Bibliografia

1. Baron-Cohen, Leslie, Frith (1985) *Does the autistic child have a 'Theory of mind'?*, Cognition, 21, 37-46.
2. Bloom (2000) *How children learn the meaning of words*, MIT, Cambridge, MA.
3. Goldman (1989) *Interpretation psychologized*, Mind & Language, 4, pp. 161-185.
4. Gordon (1986) *Folk psychology as simulation*, Mind & Language 1, pp. 158-171.
5. Pinker (1997) *How the mind works*, Norton, New York.
6. Sperber (1994) *The modularity of thought and the epidemiology of representations*, in Hirschfeld e Gelman (a cura di) *Mapping the mind*, Cambridge University Press, Cambridge.
7. Sperber e Wilson (1986/1995) *Relevance: communication and cognition*, Blackwell, Oxford., trad.it. della I ed (1993), *La pertinenza*, Anabasi, Milano.

Indice

Premessa

Parte prima

L'architettura della mente

1. La mente computazionale

1.La psicologia computazionale

1.1 L'argomento della povertà dello stimolo -1.2 Turing e il funzionalismo computazionale - 1.3 La teoria della visione di Marr

2.La spiegazione psicologica e la struttura della mente

2.1 L'euristica della scomposizione e della localizzazione - 2.2 Il funzionalismo "homuncolare"
Note

2. La mente modulare

1.La modularità epistemica

1.1 La modularità epistemica nella linguistica generativa - 1.2 La modularità epistemica in psicologia dello sviluppo - 1.2.1 La fisica ingenua - 1.2.2 La matematica ingenua

2.La modularità fodoriana

2.1 L'impenetrabilità cognitiva delle illusioni percettive -2.2 La geometria ingenua -2.3 Ambiguità sintattiche - 2.4 Riconoscimento lessicale - 2.5 Altre proprietà dei moduli fodoriani

3. Oltre la modularità: la proposta di Karmiloff-Smith

4. La modularità in neuropsicologia

4.1 Un esempio di dissociazione doppia: prosopoagnosici e Capgras - 4.2 Modularità neuropsicologica e modularità fodoriana

Note

3. La modularità massiva

1. La mente non modulare

1.1 Fine di una scienza del pensiero?

2. L'ipotesi della modularità massiva

2.1 Pensiero quotidiano e pensiero scientifico: un'equivalenza fondata? - 2.2 L'argomento biologico-evoluzionistico - 2.2.1 La psicologia evoluzionistica 2.3 L'argomento della trattabilità computazionale - 2.4 Evoluzione senza demodularizzazione

3. Le proprietà dei moduli centrali

4. Come ricomporre un quadro frammentato

Note

Parte seconda

Psicologia ingenua, interpretazione pragmatica e acquisizione del lessico

4. La psicologia ingenua

1. Natura e ontogenesi della psicologia ingenua: due scoperte

2. Natura e ontogenesi della psicologia ingenua: la teoria della teoria

2.1 La teoria della teoria costruttivista e aspecifica - 2.2 La teoria della teoria modularistica

3. ToMM in un'architettura modulare moderatamente massiva

4. ToMM come modulo non parametrizzato

5. Il meccanismo di monitoraggio

6. Psicologia ingenua e intelligenza machiavellica

Note

5. Interpretazione pragmatica e acquisizione del lessico

1. Il modello del codice

2. Dal modello del codice al modello inferenziale: la pragmatica griceana

3. La teoria della pertinenza

3.1 Pertinenza e cognizione: il principio cognitivo di pertinenza - 3.2 Come calcolare costi e benefici - 3.3 Pertinenza e comunicazione: il principio comunicativo di pertinenza - 3.3.1 La procedura di comprensione guidata dalla pertinenza - 3.3.2 Implicature ed esplicature - 3.4 La teoria della pertinenza e il problema della selezione

4. L'interpretazione pragmatica come processo generale per dominio

5. L'interpretazione pragmatica come processo generale per dominio

6. La psicologia ingenua nell'acquisizione del lessico

6.1 Associazioni - 6.2 Lessico e psicologia ingenua - 6.3 Psicologia ingenua o analisi geometrica dello sguardo?

Note

Una visione d'insieme

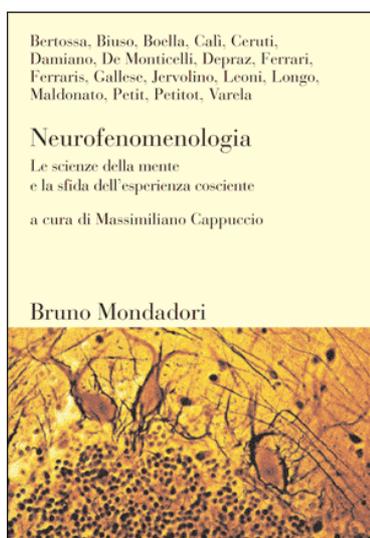
Note

Bibliografia

Recensione

**NEUROFENOMENOLOGIA. LE SCIENZE DELLA MENTE
E LA SFIDA DELL'ESPERIENZA COSCIENTE.**

Bruno Mondadori, Milano 2006
a cura di Massimiliano Cappuccio



La raccolta di saggi curata da Massimiliano Cappuccio costituisce un progetto affascinante. Accattivante già dal titolo, nonostante le iniziali diffidenze dovute alla portata del progetto, riesce subito a superare eventuali pregiudizi, ponendo spesso al lettore lo spunto per interessanti riflessioni.

Opera prima della neo-nata collana diretta da Roberta De Monticelli, fenomenologa e studiosa di Husserl, il testo si può leggere agilmente grazie alla divisione sia per sezioni tematiche, sia per singoli saggi che, seppur leggibili individualmente, si integrano in una tematica omogenea, ispirati dalle riflessioni e dalla filosofia di Francisco J. Varela.

Analisi e considerazioni eterogenee, tra cui spiccano gli interventi di Petitot, Petit e Depraz, animano un percorso lastricato di ambiziosi sforzi, alcuni riusciti ed altri meno, che hanno il merito di sviscerare le tesi poste all'inizio del libro attraverso varie alternative d'approccio. A seguito della presentazione della collana, la lunga ed esauriente introduzione di Cappuccio ci fornisce la chiave di lettura dell'opera chiarendone gli scopi e presentando le tematiche che collegano tutti i saggi.

L'ambizioso progetto di Varela, inseritosi nel dibattito moderno tra le Neuroscienze, le scienze cognitive e la filosofia della mente, consiste nel riavvicinare due tradizioni filosofiche: analitica e fenomenologica (definita anche "continentale"), fino a ora storicamente e concettualmente inconciliabili.

Nello specifico, le intuizioni del prematuramente scomparso Varela, sostenevano una sintesi tra le scienze naturali e la fenomenologia: una *naturalizzazione della fenomenologia*. Secondo le argomentazioni di Varela, integrando i dati empirici delle Neuroscienze con la pratica di riduzione fenomenologica, è possibile dare risposta al "problema difficile" descritto da Chalmers, colmando il divario tra la ricerca scientifica in terza persona e il punto di vista soggettivo. In altre parole, una riduzione fenomenologica biologicamente

fondata realizzerebbe una scienza dell'esperienza cosciente. La teoria di Varela si snoda attraverso concetti esplicativi formulati dall'elaborazione dei risultati empirici delle scienze cognitive: l'individuo viene definito un *meccanismo autopoietico*.

L'autopoiesi è una caratteristica essenziale del vivente, essa definisce la proprietà di un organismo di mantenere un equilibrio omeostatico tra il tutto e le sue parti; l'autopoiesi è un momento attivo dell'esperienza cosciente, in cui, a seguito delle informazioni elaborate dai processi cognitivi, l'organismo si riconfigura in relazione all'ambiente. Nella percezione dell'esperienza cosciente l'autopoiesi esprime lo stretto legame tra individuo e ambiente, co-implicanti e co-determinanti il fenomeno: l'individuo somatizza le dinamiche fisiche e sociali dell'ambiente, il quale costituisce l'orizzonte d'espressione del fenomeno.

La cognizione è, per Varela, un processo comune a tutti gli organismi e determina la competenza nell'interagire con l'ambiente; intesa non come una rappresentazione, ma un'attività che ri-crea l'equilibrio organismo-ambiente per fini adattativi: i processi cognitivi permettono al soggetto di riconfigurare il suo assetto chimico/organico in relazione all'ambiente. Questa chiave di lettura della cognizione di un'unità autopoietica è necessaria per un approccio *enattivo* dell'esperienza cosciente.

Il cognitivismo enattivo si propone come alternativa al cognitivismo computazionale e connessionistico fornendo descrizioni di livello superiore a quelle di tipo sub-simbolico e logico-proposizionale, coinvolgendo la sfera fenomenica del vissuto psichico, quale livello contiguo all'emergere della coscienza. Il cognitivismo enattivo descrive una *mente incorporata*: «il sé cognitivo è la sua propria implementazione: la sua storia e la sua azione formano un blocco unico»¹.

L'azione incorporata non è riducibile né a un evento meccanico, né di elaborazione delle informazioni, ha un senso ed è intenzionalmente orientata a una finalità, che tiene presente tanto le implicazioni fisico-neuronali, che il vissuto fenomenico del soggetto, ovvero le sue esperienze qualitative. Dall'analisi di tutti i fattori costitutivi del fenomeno, Varela avanza l'ipotesi che le strutture cognitive *emergano* da ricorrenti schemi (*patterns*) di attività senso-motoria.

Il concetto di emergenza viene definito da Varela prendendo come riferimento le teorie della complessità delle scienze fisiche e matematiche; questa definisce la realizzazione, integrabile su più livelli, di fenomeni auto-organizzati, prendendo le distanze da emergentisti di matrice connessionista. L'emergenza manifesta così un legame tra

¹ Massimiliano Cappuccio, a cura di, *Neurofenomenologia. Le scienze della mente e la sfida dell'esperienza cosciente*, (Bruno Mondatori editore, 2006), p. 25.

molteplici livelli descrittivi: il piano elettro-chimico dei processi neuronali, il piano dell'astrazione logico-simbolica e quello dell'esperienza fenomenica.

Nel cognitivismo enattivo, il fenomeno si forma dalle complesse dinamiche neuronali per manifestarsi nel livello fenomenico, il quale, a sua volta, influenza i livelli inferiori. La riduzione fenomenologica combinata con le ricerche delle neuroscienze permetterebbe di confrontare i risultati oggettivi, con l'analisi della visione soggettiva, che alla scienze manca; una Neurofenomenologia, come strumento pragmatico, è necessario alle scienze della mente.

Fresco dell'introduzione esplicativa di Cappuccio, il lettore si trova davanti al testo di Varela già in possesso degli strumenti per gettare le fondamenta di una ricerca lunga tutto il libro. I lavori sono distribuiti nel testo attraverso le sei sezioni, costituite in ordine da: I - Matematismo e formalizzazione, II - Corpo vivo e percezione spaziale, III – Metodologia di ricerca in prima persona, IV – Intersoggettività ed empatia, V – Ontologia fenomenologia e naturalizzazione, VI – Temporalità e coscienza.

Ogni saggio dei nostri studiosi e ricercatori aggiunge qualcosa alla ricerca inaugurata da Varela, di particolare rilievo il contributo di Petitot nella prima sezione e quello di Petit nella seconda, incentrati sul confronto con il pragmatismo americano e le teorie computazionali della mente da un lato, e la tradizione fenomenologica dall'altro, alla ricerca di una sintesi che fondi la Neurofilosofia. Da sottolineare anche le intuizioni del saggio della Depraz riguardo lo sviluppo della pragmatica neurofenomenologica.

Ogni sezione, riporta al termine del libro una bibliografia specifica per l'appassionato che volesse approfondire.

Riccardo Furi

Indice

Introduzione

Neurofenomenologia, un rimedio metodologico al "problema difficile"

SEZIONE I: Matematismo e formalizzazione

- La svolta naturalistica della fenomenologia
- La ragionevole efficacia della matematica e le sue origini cognitive

SEZIONE II: Corpo vivo e percezione spaziale

- La spazialità originaria del corpo proprio. Fenomenologia e neuroscienze

- Neuroestetica e fenomenologia. Per una teoria fenomenologica della percezione pittorica
- Il gesto fotografato. Note per un'archeologia della neurologia
- Il corpo come macchina semantica. Una prospettiva fenomenologia sull'intelligenza artificiale

SEZIONE III: Metodologia di ricerca in prima persona

- Mettere al lavoro il metodo fenomenologico nei protocolli sperimentali. "Passaggi generativi" tra l'empirico e il trascendentale
- Meditazione di presenza mentale per le scienze cognitive. Pratica del corpo e metodo in prima persona

SEZIONE IV: Intersoggettività ed empatia

- Corpo vivo, simulazione incarnata e intersoggettività. Una prospettiva neuro-fenomenologica
- L'empatia nasce dal cervello? La comprensione degli altri tra meccanismi neuronali e riflessione filosofica

SEZIONE V: Ontologia fenomenologia e naturalizzazione

- Persona e individualità essenziale. Un dialogo con Peter Van Inwagen e Lynne Baker
- Perché è meglio che la sintesi sia passiva. Trascendentalismo e naturalizzazione

SEZIONE VI: Temporalità e coscienza

- Coscienza della temporalità e temporalità della coscienza
- Ricoeur: la fenomenologia della memoria e il confronto con le scienze cognitive

Bibliografia

Indice dei nomi

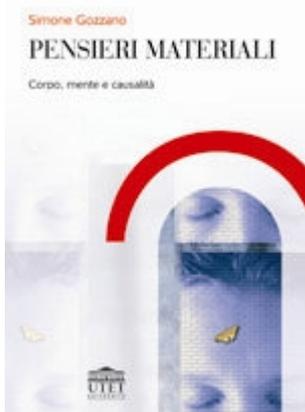
Gli autori

Recensione

PENSIERI MATERIALI

Utet, Milano 2007

Simone Gozzano



La causazione del mentale sul fisico è sempre stata un argomento filosofico delicato: come è possibile che questi due ambiti così distinti abbiano influenza l'uno sull'altro? Ogni soluzione è contraria ad alcune intuizioni del senso comune. Infatti, sostenere la causalità del mentale richiede di spiegare in che modo questa è possibile; viceversa, negarla porta a eliminare il ruolo autonomo della mente rispetto al corpo, ruolo che è invece riconosciuto dal senso comune. In questo libro Simone Gozzano propone la sua personale soluzione al

dilemma.

Il primo capitolo affronta le radici del problema mente-corpo, sostenendo che «è una conseguenza, piuttosto che il punto di avvio, di un modello generale di causalità» basato sull'idea «secondo la quale tra causa ed effetto c'è un trasferimento di una quantità fisica conservata» (pag. IX). Il secondo capitolo mostra come tale modello della causalità sia strettamente connesso al fisicalismo, approfondisce tale tesi e la mette in relazione al principio di chiusura causale del mondo fisico. La congiunzione di queste due ipotesi esclude la causalità del mentale e lo rende epifenomenico (Kim 1998). Secondo Gozzano l'unica strada per difendere a questo punto la causazione del mentale senza rinunciare al fisicalismo è quella di identificare fisico e mentale. Il terzo capitolo è così dedicato alla difesa di una particolare versione della teoria dell'identità dei tipi dalle obiezioni portate dalla teoria delle identità necessarie di Kripke e da quella della realizzabilità multipla. Il quarto capitolo, infine, prosegue la critica a quest'ultima tesi e fornisce degli argomenti a favore della teoria dell'identità.

Scendendo più nel dettaglio, Gozzano dedica il primo capitolo a un'analisi della causalità. Questa relazione si riferisce a eventi e si caratterizza per (i) la *regolarità* che si stabilisce tra due di essi (un evento segue *necessariamente* l'altro), (ii) la *contiguità spaziotemporale* tra gli stessi, e (iii) la loro *asimmetria* (l'effetto viene sempre dopo la causa). Due grandi classi di teorie fondano tali proprietà sulla base di valutazioni differenti. Da una parte, le teorie *concettuali* (nomologiche, probabilistiche e controfattuali) fanno

dipendere l'individuazione di una relazione causale dal nostro apparato epistemico. Esse non pongono petizioni di principio a una nozione di causalità tra eventi mentali e fisici, come fa invece il gruppo delle teorie *empiriche*. Queste presuppongono la presenza di una proprietà comune nella causa e nell'effetto, proprietà che rende possibile il trasferimento di una quantità conservata dalla prima al secondo. La natura delle grandezze che possono intervenire in questo processo è definita dai principi di conservazione della fisica. Essa stabilisce dunque quali entità possono rientrare nel modello e sono quindi considerate esistenti (fiscalismo). Poiché la fisica non descrive delle grandezze comuni al fisico e al mentale, una relazione di causazione psico-fisica risulta difficilmente ammissibile.

Gozzano dedica il secondo capitolo a precisare la tesi fiscalista. Considerando che la fisica cambia nel tempo, il fiscalismo corre il rischio di essere empiricamente confutato se assume come entità esistenti quelle definite *attualmente* dalla fisica, mentre può svuotarsi di contenuto se presuppone una teoria futura cui la fisica converge. Esso è identificato dalla tesi che le nostre attuali conoscenze scientifiche, seppur rivedibili, sono tutto quello di cui abbiamo bisogno per comprendere il mondo (e dunque il mentale). In base a questa visione, le entità mentali possono ancora essere incluse nell'ontologia se sono sopravvenienti su quelle fisiche (cioè se la modifica di proprietà fisiche porta necessariamente alla modifica di proprietà mentali).

Gozzano tenta quindi di spiegare da cosa è determinata una relazione causale tra due eventi. Dapprima difende la definizione di evento come tripla <oggetto, proprietà, tempo> (Kim 1973). Gli eventi sono quindi individuati ontologicamente da una proprietà; questa può essere descritta in modi diversi, ma è individuata dal suo ruolo causale. In secondo luogo afferma che la proprietà che identifica un evento è anche determinante nelle relazioni causali che questo istituisce con altri eventi: «le relazioni causali sono relazioni tra eventi determinate da proprietà» (pag. 80). È qui presente il rischio di una spiegazione circolare: la nozione di proprietà è allo stesso tempo definita e determinante per il potere causale di un evento. Secondo Gozzano la circolarità è solo apparente. I ruoli causali hanno un valore prioritario nell'individuazione delle proprietà, e dunque sul piano epistemologico. Sul piano ontologico, tuttavia, la determinazione di una proprietà avviene tramite una fissazione *progressiva* di suoi ruoli causali: tale processo presuppone la priorità della proprietà rispetto ai ruoli stessi. Proprietà e relazioni causali sono quindi nozioni interdipendenti, ma non si definiscono in modo circolare.

A questo punto Gozzano introduce una tesi complementare al fiscalismo, il Principio di Chiusura Causale del mondo fisico (PCC), del quale viene difesa la formulazione di Robb

(1997): «ogni evento fisico ha nella propria storia causale solo eventi fisici e proprietà fisiche» (pag. 88). Il PCC nega quindi che una causa mentale e una fisica possano entrambe essere sufficienti per il darsi di un effetto (*sovradeterminazione* del mentale sul fisico).¹ Il PCC non esclude tuttavia che eventi mentali possano provocare singolarità fisiche non spiegate, capaci di indurre effetti fisici. Il problema della causazione mentale origina quindi soltanto quando la *completezza* della spiegazione fiscalista predicata dal PCC si unisce alla metafisica fiscalista vera e propria, che afferma la *necessità* di tale spiegazione. Riassumendo, la metafisica fiscalista è connessa alla (i) tesi della chiusura causale del mondo fisico, ed esclude (ii) la sovradeterminazione causale del mentale sul fisico. Ammette tuttavia (iii) la sopravvenienza del mentale sul fisico. Gozzano accetta l'argomento di Kim (1995) che la congiunzione di (i-iii) insieme alla tesi dell'irriducibilità tra proprietà fisiche e mentali porta all'epifenomenismo (inerzia causale) di quest'ultime. L'unica strada affinché il fiscalismo si concili con la causalità mentale è allora quella di appoggiare una teoria dell'identità di stati fisici e mentali.

Gozzano discute la formulazione dell'identità dei tipi all'inizio del terzo capitolo. Per la maggior parte dei funzionalisti gli stati mentali sopravvengono sugli stati materiali in virtù del loro ruolo funzionale (funzionalismo del second'ordine). Riprendendo Armstrong (1968), egli identifica invece gli stati mentali con gli occupanti in senso stretto dei ruoli causali (funzionalismo del prim'ordine): una volta che l'implementazione fisica di uno stato mentale è stata identificata, il ruolo funzionale non definisce quindi uno stato ontologicamente autonomo. Il funzionalismo si presta così a un'analisi riduzionista delle proprietà mentali, e il problema mente-corpo trova una soluzione immediata e conseguente.

Gli intenti riduzionisti del funzionalismo del prim'ordine minimizzano l'indipendenza del livello di descrizione funzionale dell'attività mentale. L'identità dei tipi è stata allora criticata sottolineando l'autonomia descrittiva di tale piano con il ricorso alla tesi della realizzabilità multipla. Questa fa causa comune con un argomento proposto da Kripke (1980). In base alla sua *teoria del riferimento diretto* le proprietà mentali si comportano come designatori rigidi. Se sono identiche a proprietà fisiche, tale identità deve quindi essere necessaria e non sarà possibile avere una proprietà mentale senza la corrispondente proprietà fisica. La componente fenomenologica di uno stato mentale è

¹ L'argomento di Crane (2001) a sostegno della sovradeterminazione viene abilmente smontato da Gozzano distinguendo tra casi di sovradeterminazione vera e propria, in cui catene causali indipendenti hanno lo stesso effetto, e casi di sovradeterminazione in cui la stessa catena causale è descritta a livelli diversi. Crane dimostrerebbe l'esistenza solo di questo tipo di sovradeterminazione del mentale, poco rilevante sul piano ontologico. Si osservi come Gozzano utilizza nella dimostrazione la teoria, poi confutata, della realizzabilità multipla.

tuttavia pervasiva ed essenziale: prendendo il dolore, per esempio, si può dire che esso non è altro che *la dolorosità* di una sensazione. È quindi possibile immaginare un dolore non associato ad uno stato fisico, contrariamente a quanto previsto dalla teoria del riferimento diretto. Bisogna così concludere che la premessa su cui tutto il ragionamento si basa, l'identità di proprietà fisiche e mentali, è errata.

L'argomento è fondato sull'identificazione del dolore con la sensazione di dolore, fatto che rende Kripke un «*eliminativista* del lato corporeo delle sensazioni di dolore»². Contro questa posizione Gozzano propone una teoria causale delle sensazioni. I recettori di una modalità sensoriale devono essere connessi causalmente (cioè in modo legisimilmente affidabile) alle cause distali (esterne) che li hanno attivati; non sarebbe altrimenti possibile giustificare la capacità referenziale delle sensazioni. Accanto a questa componente *proto-intenzionale*, una sensazione presenta inoltre una *rappresentazione fenomenologica* correlata. Tale rappresentazione è *distintiva*, cioè «diversa da tutte le altre e associata a determinate condizioni di attivazione»³: se così non fosse essa sarebbe un segnale ambiguo e poco efficiente. Il fatto che le condizioni di attivazione siano esterne fa sì che la distintività di una sensazione sia «fissata da condizioni non fenomenologiche»⁴, nella fattispecie causali. Se la teoria causale della percezione proposta è giusta, le sensazioni richiedono per essere definite anche un riferimento a condizioni di correttezza basate su criteri esterni alla sensazione stessa: non è allora possibile immaginare, come fa Kripke, il caso di un dolore che non sia fisicamente localizzato. La confutazione di questa premessa dell'argomento di Kripke difende quindi la teoria dell'identità da un'importante obiezione. Tale linea difensiva è considerata risolutiva anche contro l'argomento della realizzabilità multipla, che sempre dall'argomento di Kripke dipende.

Accettare l'identità dei tipi significa svuotare il funzionalismo di molto significato. Il quarto capitolo è dedicato alle conseguenze di questo atto. Il funzionalismo del prim'ordine integra analisi concettuale ed empirica nell'individuazione degli stati mentali. L'analisi concettuale è richiesta per individuare il loro *ruolo causale*. Gozzano propone di identificare questo con l'intersezione massimale di determinate strutture funzionali esemplificate *in maniera univoca* «in una specifica entità fisica»⁵. Egli sta dunque proponendo una versione di funzionalismo in cui i tipi funzionali sono molto vicini ai tipi fisici.

² Pag. 141.

³ Pag. 130.

⁴ Pag. 131.

⁵ Pag. 173.

Un'obiezione naturale è allora che la sua teoria assottiglia, fino ad eliminare, le differenze tra la generalità delle proprietà psicologiche e la specificità di quelle fisiche ad esse sottostanti. A tale obiezione, Gozzano replica problematizzando la tesi della realizzabilità multipla. In base ad essa una proprietà di alto livello può essere istanziata in casi particolari con poteri causali diversi, e dunque proprietà di basso livello diverse. Tuttavia vengono adottati «degli standard molto rilassati quando si tratta di fissare le condizioni di individuazione delle proprietà di altro livello come *avere dolore*»⁶, mentre «quando si considerano i predicati fisici, ecco che le condizioni di individuazione si fanno assai stringenti»⁷. Gozzano afferma quindi che «l'argomento della realizzabilità multipla individua solo una tecnica euristica di approssimazione al problema empirico di identificare i realizzatori delle proprietà mentali [...]. Allorquando si scende nei dettagli, tale relazione mostra immediatamente la propria debolezza e ci si affida a identità specifiche»⁸.

La spiegazione dell'identità dei tipi non intende quindi negare che proprietà (mentali) di alto livello e (fisiche) di basso livello siano descritte in termini funzionali diversi. Soltanto «il fisicalismo e la teoria dell'identità sono delle tesi ontologiche»⁹¹⁰, e come tali propongono una forma di riduzionismo che semplicemente «asserisce che due entità supposte autonome sono di fatto la medesima entità»¹¹. Tale riduzionismo *ontologico* non risolve i compiti di un altro tipo più interessante di riduzionismo, quello *epistemologico*, in base al quale «le leggi relative a un determinato ambito di fenomeni sono deducibili, o in qualche senso ottenibili, dalle leggi relative a un dominio distinto»¹². Confondere questi due piani spinge molti filosofi a una critica a spada tratta del fisicalismo e della teoria dell'identità. L'accettabilità di queste posizioni si mostra invece quando l'identità tra stati mentali e fisici viene intesa in modo più rilassato, e si rinuncia alla petizione di principio di mantenere il dominio del mentale e del fisico distinti.

Il libro è orientato a un pubblico specialista. Difficile da seguire appare la scelta di anticipare e poi riprendere alcune questioni, come il problema della definizione di evento - introdotto nel primo capitolo e poi ripreso (pagg. 68-80) - o la spiegazione della problematicità della causalità del mentale. A proposito di questa Gozzano prende due

⁶ Pag. 191.

⁷ Pag. 192.

⁸ Pag. 188.

⁹ Questo permette di evitare le critiche di argomenti classici contro di esso come quelli di Nagel (1974), o quello di Mary e la stanza dei colori (Jackson 1986), che sono argomenti epistemici, non metafisici.

¹⁰ Pag. 207.

¹¹ Pag. 198.

¹² Pagg. 198-199.

linee argomentative diverse. Nei primi due capitoli la fa dipendere dall'accettazione del modello causale del processo fisico; questo viene poi però connesso strettamente al fisicalismo e all'argomento di Kim (1998) (pagg. 100-103). Il lettore si ritrova così con due argomenti connessi, ma indipendenti, per introdurre il problema della causazione mentale: l'assenza di leggi psico-fisiche secondo un particolare modello di causalità e l'accettazione di una metafisica fisicalista incompatibile con il dualismo delle proprietà.

La principale linea argomentativa di Gozzano è chiara e consequenziale. Una distinzione molto importante è quella tra livello ontologico e epistemologico, che viene chiamata in causa, ad esempio, per distinguere tra la determinazione del ruolo causale di una proprietà e la sua individuazione: in questo caso la dimostrazione fornita (pagg. 81-84) non sembra tuttavia sufficiente ad allontanare il rischio di circolarità tra le due. Un'obiezione più sostanziosa può essere invece rivolta alla teoria causale della percezione sostenuta contro l'argomento di Kripke. Ci si può allontanare dalla prospettiva kripkeana anche negando l'esplicito ruolo referenziale delle sensazioni e continuando a difenderne la distintività solo in virtù dell'univocità dei comportamenti da esse causati. Faccio riferimento all'esempio di una persona che avvicina la mano al fuoco e, scottandosi, la ritira. In questo caso il processo che porta dalla sensazione di dolore alla corrispondente reazione motoria di allontanare la mano è spiegabile puramente in termini di stimoli e reazioni dell'organismo, senza far riferimento a circostanze esterne. Questa non costituisce una spiegazione *razionale* dell'accaduto, ma in termini *causali* è più che sufficiente. Sarebbe quindi ugualmente vero che «impariamo a distinguere tra *qualia* simili tra loro perché siamo attenti a ciò che ci accade quando essi si presentano alla coscienza, non il contrario»¹³; tuttavia *quello che ci accade e a cui stiamo attenti* non sono le circostanze esterne di innesco di una sensazione, ma lo stato interno di attivazione dei nostri recettori.

Marco Fenici

Bibliografia

- Armstrong (1968), *A materialist theory of the mind*, Routledge, London.
- Jackson (1986), *What Mary didn't know*, The Journal of Philosophy, 83, pagg. 291-295.
- Kim (1998), *Mind in a physical world*, MIT Press, Cambridge, Massachusetts.

¹³ Pag. 144.

- Kripke (1980), *Naming and necessity*, Blackwell, Oxford.
- Nagel (1974), *What is like to be a bat?*, *Philosophical Review*, 83.

Indice

Gozzano, pensieri materiali

Introduzione

1. Il problema mente corpo
2. Fisicalismo e causalità
3. Sensazioni e identità
4. Una nuova teoria dell'identità

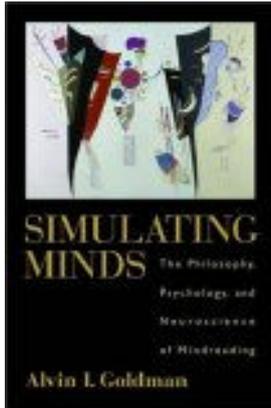
Bibliografia

Indice dei nomi

Recensione
**SIMULATING MINDS: THE PHILOSOPHY, PSYCHOLOGY
AND NEUROSCIENCE OF MINDREADING**

Oxford University Press, 2006

by Alvin Goldman



In *Simulating minds* (2006) Alvin Goldman provides arguments for a hybrid simulationist approach to mindreading capacities. Mindreading is a second-order activity based on producing beliefs about a target's first order mental states. In particular, *detecting* is the capacity to point out current mental states of a target and *predicting* is the capacity to predict the behavior of a target by using information about his underlying mental states. The two most important theories of mindreading are the Theory-theory (TT) and the Simulation theory (ST). ST is about predicting a target's decisions by putting oneself into her shoes and going through the same decision-making process. On the contrary, TT maintains that mindreading is an information-rich mechanism that processes theories about the target's mental states. The core chapters of Goldman's books are chapters two, six and seven where he explains his simulationist theory of low-level and high-level mindreading, and then chapters nine and ten where he focuses on introspection.

The low-level mindreading prototype is a mirroring type of simulation process. It is a primitive, automatic, largely unconscious process. In chapter six Goldman focuses on a special class of mindreading tasks, namely recognizing someone's face as expressive of certain emotions. Empirical findings demonstrate that when we look at someone making an expression *e*, we are also affected by a type of emotion which is actually related to *e*. There may be a causal relation between a basic emotion *E* and an expression *e* that, when observed, eventually causes the same type of emotion *E* in the observer. That is, the perception of fear in the target's face directly elicits activation of the same neural substrate in the observer, who introspects and classifies that emotion as fear and then projects fear onto the target. However, these empirical data only suggest that the expression *e* is a signal that elicits a certain kind of emotion. A particular cue may elicit fear by signaling danger. This does not entail that the observer would consider that

particular cue as an expression of fear in the first place. That is, mirroring effect is not the same as mindreading.¹ However, this “resonating mechanism” may play a role in mindreading. As a matter of fact, people who have a relevant emotion system E damaged are also impaired in recognizing the emotion E in others’ faces. So apparently the deficiency in seeing threats corresponds to a deficiency in recognizing scared expressions. However, *pace* Goldman that does not prove that low-level mindreading is based on simulation, as we can assume that the impairment may exclusively concern the cognitive capacity to respond appropriately to signals of threat. Hence, in order to attribute an emotion to the target we probably need an information-rich process that may not rely on a simulation routine. Therefore, it is not so obvious that simulating is the foundation of low-level mindreading.

High-level mindreading uses pretense or enactment-imagination. According to ST, the mindreader takes the target’s mental states by creating in her (pretended) desires, beliefs and intentions. Then, given those premises the mindreader herself runs a genuine practical reasoning process, which is supposed to be the same as the one run by the target. The outcome of this process is a (pretended) decision that the mindreader applies to the target. That is, the mindreader introspects and then projects a first order mental state of hers onto the target. Moreover, she uses E-imagination to resemble others’ mental states and to put herself in the target’s shoes. ST maintains that mindreading is mainly a matter of assuming the target’s prospective and inhibiting a self prospective when this is not functional to the mindreading process. That does not mean that the mindreader elaborates a set of beliefs about the target’s mental states. On the contrary, mindreaders first try to adopt the target’s prospective² and then adjust it by using an informed theory process. That is, mindreaders need to correct their preliminary simulation by adjusting some of their egocentric bias in order to fit the target’s point of view. Although inhibiting self-prospective is essentially a simulation practice, this adjustment step requires an element of theorizing, because the capacity to scrutinize the target’s mental set entails a theory-like body of knowledge. That is, the detection of the target’s initial mental states has to be guided by theory, as simulation process works only in predictive cases and cannot run “backward”, from observed behavior to mental states.

¹ On the other hand, if being scared is due to an interpretation of a target’s face as scared, that means that the emotion-resonance is *caused* by mindreading. Goldman could not accept such a view because it would imply that the mindreading occurs before the simulating process has taken place.

² The difficulty in inhibiting self-prospective is usually responsible for egocentric errors on several mindreading tasks, like the false-belief task.

Therefore, although Goldman sustains that mindreading is by default a simulation-practice, he admits that mindreading entails a form of theorizing and therefore he argues for a hybrid simulationist approach. On the other hand, theory-theorists do actually deny that simulation routine does play a role in mindreading practice (Nichols and Stich, 2003). In this respect it is hard to distinguish Goldman's views from weakened forms of theory-theory.

The main difference between TT and ST lays in how they explain the relation between third-person mindreading and self-attribution. ST maintains that there is an asymmetry between first and third person mindreading. Although people do not use simulation routine in first person mindreading, third person mindreading employs classification of one's own mental states in order to classify states of others. That is, simulation mindreading requires undertaking a first personal outlook in order to predict and imagine the target's point of view. That also implies that children acquire the capacity to attribute mental states to themselves before they acquire the capacity to attribute such states to others. Therefore, Goldman embraces the theory of *privileged access*, namely that people are in an authoritative position in relation to their own mental states. Moreover, he assumes that there is a *special method* of accessing one's own mental states, namely a quasi-perceptual process that works only in first-person attribution. He thinks that this introspection system takes features of concrete mental states as inputs and yield classifications of those states as outputs. Introspection takes place thanks to an introspective code (I-code) used to *represent* mental concepts and to *classify* currently observed mental state tokens.

Goldman's approach to self-knowledge faces several difficulties. For example, there is empirical evidence that people who are not able to access to their own mental states nonetheless explain their own behavior in terms of intentions and decisions (Gazzaniga, 1995). These data seem to support the thesis that self-knowledge is mostly a matter of self-interpretation rather than introspection (Carruthers, forthcoming). Goldman's theory faces some major problems also when it comes to introspection in autistic children. Goldman maintains that introspection is intact in autism, although autistic kids face major difficulties in third-person mindreading. But there are also empirical data which indicate that autistic people are also impaired in ascribing mental states to themselves. That may suggest that there is only a single faculty for both mindreading and self-knowledge (Carruthers, forthcoming). In this respect, the debate on autism may hold the key to some of the most relevant questions about self-knowledge and mindreading.

In conclusion, Goldman's book provides an extensive analysis of the simulationist approach by bridging cognitive neuroscience, developmental psychology, and philosophy of mind and therefore represents an essential reading for all who are interested in mindreading.

Marianna Bergamaschi Ganapini

- Carruthers, P. (forthcoming). Introspection: divided and partly eliminated. *Philosophy and Phenomenological Research*.
- Gazzaniga, M. (1995). Consciousness and the cerebral hemispheres. In M. Gazzaniga (ed.), *The Cognitive Neurosciences*, MIT Press.
- Nichols, S. and Stich, S. (2003). *Mindreading: an integrated account of pretence, self-awareness, and understanding other minds*. Oxford University Press.

Table of Contents

Chapter 1 *Philosophical and Scientific Perspectives on Mentalizing*

Chapter 2 *Conceptualizing Simulation Theory*

Chapter 3 *The Rationality Theory*

Chapter 4 *The Child-Scientist Theory*

Chapter 5 *The Modularity Theory*

Chapter 6 *Simulation in Low-Level Mindreading*

Chapter 7 *High-Level Simulational Mindreading*

Chapter 8 *Ontogeny, Autism, Empathy, and Evolution*

Chapter 9 *Self-Attribution*

Chapter 10 *Concepts of Mental States*

Chapter 11 *The Fabric of Social Life: Mimicry, Fantasy, Fiction, and Morality*

References

Author Index

Subject Index

Idee per una rilettura

LA MENTE E IL MONDO FISICO

McGraw-Hill, Milano 2000

di Jaegwon Kim



“La mente e il mondo fisico” di Jaegwon Kim è stato pubblicato in Italia nel 2000, da McGraw-Hill, con due anni di ritardo rispetto all’originale “Mind in a Physical World”. Si tratta di una raccolta di quattro relazioni presentate dall’autore presso l’università della California nel marzo del 1996. La finalità principale dell’opera è, da un lato, quella di chiarire alcune confusioni riguardanti il concetto di sopravvenienza e, dall’altro, quella di definire una teoria

riduzionista a proposito del problema mente-corpo.

A cosa serve la sopravvenienza?

Uno dei principali fraintendimenti legati al concetto di sopravvenienza è la presunta maniera in cui esso dovrebbe risolvere il problema mente-corpo. Uno degli obiettivi critici di Kim in questo senso è la teoria del *monismo anomalo* di Donald Davidson. L’errore principale di quest’autore sarebbe quello di considerare la sopravvenienza degli stati mentali su quelli fisici come una relazione metafisicamente profonda, capace di spiegare l’anomalia dei primi rispetto ai secondi, mentre in realtà si tratta di una semplice relazione fenomenologica.

Il concetto di sopravvenienza si limita a descrivere la co-varianza di determinate proprietà, non rappresenta un particolare tipo di relazione di dipendenza, come per esempio potrebbe essere la dipendenza causale. In relazione al problema mente-corpo la sopravvenienza non deve essere considerata quindi come una soluzione, ma solo come una giusta intuizione a proposito della natura fondamentale fisica dei nostri stati mentali.

Un altro pregiudizio riguardante il concetto di sopravvenienza è quello che, in qualche maniera, garantisce l’irriducibilità delle proprietà mentali a quelle fisiche. Tale idea deriva, secondo Kim, da un’interpretazione scorretta delle modalità proprie della riduzione fra proprietà. Davidson, come altri autori, si basa sulla visione di Nagel secondo la quale la

riduzione fra due classi di proprietà è possibile solo in presenza di leggi-ponte (assenti fra proprietà mentali e fisiche). Questa esigenza di "isomorfismo interteorico" è però così estrema da rendere rarissimi i casi di riduzione possibili.

Il modello di riduzione proposto dai teorici della *realizzabilità multipla* è per Kim, al di là della validità della teoria in se stessa, molto più coerente con la prospettiva fiscalista che egli intende difendere. Nell'ottica dei funzionalisti infatti qualsiasi proprietà mentale sopravveniente è nomologicamente riducibile al suo realizzatore fisico, non possiede cioè un dominio causale indipendente. Ciononostante, non significa che vi sia un'identità assoluta fra proprietà mentali e fisiche (il che comporterebbe una riduzione metafisica o logica): stati fisici uguali, in mondi possibili governati da diverse leggi fisiche, potrebbero istanziare proprietà mentali differenti.

Alla luce di quanto detto, Kim riconsidera la possibile utilità del concetto di sopravvenienza e individua una questione filosoficamente fondamentale: il problema della causazione mentale. Si tratta di un tema sempre attuale, in quanto profondamente legato alle nostre intuizioni quotidiane. La possibilità di considerarci, come esseri umani, protagonisti delle nostre azioni e capaci di un'interazione gnoseologica con l'ambiente che ci circonda, così come l'autonomia di una scienza psicologica, richiedono l'esistenza di cause mentali.

Uno degli argomenti più forti, dal punto di vista naturalistico, contro la possibilità della causazione mentale è quello dell'*esclusione mentale*. Accettare la causazione mentale significa accettare che un evento mentale m possa causare un evento fisico p . Ciononostante, a meno che non si voglia ricadere in una sorta di dualismo interazionista fra due domini separati e rifiutare la chiusura causale del mondo fisico, dobbiamo ammettere che esista anche una causa fisica p^* per p . Ma allora: che ruolo potrà mai giocare la causa mentale? All'interno di un dominio fisico, sostiene Kim, essa viene inevitabilmente *esclusa*.

La vendetta di Cartesio

Come risolvere il problema dell'esclusione causale? Autori come Fodor o Searle sostengono che la sopravvenienza delle proprietà mentali su quelle fisiche è una sufficiente garanzia del potere causale delle prime. Entrambi gli autori concordano sulla chiusura causale della fisica e rifiutano il dualismo cartesiano: il mentale non costituisce un dominio ontologico indipendente rispetto al fisico. L'attribuzione di un potere causale al mentale non implica quindi per costoro un'infiltrazione ingiustificata da un dominio

all'altro, ma solo l'accettazione del fatto che un sistema si possa interpretare secondo diverse prospettive.

Kim ritiene che questa soluzione sia assolutamente illecita. L'errore di base consiste nel considerare il concetto di sopravvenienza come una spiegazione del problema mente-corpo, in quanto garante del ruolo causale autonomo del mentale. Se si prende seriamente il problema dell'esclusione mentale ci si rende presto conto che per spiegare la causazione mentale l'unica alternativa valida è il dualismo cartesiano che, chiaramente, non rappresenta un'alternativa accettabile.

Si prenda il caso in cui uno stato mentale ne causa un altro (M è causa di M^*). I teorici della sopravvenienza accettano che, in presenza di uno stato mentale, deve esistere anche uno stato fisico che ne garantisca l'apparizione. Si consideri quindi che M^* possiede una base fisica subveniente P^* . A questo punto possiamo lecitamente chiederci che cosa ha davvero permesso l'occorrenza di M^* , se la supposta relazione causale con M , o la presenza della base fisica P^* .

C'è una chiara asimmetria tra le due pretese relazioni causali. Mentre infatti possiamo immaginare l'esistenza di M^* indipendentemente da quella di M , non accade lo stesso in assenza di P^* . La strategia dei sostenitori della sopravvenienza è quindi quella di *mediare* il potere causale di M attraverso P^* , è cioè dire che: M ha causato M^* *attraverso la causazione* di P^* . Ma è evidente che a sua volta, come ogni proprietà sopravveniente, anche M deve possedere una sua base fisica (che chiameremo P).

Se P non fosse apparso, non sarebbe apparso neppure M , e di conseguenza neppure P^* e M^* . M e P sembrano entrambe cause papabili ma purtroppo, a meno che non si voglia accettare una sovradeterminazione causale, la causa fisica deve essere considerata come privilegiata. Lo schema più plausibile secondo Kim è quello di considerare le relazioni causali unicamente fra gli stati fisici (P ha causato P^*), e gli stati mentali sopravvenienti come indipendenti fra loro (M è sopravvenuto su P , e M^* è sopravvenuto su P^*).

La causazione mentale sarebbe dunque un'illusione, ingiustificata nella misura in cui si accettano determinati principi naturalistici come la chiusura causale della fisica. Il problema dell'esclusione causale rappresenta "La vendetta di Cartesio" contro i fisicalisti, nel senso che solo accettando il dualismo delle sostanze si può ignorare la sovrabbondanza di cause presente nei presunti casi di causazione mentale.

Searle propone di non considerare la relazione di sopravvenienza come relazione causale, Fodor si appella alla presenza di leggi autonome nell'ambito delle scienze sociali. Sta di

fatto che, per Kim, si tratta solo di soluzioni *ad hoc* per nascondere le reali debolezze del concetto di sopravvenienza.

Chi ha paura del riduzionismo?

Gli obiettivi critici di Kim sono molti (dal epifenomenalismo di Block alle proposte di rivisitazione dei principi metafisici impliciti di Burge e Baker) e non avrebbe senso qui ripercorrere interamente le argomentazioni da lui proposte (si può dire che, in generale, ruotano intorno all'inintelligibilità della causazione mentale e alla futilità di rinvenire leggi-ponte). Molto più interessante, e forse anche più attuale, mi sembra il tentativo di *snellire* il concetto di riduzione sulla base di quanto detto a proposito della sopravvenienza.

Kim nega che il concetto classico di riduzione, quello nageliano, possa avere un valore filosoficamente interessante. In se stessa la riduzione non ha un valore per nulla negativo: nel caso della sensazione di dolore e la stimolazione delle fibre-C, come in molti altri, ci permette di comprendere determinati fenomeni sulla base di fenomeni più primitivi. Compie in pratica una delle finalità più elementari della scienza. La riduzione nageliana, soprattutto nel senso in cui è interpretata da Davidson, si oppone a tale principio nella misura in cui, allo scopo di tutelare la scienza psicologica, accetta la sopravvenienza mente-corpo come un fatto bruto e inesplicabile.

La riduzione deve essere in primo luogo *semplificazione*. La ricerca di leggi-ponte segue una direzione contraria. Il prezzo che si deve pagare per ottenere delle leggi-ponte è quello di una nuova terminologia e di una nuova ontologia di base. Quello che propone Kim è l'identificazione vera e propria fra proprietà mentali e proprietà fisiche. Il fatto che normalmente si presentino insieme può unicamente significare che *sono la stessa proprietà*.

Una clausola fondamentale di questo tipo di riduzione è che le proprietà mentali siano considerate come *funzionali*. Ciò significa che la riduzione fra i due tipi di proprietà non è necessaria metafisicamente, ma solo in relazione all'aspetto causale/nomico che le identifica in un determinato dominio e che, come si è già detto, potrebbe essere differente in un altro mondo possibile.

Il concetto di sopravvenienza non ha né la possibilità né le caratteristiche per risolvere il problema della causazione mentale: al contrario, aiuta a comprendere la relazione fra proprietà mentali e fisiche nei termini di una riduzione funzionale del potere causale delle prime a quello delle seconde.

Fiscalismo minimalista

Lo scenario che Kim presenta sembra essere quello di un mondo in cui gli stati mentali, privati del loro potere causale, vengono irrimediabilmente ridotti al loro componente fisico. In realtà ne “La mente e il mondo fisico” appare evidente la consapevolezza del problema rappresentato dall’esistenza dei *qualia*, considerato più estesamente in scritti posteriori¹. La riduzione funzionale si riferisce infatti all’aspetto intenzionale della coscienza, non ai cosiddetti stati qualitativi. Mentre per l’autore è inconcepibile l’esistenza di un mondo possibile che sia un esatto duplicato del nostro mondo ma privo d’intenzionalità, che i *qualia* siano distribuiti in modo differente o che siano totalmente assenti non rappresenta di certo un’impossibilità.

Kim sostiene che non ci sia alcuna maniera di funzionalizzare i *qualia* e rendere quindi conto della loro efficacia causale all’interno di uno schema fiscalista. Realisticamente, le uniche due opzioni disponibili sono l’*epifenomenismo* (secondo il quale la sopravvenienza degli stati mentali non implicherebbe la loro efficacia causale) e l’*eliminativismo*. Entrambe cadono però nell’*irrealismo mentale* in quanto, in ultima istanza, è proprio il potere causale ciò che da dignità ontologica a una cosa.

“La mente e il mondo fisico” si chiude con toni quasi pessimistici: forse il fiscalismo, pur sembrando così ragionevole, paga un prezzo troppo elevato e finisce per essere inconsistente con le nostre intuizioni quotidiane. Nonostante ciò, né il dualismo cartesiano, né le posizioni intermedie rappresentano una valida alternativa. L’unica possibilità di salvaguardare il mentale è quella di considerarlo come parte del fisico. Qualsiasi altra soluzione è effimera.

Kim non è comunque un dogmatico: il fiscalismo rappresenta una prospettiva di ricerca, non una dottrina definitiva. Si tratta del punto di partenza su cui fondare l’analisi della coscienza, consci di tutte le difficoltà a essa relazionate.

Lorenzo Baravalle

¹ In particolare in J.Kim, *Physicalism, or Something Near Enough*, Princetown University Press (2005).

Bibliografia

Baker L. R., *Metaphysics and mental causation*, in J.Heil, A. Mele (a cura di), *Mental causation*, Clarendon, Oxford, 1993.

Burge T., *Mind-body causation and explanatory practice*, in J.Heil, A. Mele (a cura di), *Mental causation*, Clarendon, Oxford, 1993.

Davidson D., *Essays on Actions and Events*, Oxford University Press, Oxford, 1980 (trad. it., *Azioni ed eventi*, Il Mulino, Bologna, 1992).

Fodor J., *The Modularity of Mind*, MIT Press, Cambridge, Ma, 1983 (trad. it. *La mente modulare*, Il Mulino, Bologna, 1999).

Jackson F., Pettit P., *Functionalism and Broad Content*, *Mind*, 97, 1988.

Kim J., "Downward Causation" in *Emergentism and Nonreductive Physicalism*, in A.

Beckerman, H. Flohr, J. Kim, *Emergence or reduction?*, de Gruyter, Berlin, 1992.

Kim J., *Supervenience and Mind*, Cambridge University press, NY, 1993.

Kim J., *Mind in a Physical World*, MIT press, Cambridge, Mass, 1998 (trad.it., *La mente e il mondo fisico*, Mc-Graw Hill (Dynamie), Milano, 2000).

Kim J., *Physicalism, or Something Near Enough*, Princetown University Press, 2005.

Nagel E., *The Structure of Science*, Routledge & Kegan, London, 1961 (trad. it., *La struttura della scienza*, Milano, Feltrinelli, 1967).

Putnam H., *Realism with a human face*, Harvard University Press, Cambridge, Mass., 1990.

Searle J. R., *The Rediscovery of the Mind*, Bradford Books, Montgomery, Vt., 1992 (trad. it., *La riscoperta della mente*, Boringhieri, Torino, 1992).

Idee per una rilettura
LA MENTE E IL COMPUTER,
INTRODUZIONE ALLA SCIENZA COGNITIVA

Il Mulino, Bologna 1990

di Philip N. Johnson-Laird

Gli esseri umani sono animali (Darwin)

Gli animali sono macchine (Cartesio)

Gli esseri umani sono macchine (La Mettrie)

Scopo del libro di Philip N. Johnson-Laird, psicologo inglese e professore di psicologia cognitiva e scienze cognitive presso l'Università di Princeton, è di "esaminare i processi mentali dal punto di vista della teoria della computazione" (pp. 6).

Il libro, pensato per studenti e per un pubblico interessato allo studio della mente dal punto di vista del paradigma *cognitivista*, è uno dei pochi manuali in circolazione che introduce espressamente ai metodi e ai modelli teorici delle scienze cognitive.

L'orizzonte teoretico di riferimento può essere individuato in quello che l'autore, in apertura, chiama un 'funzionalismo debole' secondo il quale si ammette la possibilità "che un programma per computer possa simulare aspetti dell'intelligenza umana".

L'autore, pur riconoscendo la natura non completamente computazionale della maggior parte dei processi mentali (visione, emozioni, udito, per esempio), resta un teorico dell'approccio simbolico in psicologia cognitiva, tanto da rievocare, in apertura, lo slogan caro ai funzionalisti e ai teorici di IA: "La mente può stare al cervello nello stesso modo in cui un programma può stare ad un computer" (pp. 28).

Naturalmente, ammette Johnson-Laird, si tratterebbe di una congettura destinata a rimanere tale, se i tentativi di risoluzione da parte degli scienziati cognitivi fossero di natura puramente *speculativa*. La questione, infatti, "è di natura empirica": costruire un robot in grado di percepire un mondo, interagire con esso ed esibire attività cognitive di alto livello.

Johnson-Laird non è solo uno scienziato cognitivo interessato alle questioni filosofiche ed epistemologiche della propria disciplina, ma è anche uno dei pochi psicologi sperimentali che scrive programmi per calcolatore con il preciso scopo di realizzare modelli computazionali delle sue teorie. Scrivere programmi al calcolatore per una data teoria psicologica, equivale, per lo scienziato cognitivo, a 'comprendere' i processi mentali alla

base del comportamento indagato (ragionamento, percezione, memoria, risoluzione di problemi).

L'idea che fa da sfondo a tutti i capitoli del libro è quella di *modello mentale*: una struttura cognitiva fondamentale per la comprensione dei processi percettivi (il risultato della percezione è un modello o sono modelli 3D del mondo) per la spiegazione dei processi di pensiero (il pensiero è la manipolazione *interna* di tali modelli che diventano, così, *mentali*) e per la formulazione di una teoria della competenza semantica (il linguaggio *serve* per comunicare il *contenuto* dei modelli mentali).

Nulla vieta, infatti, di dotare un robot dei più efficaci modelli mentali e delle procedure ricorsive in grado di manipolare tali strutture per ragionare, percepire e comprendere il linguaggio umano.

Se si dovesse scrivere, oggi, un libro introduttivo alle scienze cognitive, non potremmo fare a meno di dedicare almeno un capitolo alle neuroscienze cognitive, che in questi ultimi decenni hanno fatto veri e propri passi da gigante. Johnson-Laird fa sì riferimento ai contributi sperimentali offerti, per esempio, dalle teorie neuropsicologiche della visione, della coscienza, o allo studio dei famosi casi di 'split-brain' da parte del neuroscienziato Michael Gazzaniga, ma sempre privilegiando i modelli *computazionali* di tali attività cognitive.

Conosciamo le difficoltà che l'impostazione classica in scienza cognitiva affronta in sede di progettazione di automi o agenti artificiali già a un livello considerato basilare, come la percezione di oggetti, il riconoscimento di volti, o la comprensione del linguaggio. Consapevole di tali difficoltà, l'autore ammette che incorporare *in un* programma una teoria psicologica è, comunque, già un risultato significativo per lo scienziato cognitivo.

Oltre a illustrare i progressi della disciplina nei più importanti settori di indagine della mente (percezione, memoria, ragionamento, coscienza, linguaggio) il libro contribuisce a chiarire l'atteggiamento metodologico dello scienziato cognitivo nei confronti degli 'oggetti mentali'. Secondo Johnson-Laird, ciò che differenzia la scienza cognitiva dalle altre scienze, è il ricorso sistematico al computer, sia come metafora concettuale di riferimento nella comprensione dei processi mentali, sia come strumento in grado di incorporare, previa traduzione in un linguaggio di programmazione, una qualsiasi teoria della mente.

Una lettura più debole del paradigma funzionalista suggerisce, infatti, che un computer possa incorporare una teoria del mentale, o che sia in grado di simulare computazionalmente specifici processi mentali, senza, per questo, riconoscere che il

computer, o un ipotetico robot umanoide, siano dotati di un accesso interno ai propri pensieri (computazioni).

Un atteggiamento funzionalista ‘moderato’ non è detto, però, che sia quello corretto. Soltanto i progressi della robotica e dell’intelligenza artificiale potranno stabilire se un robot sarà in grado di possedere una vita mentale. I computer odierni non sono in grado di comprendere il linguaggio, perché non sono capaci di “metterlo in relazione con il loro dominio appropriato di interpretazione” (pp. 424). Ciò non esclude che in un futuro prossimo, risulti ancora impossibile “dare a queste macchine i meccanismi necessari per motivazioni, sensazioni interne e coscienza” (pp. 425).

La ‘computazione cognitiva’, aggiunge Johnson-Laird, suggerisce un’alternativa anche alle teorie dominanti in filosofia della mente: “I processi mentali sono le *computazioni* del cervello” (pp. 428).

Tale impostazione, sostiene l’autore, rigetta qualsiasi tipo di dualismo, assume l’esistenza di ben determinate caratteristiche della materia responsabili della possibilità di ‘esperire un mondo’, mette in primo piano l’organizzazione dei processi (mentali) e non esclude un’indagine sperimentale dei primitivi neurali (impulsi nervosi ed eventi elettrochimici in grado di produrre attività simbolica nella precisa struttura dei formati rappresentazionali). Ammette, inoltre, che non tutto è computabile in un senso classico e non rifiuta a priori che determinate computazioni, come quelle relative al pensiero deduttivo, possano essere conservate in un programma di un computer, così da costituire una ‘forma di rappresentazione dinamica dell’intelletto’ incorporabile in un modello mentale di un robot antropomorfo.

Tale impostazione sposa alcune congetture che non siamo obbligati ad accettare tanto facilmente: che la mente computi, o che il cervello computi, nel senso classico di Turing - von Neumann, non sembra affatto un risultato empirico *acquisito*. Che la maggior parte dei nostri processi cognitivi sia di natura computazionale non spiega come sia possibile concepire processi mentali non computabili: si pensi al settore della logica matematica e ai teoremi limitativi dimostrati dallo stesso Turing; se la mente di Turing è qualcosa che funziona come una macchina di Turing, allora, come spiegare il problema della fermata? Come hanno fatto le computazioni della mente *di* Turing a concepire un teorema che ‘dimostra’ la non computabilità dei suoi stessi processi mentali (delle sue stesse computazioni) se si accetta la computabilità di tali processi? Se un dato processo cognitivo è computabile fino a un certo punto, la parte non computabile è ancora da considerarsi un processo cognitivo nell’accezione che ne dà Johnson-Laird? E se non lo fosse, cosa

sarebbe? Di cosa scriviamo programmi? Di parti di processi mentali? Secondo questa impostazione, sarebbe più corretto definire la scienza cognitiva (classica) come la scienza che studia *le parti computabili* delle attività cognitive, ma se accettassimo tale 'condizionamento operativo' ne deriverebbe una visione molto circoscritta della mente.

In robotica, attualmente, viene data grande attenzione da parte dei progettisti ai contributi offerti dalle neuroscienze e dalla fisiologia animale. La ricerca attuale vede una sempre più marcata sovrapposizione degli studi classici sui sistemi artificiali con quelli basati sui sistemi naturali (si pensi alla 'cognizione incorporata' in IA, o ai robot realizzati dal ricercatore Rodney Brooks presso il Laboratorio di Intelligenza Artificiale del MIT). I sistemi artificiali odierni sono, però, molto lontani da fornire prestazioni accettabili in compiti 'apparentemente' privi di difficoltà per gli esseri umani come percepire, ricordare, o muoversi, il tutto intenzionalmente, ovviamente.

Le aspettative del cognitivista nei confronti della progettazione e della realizzazione di un robot che incorpori le più importanti funzioni cognitive umane rimangono, ad oggi, ancora inappagate. Per essere obbiettivi, l'autore del libro, è sì un cognitivista classico, ma è anche, potremmo aggiungere, un cognitivista 'critico'. I robot attuali, sottolinea a più riprese Johnson-Laird, non sono in grado di manifestare attività cognitive intenzionale, perché non sono capaci di mettere in relazione il livello delle computazioni con la varietà e la ricchezza sensoriale del mondo percepibile. Se, però, è possibile studiare la mente e le sue 'risorse' (termine recentemente introdotto da Marvin Minsky nel suo ultimo libro *The Emotional Machine*, per sostituire il ben più noto termine di 'agenti' comparso in *The Society of mind*) attraverso diverse classi di algoritmi, cercando, nel particolare comportamento di determinate computazioni la spiegazione dei fenomeni mentali, realizzare fisicamente queste computazioni in un robot diventa un banco di prova basilare per il teorico e per il progettista di agenti cognitivi artificiali (la mente sarebbe individuabile *nelle* computazioni del nostro agente artificiale).

Perché non denominare, un po' provocatoriamente, tale visione della mente, 'artificial behaviorism'? Resta da capire come un robot possa sviluppare autonomamente una teoria dei propri processi computazionali, ovvero, come possa *maturare* una psicologia se in lui esiste soltanto il livello delle computazioni. Tale argomento varrebbe anche se il robot fosse dotato di un sistema di interfacce visive, cinetiche, uditive, o altro; anche se possedesse delle istruzioni per costruire modelli mentali a partire da dati di tipo percettivo, o motorio, così da 'interpretare' le sue computazioni in un modello 3D del mondo, il robot non sarebbe in grado di sviluppare un accesso interno alle proprie

rappresentazioni: per fare questo, occorrerebbe inserire tali stati epistemici nella sua base di conoscenze in un formato simbolico o sub-simbolico, scegliendo un tipo di logica opportuna, o altre strutture come frames, script o appropriate reti neurali; ciò nonostante, i suoi modelli mentali continuerebbero a ignorare di possedere tali stati epistemici.

Dovremmo, inoltre, supporre che le computazioni cognitive siano, per il robot, 'sopravvenienti' rispetto agli impulsi elettrici che sfrecciano nei suoi microchip, ma questo non sembra essere sufficiente per sviluppare intenzionalità o memoria (a meno di considerare, per esempio, le schegge presenti in un sasso come 'memoria', nel senso di una proprietà fisica appartenente al sasso e ai *suoi* atomi). Questo è in parte corretto: ciò che chiamiamo memoria dipende da precise caratteristiche della materia e concordiamo con Johnson-Laird che questo valga anche per il pensiero; ma, se risulta *così fondamentale* il substrato che incorpora la 'natura dei processi psicologici', come facciamo a sostenere ancora la multi-realizzabilità del software mentale? Perché considerare soltanto i programmi per calcolatore, come l'equivalente dei processi cognitivi superiori negli esseri umani? Perché non considerare, per esempio, *pensiero* già i processi biochimici neurali? Non basta dire che l'intelligenza umana e artificiale si manifestano nella veste di un formato simbolico: un robot *non sa* di manipolare simboli, semplicemente esegue programmi installati dall'utente. I programmi di alto livello, così come le strutture rappresentazionali, non sono parte dell'ontologia interna del robot. Per il robot, non esiste neanche il livello delle computazioni, nel senso che *non sa* di stare effettuando computazioni, così come *per* la caffettiera non *esiste* il caffè come elemento intrinseco all'essere una caffettiera. Questo, si potrebbe argomentare, vale anche per noi umani: per esempio, la maggior parte di noi ignora il numero di cellule del corpo che stanno morendo in questo momento, ma possiamo dire di ignorarlo, appunto. Se il livello della manipolazione sintattica è ritenuto sufficiente per sviluppare un comportamento intelligente, non si vede perché non possano esserlo anche gli impulsi a livello delle sinapsi, o i fenomeni elettrochimici a livello molecolare presenti nei motoneuroni di una zampa di cavalletta, per esempio.

La mente è l'insieme dei processi sensoriali, percettivi, emotivi e cognitivi che il corpo è in grado di attivare. Quella tra mente e corpo è un'unità biologica vera e propria: la mente (umana e animale) semplicemente non esiste *priva* di un corpo; inoltre, il possesso di una mente non può inferirsi senza un comportamento osservabile e, si badi, il comportamento stesso non è soltanto qualcosa di osservabile, ma di *inferibile*.

Non sembra che la vita mentale si strutturi a partire da vuote occorrenze di un misterioso formato da riempire di volta in volta a seconda del dominio di esperienza considerato: il mondo che percepiamo non varia gran che, esso è la nostra forza di gravità, ciò che varia è il mondo interno (anche se non quanto vorremmo, purtroppo); ma se questa ricchezza di esperienza si colloca, per quanto ne sappiamo, a un livello 'subcognitivo' (forse citologico, o biochimico) cosa *intercettiamo* pensando? Dal punto di vista della vita, una scena percepita, un suono udito, una sensazione vissuta, un certo tipo di movimento, sono tutti ingredienti fondamentali per la genesi del pensiero, e, forse, sono già essi stessi *pensiero* puro (nel senso di *parti di mondo* sensoriale e percettivo *proiettate* nel corpo).

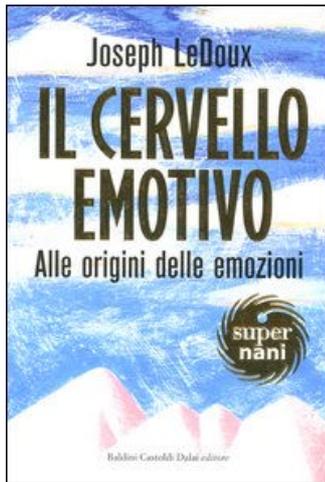
L'attualità del pensiero di Johnson-Laird, in linea con gli sviluppi di ricerca contemporanei, sta nel mettere in risalto, l'importanza della dimensione *interattiva* dell'intelligenza e della cognizione. Una cognizione, potremmo dire, condivisa, posseduta in comune con il mondo (in forma di modelli percettivi e mentali di questo) che fa da sfondo anche ai processi di ragionamento *astratto* e allo sviluppo della competenza semantica per i linguaggi naturali. Quando le aspirazioni del cognitivista si trasformeranno in realtà (robot antropomorfi) allora, i processi cognitivi *di tali* robot saranno molto simili ai nostri. Le operazioni mentali superiori coinvolte nella genesi di una teoria psicologica 'ingenua' saranno 'riproducibili' anche in un robot e un qualsiasi manuale di scienza cognitiva scritto per tale robot - che sarà in grado di aggiornare *automaticamente* i suoi modelli mentali collegandosi al web - non sarà, poi, così tanto diverso da quello offerto da Johnson-Laird.

Alberto Binazzi

Idee per una rilettura
IL CERVELLO EMOTIVO
ALLE ORIGINE DELLE EMOZIONI

Baldini e Castoldi, Milano 2003¹

Joseph LeDoux



Da William James a Damasio, LeDoux ripropone la storia di uno dei percorsi più importanti delle neuroscienze. Attraverso l'analisi di oltre cento anni di studi sulle emozioni, in questo libro sono analizzate le progressive conquiste nella comprensione delle funzioni e della struttura del cervello. Ci si trova di fronte ad un'analisi storiografica di rara completezza che, attraverso le ipotesi, gli errori, le teorie scientifiche e filosofiche, svela a poco a poco al lettore quanto sappiamo sulla biochimica delle emozioni.

La costruzione storiografica del testo permette di addentrarsi nella materia gradualmente, senza dover affrontare lo scoglio di un linguaggio e di una dettagliata topografia del cervello. LeDoux inizia la propria analisi dagli schemi funzionali di inizio secolo, sostituendo, scoperta dopo scoperta, le strutture fisiche alle funzioni prima solo rappresentate o ipotizzate. Il libro permette dunque di conquistare una buona conoscenza delle strutture fisiche e cerebrali che svolgono un ruolo funzionale nella genesi delle emozioni. Parallelamente, LeDoux svela il ruolo delle emozioni nello sviluppo di tutte le altre funzioni cerebrali. Dagli stimoli condizionati all'apprendimento, fino alle risposte subconscie, l'autore osserva le emozioni in tutto il loro campo d'azione.

Il testo consente inoltre un ottimo approccio alle tecniche sperimentali delle neuroscienze. Si passa rapidamente da un'analisi comportamentistica, come nel caso dei riflessi condizionati degli studi di Pavlov, allo studio delle lesioni cerebrali riproponendo le ultime indagini di Damasio, alle indagini sul ruolo di immagini e messaggi subliminali nella formazione di pensieri ed emozioni coscienti. Nel far ciò LeDoux descrive le tecniche di indagine e le accortezze che hanno permesso di creare le attuali generalizzazioni e teorie a partire da singoli casi, facendo molta attenzione a spiegare le modalità che permettono di comprendere le funzioni del cervello umano a partire da esperimenti sugli animali.

¹ *The Emotional Brain. The Mysterious Underpinnings of emotional Life*, di Joseph LeDoux (1996), tr. it. *Il Cervello emotivo. Alle origini delle emozioni*, Baldini e Castoldi, 1998.

Oltre ad essere un più che completo approccio alle ricerche delle neuroscienze, *Il cervello emotivo* risulta essere anche un'ottima palestra per analisi filosofiche sugli esperimenti, sulla completezza delle teorie e sui livelli di spiegazione scientifica. Di particolare interesse filosofico è il capitolo dedicato alle teorie del novecento sul ragionamento umano. Attraverso la storia del comportamentismo, del funzionalismo e delle macchine pensanti, già dal secondo capitolo LeDoux evidenzia l'apporto filosofico alle ricerche che in seguito analizzerà, ponendo da subito le basi per una visione di insieme. Efficace è anche l'ultimo capitolo, dove forte del *background* fin qui creato, viene proposta l'alternativa neuroscientifica al problema della coscienza.

Daniele Romano

Indice del libro

Prefazione

1. Che cosa c'entra l'amore?
2. L'anima nel freezer
3. Sangue, sudore e lacrime
4. Il Santo Graal
5. Com'eravamo
6. Alcuni gradi di separazione
7. Alla ricerca delle emozioni perdute
8. Il lato selvaggio
9. Da capo, con sentimento

Note

Bibliografia

Indice dei nomi

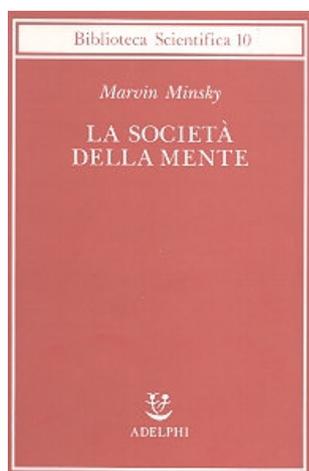
Indice tematico

Idee per una rilettura

LA SOCIETÀ DELLA MENTE

Adelphi, Milano 1989

di Marvin Minsky



Rileggere oggi il libro di Marvin Minsky del 1985 (anno dell'edizione originale) significa quasi fare quel che si dice "un tuffo nel passato". Eppure sono trascorsi solo (!) 23 anni, ma sembrano tanti; ed anche Minsky ha già superato l'ottantina (è nato a New York il 9 agosto 1927). Il guaio è che proprio la disciplina che insieme ad altri ha contribuito a fondare ha avuto una crescita rapidissima, si è allargata ed ha invaso praticamente tutti i campi della tecnologia e molti della scienza. Ed anche in ambiti più astratti, come la filosofia della

mente si è conquistata una posizione di assoluto rilievo. Il suo libro apparve subito come un'opera fondamentale, destinata a diventare un classico. Anche la struttura dell'opera era originale: 31 brevi capitoli suddivisi a loro volta in paragrafi di una o due pagine ciascuno.

Ma l'originalità era rappresentata dalla sua non-sistematicità e dalla sua ampiezza: una panoramica completa su tutti i temi costituenti il nocciolo duro del problema centrale della I.A. classica, ossia il rapporto tra la mente e il mondo, tra la mente e le "macchine".

Minsky aveva contribuito a definire ed avviare il Progetto di Intelligenza Artificiale del MIT nel 1959 insieme a McCarthy e ad altri collaboratori. Ma le differenze di ispirazione e di metodo li separarono ben presto.

Mentre McCarthy privilegiava un approccio linguistico, logico-matematico e quindi si stava impegnando ad approntare strumenti derivati dalla logica matematica per la simulazione del ragionamento umano (vedi ad es. la realizzazione del LISP, il primo linguaggio di programmazione derivato dal lambda-calcolo di Church), Minsky aveva già raggiunto la convinzione che il cosiddetto "buon senso", il comportamento intelligente, non potesse essere ridotto alle regole formali della logica ma si ispirasse, nella vita quotidiana, a criteri di tipo geometrico-analogico come il riconoscimento delle forme, le somiglianze e le metafore.

Il tentativo di riprodurre artificialmente appunto il "buon senso", proprio dal punto di vista pratico, è stato il leitmotiv che ha accompagnato tutta la carriera scientifica di

Minsky. In buona sostanza, dal suo punto di vista, l'I.A. non può che essere una scienza sperimentale, molto più vasta della psicologia (nel senso che la contiene) e soprattutto libera da vincoli teorici e metodologici.

Per quanto riguarda l'hardware, Minsky prende le distanze dalla architettura alla Von Neumann ed esplora altri modelli computazionali. Ed è uno dei suoi allievi più brillanti, D. Hillis che realizza, nel 1985, la "Connection Machine", un computer massivamente parallelo, capace di miliardi di operazioni elementari al secondo. Solo disponendo di macchine così potenti, sosteneva Minsky, sarebbe stato possibile realizzare esperimenti mentali per verificare ipotesi sulla mente e l'intelligenza umana, esperimenti finora neppure immaginabili: potremmo finalmente parlare di "epistemologia sperimentale".

La sua con Samuel Papert lo aveva infatti portato alla convinzione che neppure il modello a reti neurali potesse offrire uno schema valido per il funzionamento della mente. Ed aveva perfino contribuito a bloccare il flusso di finanziamenti per quel tipo di progetto, anche se poi aveva rivisto questo suo giudizio negativo.

Nella *Società della mente*, dove appunto riassume le tesi principali sostenute nel corso della sua attività, viene presentata e precisata la sua idea complessiva sulla struttura della mente e il concetto di *frame*.

Per Minsky la "mente" è ciò che "fa" il cervello e il cervello è una macchina di estrema complessità: nessuna descrizione lineare è in grado di render conto del suo funzionamento. Dobbiamo perciò immaginare tale funzionamento come l'effetto risultante delle azioni di un enorme numero di agenzie cerebrali indipendenti, altamente specializzate, decentrate, ciascuna delle quali agisce singolarmente, collaborando o confliggendo con le altre: una sorta di società di agenti mentali che appunto concorrono a produrre la risposta ad uno stimolo, a far riaffiorare un ricordo, a risolvere un problema o ancora a manifestare un'emozione, cioè a realizzare quella che comunemente si chiama "vita mentale". Una vita mentale, si badi bene, che per Minsky si può riassumere nel "senso comune": il comportamento di un uomo comune nella vita di tutti i giorni, in circostanze ordinarie, quotidiane.

Con il concetto di *frame* Minsky vuole proprio precisare la nozione intuitiva di "contesto", "quadro di riferimento", "situazione concreta". In pratica si tratta di realizzare un programma di tipo misto, procedurale e dichiarativo, che inquadra la situazione fornendo una intelaiatura di fatti e condizioni di tipo generale che vale per *default*, ossia "fino a prova contraria": e questo per ogni situazione presa in considerazione, in cui il soggetto può trovarsi a doversi confrontare. Il problema diventa allora quello di fornire il numero

giusto di *frames* sufficienti a garantire risposte adeguate nelle situazioni più diverse. Minsky sapeva già che questo è possibile se ci limitiamo ai famosi "micromondi": e sapeva anche che era proprio questo il limite di quella proposta.

Nel mondo reale la plasticità e l'apertura, la capacità di "bricolage" sono potenzialmente infinite, l'adattabilità e l'intuizione sono le caratteristiche fondamentali della mente umana nella sua attività quotidiana. E quindi com'è possibile scegliere un criterio valido per decidere la catena dei *frames* e sotto-*frames* applicabili in quella situazione? Ogni scelta sembra arbitraria e contribuisce a irrigidire un processo che deve invece restare libero da costrizioni e influenzato solo dalle condizioni di fatto nel rapporto soggetto-mondo.

Il testo di Minsky è, ovviamente, ricchissimo di osservazioni e di spunti su tutta la tematica relativa alla architettura della mente tanto da non essere facilmente riassumibile in poche parole. Ma viene la sensazione, rileggendolo, di trovarsi di fronte ad una vera e propria esaltazione del potere equilibratore del "mercato", del tipo di quelle che ogni tanto risuonano tra gli economisti, questa volta riferita alle "agenzie mentali" anziché agli agenti economici, una specie di "liberismo-mentale", di liberismo applicato ai fenomeni mentali. Come per l'economia, già da tempo ci si è resi conto che non basta la libera concorrenza ma servono anche regole semplici e chiare per produrre risultati apprezzabili e duraturi.

Tanti nuovi tipi di approccio caratterizzano oggi lo studio della mente e non è certamente questa la sede adatta per illustrarli o discuterli. Mi basta accennare al fatto che già dagli anni immediatamente successivi alla pubblicazione del testo di Minsky è venuta proponendosi sempre più l'importanza degli aspetti spaziali e interattivi della robotica, lo sviluppo delle cosiddette nano-tecnologie ed è grandemente cresciuta l'attenzione per le caratteristiche evolutive delle macchine "intelligenti" (oggi si parla infatti di robotica evolutiva). L'idea che la mente non risieda soltanto nella massa cerebrale e che il comportamento intelligente debba la sua origine nelle strutture della corporeità elementare e nella loro interazione spaziale con la realtà, a sua volta strutturata, del mondo fisico circostante sta conquistando il primato tra i ricercatori. La teoria dei sistemi dinamici complessi e lo studio delle caratteristiche strutturali ed auto-organizzanti degli organismi viventi sono oggi al centro della scena. Allo stesso tempo si fa strada la convinzione di essere ancora molto lontani dalla soluzione dei problemi centrali dell'IA: la consapevolezza, la coscienza, l'intenzionalità.

Il grande merito di Minsky e della sua "mente sociale" sta forse proprio nell'aver chiaramente indicato già da allora che la strada era ed è ancora molto lunga, che non esistono soluzioni semplici ad un problema così complesso e che la riproduzione di una mente umana è un'impresa collettiva, difficile ed impervia. Quando il libro uscì il paradigma linguistico-funzionalista era ancora in auge e, per alcuni in particolare, prometteva grandi novità a breve scadenza. Come tutti i grandi creatori, a mio avviso, Minsky ha avuto il grande merito di non seguire mai il pensiero corrente, di non affezionarsi troppo alle sue stesse idee e di mantenere vivo un positivo scetticismo in relazione ai facili entusiasmi.

E mi piace concludere sottolineando che forse la vera ricerca ha bisogno tanto di quelli che stabiliscono punti fermi quanto di quelli che mantengono aperte le strade del dubbio e della fantasia, di quella sana intuizione disposta a rimettere in discussione anche le verità più ovvie.

Umberto Maionchi

Idee per una rilettura

EMPIRISMO E FILOSOFIA DELLA MENTE

Einaudi, Torino 2004

di Wilfrid Sellars

1. Introduzione

In *Empirismo e filosofia della mente*, Wilfrid Sellars conduce una critica all'idea che esistano delle entità delle quali abbiamo consapevolezza in modo immediato, idea riassunta sotto la formula di "mito del Dato". Sellars non si riferisce con questo alla sensibilità che mostriamo per certe caratteristiche percettive degli oggetti. La critica è piuttosto diretta contro l'ipotesi che esistano delle credenze esplicite, tipicamente espresse in giudizi osservativi, che non presuppongono altre credenze e che hanno un valore fondativo per l'intero edificio del sapere.

L'argomentazione è così orientata contro l'uso del concetto di dato in teoria della conoscenza. Sellars ritiene che questa abbia una forma essenzialmente linguistica: è riferita a fatti, la struttura formale dei quali è espressa nel linguaggio. Secondo Sellars, inoltre, la conoscenza implica la pratica del dare e chiedere ragioni: un enunciato conosciuto deve essere la premessa, o la conclusione in un ragionamento di tipo inferenziale. Questa visione conduce – allo stesso modo di quanto fatto negli stessi anni da Quine in *Due dogmi dell'empirismo* – a una concezione olistica del sapere, nel quale non vi sono fondamenta ultime ma ogni parte contribuisce a giustificare le altre ed è dalle altre giustificata.

L'impostazione anti-fondazionalista costituisce il nucleo centrale della critica alla tradizione empirista, tanto nella sua forma classica, quanto nella versione dei neopositivisti logici: ignorando il carattere dinamico e auto-correttivo dell'indagine conoscitiva, essa assume l'esistenza di un insieme di enunciati tra loro indipendenti (enunciati protocollari) che costituiscono le conoscenze di base a partire dalle quali tutte le altre vengono inferite (atomismo). Tale insieme è ritenuto stabile e non rivedibile: i suoi enunciati non possono essere giustificati da nessun'altra proposizione. Ecco perché si arriva alla conclusione che essi si debbano fondare direttamente nell'esperienza. Sposando la posizione che definirà come "nominalismo psicologico", Sellars critica invece l'ambiguità degli enunciati osservativi, che da una parte si basano sull'esistenza di contenuti sensoriali materiali, ottenuti attraverso i sensi, e dall'altra pretendono di definire una conoscenza non inferenziale di tipo normativo.

L'impostazione anti-fondazionalista si rivolge però anche contro l'epistemologia razionalista cartesiana. Nel suo tentativo di superare il dubbio scettico, infatti, Descartes fonda la conoscenza su di un insieme di enunciati relativamente ai quali il soggetto conoscente non può essere né ignorante, né in errore: se li crede sa anche di crederli, e viceversa. La trasparenza e l'infallibilità che caratterizzano tale atteggiamento portano all'identificazione tra vita psichica e vita cosciente: ciò che è osservabile alla coscienza è anche da essa conoscibile. Tale conclusione continua a identificare la percezione di contenuti sensoriali e la loro conoscenza, mostrando così come la prospettiva empirista e quella cartesiana condividano le stesse problematiche.

Il problema fondamentale è che il riconoscimento di un dato prevede l'associazione di un predicato a un insieme di particolari che si assomigliano tra loro, e tale associazione è un atto epistemico che avviene all'interno del linguaggio. Facendo riferimento a una capacità di osservazione di carattere non teorico, tanto gli empiristi quanto Descartes assumono uno spazio logico strutturato e familiare la cui consapevolezza è pre-linguistica. Per Sellars, al contrario, non può esistere una qualsiasi consapevolezza di tipi e rassomiglianze antecedente, o indipendente, dall'acquisizione della lingua. L'osservazione ha un intrinseco carattere teorico che le conferisce immediatamente un valore epistemico; bisogna invece rinunciare a parlare di un contenuto sensoriale dell'esperienza.

Questa posizione non intende negare valore intersoggettivo agli enunciati di percezione, e in questo si discosta tanto dall'eliminativismo comportamentista di Ryle, quanto dalla critica di Wittgenstein. Rispetto al primo, che aveva negato l'esistenza delle impressioni sensoriali sulla base della loro natura non manifestamente osservabile, Sellars replica che il concetto di impressione sensoriale ha comunque uno statuto teorico, che conferisce una controllabilità intersoggettiva alle asserzioni nelle quali compare; la mancanza di valore osservativo non è quindi sufficiente a privare un concetto di ogni valore funzionale. Rispetto al secondo, che aveva escluso la conoscibilità degli episodi privati sottolineando la natura pubblica del linguaggio, Sellars ribatte mettendo in luce proprio la natura linguistica intersoggettivamente afferrabile delle impressioni sensoriali, che le rende oggetto possibile di conoscenza: si deve piuttosto dubitare del loro carattere privato. Alla critica e al tentativo di destrutturazione del concetto di dato sensoriale, Sellars fa quindi seguire una proposta alternativa che ne concili il carattere intersoggettivo con quello privato.

2. La critica al mito del Dato

Contro il mito del Dato Sellars svolge due argomenti principali, uno basato sulla contraddittorietà del concetto, l'altro su di una spiegazione alternativa dei contesti linguistici nei quali il termine di impressione sensoriale compare. Riguardo al primo, Sellars parte dall'individuazione della classe degli episodi interiori, alla quale appartengono le impressioni sensoriali e i pensieri. Egli nota che raccogliere sotto uno stesso genere due tipi di entità così diverse è problematico, dal momento che le prime derivano un processo di percezione che coglie particolari, mentre i secondi esprimono atti di conoscenza riferita ai fatti. Inoltre, mentre l'esperienza di un contenuto sensoriale sembra darsi in maniera immediata, il possesso di una conoscenza presuppone il passaggio attraverso un processo d'apprendimento. I due generi di episodi interiori sono quindi, secondo Sellars, apparentemente ben distinti; il teorico dei dati materiali, invece, li collega sostenendo che «la sensazione di un triangolo rosso è il paradigma della conoscenza empirica» (§7). Al contrario, Sellars rileva l'ambiguità del concetto di dato materiale: le impressioni sensoriali da una parte sono il frutto di osservazioni, e quindi di un processo *fisico-causale* attraverso il quale il soggetto conoscente viene in contatto con il mondo; dall'altra hanno valore *normativo-epistemico*, in quanto costituiscono la base attraverso la quale le osservazioni di una teoria scientifica vengono collegate con il mondo. Tali aspetti sono difficilmente assimilabili.

Il secondo importante argomento contro il mito del Dato tenta di chiarire in che modo si arriva a inferire l'esistenza delle impressioni sensoriali. Vediamo qui un caso paradigmatico dell'analisi concettuale di Sellars: egli trasferisce una questione ontologica sull'esistenza, o meno, delle impressioni sensoriali all'analisi linguistica degli enunciati che parlano di queste entità (gli enunciati del tipo 'x appare F a i al tempo t'), fornendo così una spiegazione funzionale che non ha più bisogno di presupporle. Sellars propone come esperimento mentale il caso di John, un commesso che lavora in un negozio di cravatte. Dopo che la luce elettrica è stata installata nel suo negozio, John è stupito dal fatto che una cravatta che vede verde sotto la lampada, sia in realtà blu se esposta alla luce del sole. Con il tempo, tuttavia, egli impara a riconoscere il colore della cravatta all'interno del negozio dicendo che essa è blu, nonostante *appaia* verde. Il concetto di impressione sensoriale trova la sua spiegazione in questa pratica linguistica. Siamo portati a credere che John stia riportando un fatto oggettivo dentro di sé quando dice che la cravatta gli sembra verde; in realtà sta semplicemente riportando il risultato del ragionamento che sotto alla luce della lampada le cose non sono del loro colore. Il predicato 'apparire', o

‘sembrare’, non esprimono pertanto una conoscenza diretta di tipo osservativo, essi fanno la loro comparsa nel linguaggio per allentare l’impegno ontologico in relazione ad alcune asserzioni, che possono essere così *affermate* senza essere *sottoscritte*.

L’esempio di John è volto a mostrare che il linguaggio delle impressioni sensoriali origina all’interno della pratica linguistica dell’apparire. Questa constatazione fonda il *nominalismo psicologico* di Sellars, in base al quale non è sensato parlare del riconoscimento di proprietà se non facendo riferimento agli enunciati di percezione attraverso cui tale riconoscimento si esprime, e dunque qualsiasi consapevolezza di tipi e rassomiglianze - quindi qualsiasi capacità di categorizzazione degli stimoli - è epistemicamente connotata e dipende dall’acquisizione della lingua. Sellars identifica così l’impressione sensoriale con l’enunciato osservativo che ne esprime il riconoscimento: l’osservazione si spoglia della sua tradizionale natura empirica per acquisire un carattere intrinsecamente teorico. Di conseguenza appare anche illusorio che si possa caratterizzare in modo diretto il contenuto descrittivo di un’esperienza facendo riferimento alle sue qualità ; esso è piuttosto un’entità teorica che permette di assimilare casi diversi di percezione, come quelli in cui vediamo che qualcosa è di un certo tipo e quelli in cui affermiamo soltanto che le cose ci appaiono in un certo modo.

La tesi della natura epistemica dell’osservazione motiva anche l’olismo sellarsiano: se i processi causali di percezione di particolari determinati sono tra loro indipendenti, non lo è invece l’attribuzione di proprietà. Per poter affermare che qualcosa è rosso, per esempio, un soggetto deve conoscere: (i) l’intero campo di variazioni (i colori); (ii) le corrette circostanze per l’attribuzione della proprietà (il fatto che i colori si osservano alla luce del sole); (iii) il fatto di essere nelle circostanze corrette per l’attribuzione (il fatto di essere alla luce del sole). Soltanto basandosi sulla capacità riflessiva espressa dal terzo punto si può, secondo Sellars, distinguere la vera e propria capacità linguistica, che si accompagna al riconoscimento dell’autorità delle proprie asserzioni, dalla disposizione a produrre risposte verbali adeguate in risposta a determinati stimoli - quella che caratterizza il comportamento di un pappagallo, che dice ‘rosso’ quando vede un oggetto rosso, o di un fotodiodo che emette un segnale su frequenze diverse a seconda che capti o meno un raggio di luce. Tuttavia le iniziali attitudini linguistiche del bambino sembrano proprio di questo secondo tipo; bisognerebbe concluderne che questo non è in grado di attribuire proprietà alle cose. Ma allora in che momento si può invece considerare che il bambino ha appreso la lingua? La risposta di Sellars è che semplicemente un parlante impara a un certo punto a dare ragione dei propri resoconti d’osservazione, e che solo da quel

momento in poi i suoi giudizi possono essere ritenuti degli atti di osservazione. Una risposta che lascia piuttosto insoddisfatti.

3. Scienza e senso comune

Chiarita la natura teorica dei giudizi di osservazione, Sellars cerca di spiegare come il linguaggio degli episodi interiori si possa sviluppare all'interno di un linguaggio osservativo. Questo obiettivo lo spinge a fare alcune precisazioni sulla natura del linguaggio scientifico. Una teoria può essere presentata in modi diversi: da un punto di vista logico, essa è un dominio di entità il cui comportamento è descritto da alcuni principi fondamentali, e mette in correlazione complessi di entità teoriche ad oggetti e situazioni osservative. Nella pratica scientifica, tuttavia, le assunzioni fondamentali non vengono di solito elaborate come principi astratti, ma a partire da un modello, che descrive un ambito di fenomeni familiari e stabilisce per mezzo di un commentario un'analogia ben definita tra questi e i fenomeni da spiegare. Se quindi nella versione classica delle teorie scientifiche i termini teorici sono descrizioni definite che ne abbreviano la presentazione, attraverso la nozione di modello essi sono anche posti in relazione con oggetti del mondo, e in quanto tali sembrano assumere un contenuto. Questa seconda possibilità spinge a equivocare il ruolo delle impressioni e dei pensieri una volta che vengono introdotti come termini teorici nel linguaggio del senso comune, e porta a disconoscere la loro natura teorica.

Sellars si ritrova quindi a dover chiarire il rapporto tra linguaggio scientifico e quello di senso comune. Per l'uomo della strada la via privilegiata alla conoscenza degli oggetti è quella del riferimento ostensivo: in questa prospettiva è ovvio che l'autorità del discorso teorico-scientifico non può che essere vista come derivata e strumentale. La distinzione tra linguaggio scientifico e di senso comune può tuttavia essere vista anche come sfumata. La scienza, infatti, trae le proprie radici nel senso comune, e allo stesso tempo ne sostituisce gradualmente l'ontologia. I termini teorici non servono solo a impostare la spiegazione di un campo di fenomeni, ma rimangono anche aperti all'interpretazione di più discipline che operano su livelli di descrizione diversi, favorendo l'integrazione delle singole scienze in un'immagine unitaria. È quindi perfettamente possibile che alcuni concetti teorici divengano in un secondo momento parte della concezione del mondo del senso comune e per questa via, essi possono arrivare a dotarsi anche di un carattere osservativo. La distinzione tra teorico e osservativo, dunque, si riferisce ai criteri che si hanno a disposizione in un determinato momento per identificare dei concetti all'interno

di una disciplina scientifica, e non deve essere vista come la norma che stabilisce quali oggetti esistono effettivamente.

4. Il mito di Jones

Alla luce di queste considerazioni, il comportamentismo non si definisce per la pretesa di eliminare i termini mentali riducendoli alla loro manifestazione comportamentale, quanto piuttosto come metodologia rigorosa e intersoggettivamente verificabile per la definizione del vocabolario psicologico. Questa considerazione permette di comprendere in che modo deve essere intesa la spiegazione sellarsiana del lessico mentale: esso ha un'importante funzione teorica, che lo fonda prescindendo dal problema dei correlati fisici dei suoi termini. Egli propone un altro esempio immaginario, il celebre "mito di Jones". Il racconto inizia con la descrizione di una comunità di uomini primitivi, che possiedono un linguaggio dotato di un vocabolario descrittivo per oggetti e proprietà del mondo esterno, tale linguaggio (definito da Sellars "ryleano") comprende le abituali nozioni semantiche (verità, significato, riferimento, ecc.), e la possibilità di introdurre termini teorici.

Un giorno, all'interno di questa primitiva comunità viene alla luce un comportamentista *ante litteram*, Jones, che per spiegare e prevedere il comportamento dei propri compagni sviluppa una teoria in base alla quale i loro proferimenti verbali manifesti non sono altro che il momento culminante di certi episodi interiori, i pensieri. Tali episodi interiori sono entità teoriche, e sono quindi privi di un corrispettivo osservativo. Essi sono, tuttavia, costruiti sul modello degli episodi linguistici, ed ereditano da questi l'applicabilità delle categorie semantiche al mondo, fatto che ne spiega il carattere intenzionale. Una volta che la teoria di Jones inizia a circolare, continua Sellars, si scopre che banalmente un individuo fornisce resoconti attendibili su di sé senza osservare il proprio comportamento, avvalorando l'idea ingenua che questa capacità nasca da un accesso privilegiato ai propri pensieri, e al disconoscimento della natura teorica di tali entità.

Per spiegare i casi di errore percettivo, Jones sviluppa anche una teoria della percezione sensibile, in base alla quale esiste una classe di episodi interiori, le impressioni sensoriali, che sono i risultati finali dell'azione esercitata dagli oggetti fisici sugli organi di senso, ma che si possono dare anche in assenza di questi. Jones insegna agli altri esseri umani la teoria delle impressioni, e la relativa pratica linguistica, cosicché in breve tempo tutti sono in grado di collegare i propri enunciati di osservazione alla percezione di determinate impressioni. Tuttavia, questa correlazione non dovrebbe essere presa alla lettera. In quanto entità teoriche, le impressioni sono definite come stati del soggetto epistemico, e le loro

proprietà sono definite dai postulati della teoria. Tuttavia la teoria percettiva di Jones è formulata nei termini di un modello, il mondo esterno. Se l'analogia istituita dal modello viene forzata, le impressioni possono essere viste come dei particolari, e corrispondentemente si può credere di percepirle, come pure pensare che ereditino delle proprietà intrinseche dalle entità di cui il modello è costituito. Questo errore introduce nel linguaggio delle impressioni la logica del linguaggio dei pensieri (concettualismo).

Il mito di Jones giustifica il linguaggio mentale sottolineandone la natura teorica e il valore funzionale; questi aspetti fondano anche la sua autonomia e l'irriducibilità completa al comportamento manifesto. Tuttavia il carattere linguistico della teoria degli episodi interiori fa sì che le entità mentali denotino in primo luogo concetti pubblici: il loro carattere intersoggettivo è originario, e la loro privatezza soltanto derivata da questo. La natura teorica delle impressioni e dei pensieri spiega dunque come si può coniugare la *privatezza* dell'accesso privilegiato all'*intersoggettività* dei contenuti degli episodi interiori. Il mito mostra anche come si arriva all'errore di credere nell'esistenza di episodi del genere di quelli interiori: non vedendo che il linguaggio ryleano non è soltanto un modo per parlare di fatti, ma contiene in sé anche una spiegazione teorica, il teorico dei dati materiali finisce per fondare la sua teoria della conoscenza delle ipotesi teoriche pregresse.

5. Attualità di *Empirismo e filosofia della mente*

All'epoca in cui è stato scritto, *Empirismo e filosofia della mente* ha dato un contributo fondamentale al tramonto del programma neo-empirista. Questa sezione si propone di valutarne il valore in relazione al dibattito filosofico contemporaneo.

La critica alla nozione di dato sensoriale rimane valida nei confronti di qualunque programma fondazionalista che tenti di recuperare la visione scientifica del mondo a partire da un linguaggio fenomenista, sia questo rappresentato dalla teoria di Ayer (1940, 1954) - con la quale Sellars si confronta all'interno del saggio - sia esso il linguaggio impiegato nell'analisi de *La costruzione logica del mondo* di Carnap (1928), oppure quello impiegato da Wittgenstein nel *Big typescript*, o ancora la prospettiva portata avanti dalla fenomenologia husserliana. Tuttavia la critica rimane circoscritta alla concezione della conoscenza sulla quale si fonda, privilegiandone il carattere normativo rispetto all'osservazione di casi singolari. Se da una parte questo permette di giustificare l'aspirazione universalistica della conoscenza, dall'altra rimane aperto il problema del modo in cui noi esseri umani possiamo entrare in quello "spazio delle ragioni" nel quale la conoscenza si colloca. Nella nostra vita mentale, infatti, si incontrano soltanto particolari

semplici e irripetibili. Da un punto di vista strettamente fisico il nostro comportamento nei confronti di questi è pienamente descrivibile nei termini di una reazione disposizionale. Nella visione di Sellars non è affatto chiaro che cosa porti il bambino, che deve ancora apprendere una lingua, a fare quel salto che gli permette di entrare nello spazio delle ragioni. La pratica sociale del giustificare le proprie affermazioni sembra la risposta più plausibile. Eppure tale pratica può anche essere vista come l'acquisizione di disposizioni comportamentali più complesse (la capacità di attuare quei comportamenti verbali che corrispondono al gioco linguistico del dare e chiedere spiegazioni); anche un pappagallo, o un fotodiodo potrebbero in linea di principio imparare a fornire ragioni del proprio comportamento. Sellars rimane così vittima di una concezione della conoscenza idealizzata, che può essere non raggiungibile da un agente cognitivo ordinario. Al contrario, gli approcci fenomenisti sopraccitati propongono una spiegazione costruttiva di come si possa ottenere una nozione generale di conoscenza a partire dalla conoscenza di particolari singoli e irripetibili.

L'analisi di Sellars può essere anche utilizzata per confrontare la nozione di dato con quella di sensazione corporea, così come è descritta dalla fisiologia e utilizzata nelle scienze cognitive per indicare l'*input* fisico capace di attivare i recettori sugli organi di senso di un organismo. Allo stesso modo del concetto di dato, anche quello di sensazione presenta una duplice lettura, che può essere orientata in modo più o meno naturalistico. In quanto stimolo fisico essa indica semplicemente l'informazione che un agente cognitivo riceve dall'ambiente, indipendentemente dall'uso che ne può fare per costruirsi una rappresentazione del mondo; in quanto sensazione di qualcosa, invece, essa è dotata di un contenuto intenzionale e permette di mettere in contatto un sistema cognitivo con l'ambiente nel quale vive. Su questo secondo aspetto si può agganciare la critica di Sellars al dato materiale, è infatti possibile descrivere tale contenuto in termini che non sono utilizzabili dall'agente stesso, introducendo quindi una forma di concettualizzazione che fa parte del punto di vista teorico dell'osservatore del sistema cognitivo più che del sistema stesso. Il nominalismo psicologico di Sellars sottolinea che in questo caso le distinzioni concettuali non devono essere prese come la descrizione di uno stato dell'agente cognitivo; è un richiamo che sembra essere un'anticipazione dell'*atteggiamento intenzionale* su cui ha poi notevolmente insistito Dennett (1987). Se si tiene presente questa differenza tra il punto di vista dell'organismo e quello dell'osservatore del sistema organismo-ambiente, tuttavia, si può ancora parlare della capacità di un organismo di categorizzare l'informazione ambientale; il concetto di sensazione svolge in questo senso

un importante ruolo euristico, e non può essere accantonato dalla critica di Sellars, limitata all'ambito della teoria della conoscenza.

L'analisi di Sellars fornisce infine un contributo ancora valido contro possibili velleità riduzioniste dei termini della psicologia del senso comune. Il lessico psicologico, infatti, trova una propria autonomia nel linguaggio comune. Interpretare il rapporto tra questo e il linguaggio della neurofisiologia in termini di riduzione significa invece disconoscere il ruolo funzionale del discorso psicologico del senso comune. Secondo Sellars il modo giusto di procedere è valutare in che misura può essere ottenuta una correlazione tra il linguaggio delle neuroscienze e quello della psicologia. In tale senso la sua analisi è ancora attuale e riconosce la specificità del contributo che la filosofia può dare nelle scienze cognitive.

Marco Fenici

Bibliografia

1. Ayer (1940), *Foundations of Empirical Knowledge*, Macmillan, London.
2. Ayer (1954), *The terminology of sense data*, in *Philosophical Essays*, Macmillan, London, pagg. 66-104.
3. Carnap (1928), *Der logische Aufbau der Welt*, Berlin-Schlachtensee 1928.
4. Dennett (1987), *The intentional stance*, MIT Press, Cambridge, Massachusetts.
5. Sellars (1956), *Empiricism and the philosophy of mind*, trad.it, (2004), *Empirismo e filosofia della mente*, Einaudi, Torino.

Indice

Introduzione di Richard Rorty

I. Un'ambiguità delle teorie del dato sensoriale

II. Un'altra lingua

III. La logica di 'appare'

IV. La spiegazione dell'apparire

V. Impressioni e idee: una considerazione logica

VI. Impressioni e idee: un'osservazione storica

VII. La logica di 'significa'

VIII. La conoscenza empirica ha un fondamento?

IX. La scienza dell'uso ordinario

X. Episodi privati: il problema

XI. I pensieri: la concezione classica

XII. I nostri antenati ryleani

XIII. Teorie e modelli

XIV. Comportamento metodologico versus filosofico

XV. La logica degli episodi privati: i pensieri

XVI. La logica degli episodi privati: le impressioni

Indicazioni bibliografiche

Guida al testo di Robert Brandom

Idee per una rilettura
NEUROFILOSOFIA.
**VERSO UNA SCIENZA UNIFICATA DELLA MENTE E
DEL CERVELLO.**

MIT press, Cambridge, 1986
di Patricia Churchland

1. Introduzione

In the mid-seventies I discovered that my patience with most mainstream philosophy had run out.

(P. Churchland, 1986)

Possono le conoscenze biologiche rivelarsi ininfluenti rispetto alla costruzione di una teoria delle proprietà mentali? Può una scienza della coscienza affermarsi indipendentemente dagli sviluppi delle neuroscienze cognitive? Queste le domande alle quali Patricia Churchland tenta di rispondere all'interno dei lavori scritti a partire dalla prima metà degli anni Ottanta¹. Tra questi, *Neurophilosophy* (1986) rappresenta il primo tentativo di delineare un percorso di ricerca in grado di unire ambiti conoscitivi tradizionalmente separati, facendosi interprete di una tendenza teorica maturata sia grazie allo sviluppo delle nuove tecniche d'indagine dedicate allo studio del cervello, sia grazie agli incoraggianti risultati sperimentali ai quali quest'ultime hanno condotto.

Nell'introdurre i diversi aspetti che costituiscono l'opera di Patricia Churchland è indispensabile considerarne innanzitutto le intenzioni polemiche. Contro la banalizzazione del contributo scientifico nel campo della filosofia della mente, la Churchland propone una visione naturalizzata del rapporto mente-corpo, impegnandosi nella difesa di una concezione *unificata* delle conoscenze relative sia alla sfera coscienziale, sia alla sfera biologica. Domande che hanno fatto la storia della filosofia, riguardanti la natura delle rappresentazioni e della conoscenza umana, ma anche la possibilità del libero arbitrio e della causalità mentale, troverebbero risposta all'interno di una teoria *naturale* delle proprietà fisiche e mentali, in grado di superare il baratro che per secoli ha tenuto separate nozioni come quella di coscienza e di corporeità. Secondo la Churchland, anche all'interno di un dibattito tradizionalmente dominato da riflessioni filosofiche, lo studio dei

¹ Scrive la Churchland già nel 1980: "Assuming that the cognitive activity of an organism is brain activity, why should brain-based theories of such activity be thought to be the wrong kind of theory – indeed, how could they fail to be the right sort of theory?" (Churchland, 1980).

rapporti che intercorrono tra dominio del mentale e manifestazioni comportamentali passa attraverso il riconoscimento del ruolo di primo piano svolto dalle conoscenze neurobiologiche.

La proposta di Patricia Churchland è contenuta, oltre che all'interno di una serie di articoli pubblicati a partire dal 1980, in due volumi, il già citato *Neurophilosophy* (1986) e il più recente *Brain Wise* (2002). Di seguito saranno trattati alcuni tra i temi che più contraddistinguono la prospettiva "neuro-filosofica" della Churchland², verranno evidenziati alcuni possibili fraintendimenti legati alla concezione riduzionista e solo in conclusione sarà prospettata la possibilità di un'interazione tra analisi fenomenologica e neuroscienze più liberale di quella concepita dalla stessa Churchland.

2. Il riduzionismo

The most convincing answer to skepticism is, of course, explanatory progress in neuroscience.

(P. Churchland, 2002)

Neurophilosophy è il tentativo di risolvere entro il contesto naturalistico i tradizionali problemi che affliggono la comprensione del rapporto tra stati mentali e stati corporei. Ciò avviene attraverso la proposta di una *teoria unificata* all'interno della quale la conoscenza dei processi neuronali riveste un *ruolo esplicativo* rispetto all'insieme di fenomeni complessi generalmente catalogati sotto l'etichetta "mente". Svolgono così una funzione di primo piano le procedure di *riduzione*, atte a tradurre e unificare i molteplici concetti che contraddistinguono la nozione di mente.

Il processo di riduzione si delinea, innanzitutto, quale rapporto tra due teorie esplicative. Secondo la Churchland, una nuova teoria riducente deve essere in grado di mappare le espressioni della relativa teoria ridotta, traducendole all'interno del proprio dominio linguistico e dimostrando al contempo una più adeguata capacità esplicativa.

When is an old theory successfully reduced by a new one? The fast answer, to be elaborated anon, is this: a reduction obtains when the *terms* of the old theory can be mapped onto (i.e. paired with) expressions in the new theory, in such a way that the basic laws of the old theory are thereby mapped onto sentences in the new theory, where these sentences are logical consequences of the basic laws of the new theory. In effect, an image of the old reduced theory is generable in the new reducing theory, which means that the new theory can explain

² Le due maggiori opere di Patricia Churchland sono entrambe corredate da una ricca introduzione ai temi fondamentali dell'indagine neuro scientifica ed epistemologica; si tratta di capitoli all'interno dei quali non sono presenti tratti salienti della sua proposta filosofica, ma solo elementi propedeutici a un'adeguata comprensione dei temi trattati. Per questa ragione non mi soffermerò nell'analisi di tali parti.

what the old theory explained and can explain why the old theory worked as well as it did, and of course the new theory typically explains much where the old theory was buffaloes.
(*Mind brain reduction, new light from the philosophy of science*, 1981)

Sono due le caratteristiche principali della concezione riduzionista adottata dalla Churchland:

1) Il rapporto di riduzione si delinea tra due contesti teorici e comporta la traduzione del linguaggio adottato dalla teoria ridotta nel linguaggio della teoria riducente (Churchland, P., 1986, pp. 282, 283);

2) Il rapporto di riduzione tra due teorie comporta che la teoria riducente sia in grado di spiegare nuovi fenomeni oltre alla gamma di fenomeni oggetto della teoria ridotta³ (Churchland, 1986, pp. 293, 294);

In prima approssimazione, la proposta della Churchland si concretizza attraverso la definizione di un rapporto di *riduzione metodologica*, in cui una pratica d'indagine mirata a ottenere certi risultati, nel caso specifico la teoria *psicologica del senso comune* (Folk Psychology), viene sostituita da un altro dominio di conoscenze, le *neuroscienze*, ritenuto contraddistinto da più idonee competenze esplicative (Churchland P., 1986, p. 295; Churchland P., 1982, p. 1044).

Con il richiamo a una *psicologia del senso comune*, la Churchland intende riferirsi a un insieme di concezioni teoriche ampiamente condivise a livello extra-scientifico, all'interno delle quali stati mentali come le *intenzioni*, i *desideri* e le *credenze*, vengono impiegati in quanto strumenti esplicativi, al fine di permettere la comprensione e la previsione dei più diversi atteggiamenti comportamentali umani⁴ (Churchland, P., p. 299; Churchland, P., 1982, p. 1044). Contrariamente ai tentativi di fondare la psicologia del senso comune su momenti di evidenza intuitiva, la Churchland esclude la possibilità che il soggetto possa usufruire di flussi d'informazione privi di una qualche mediazione concettuale, ponendo l'accento sul *carattere teorico* che accompagna indissolubilmente ogni processo introspettivo⁵. In questo modo, *al pari di ogni altra teoria scientifica*, anche gli assunti psicologici caratteristici del senso comune risultano potenzialmente soggetti a processi di *revisione e falsificazione*, senza che alcuna condizione epistemologica "speciale" sia

³ La Churchland assume qui una concezione deduttivo-nomologica, per certi versi affine alla classica definizione di E. Nagel.

⁴ Proprio la connotazione *teorica* attribuita al senso comune risulta una caratteristica determinante per lo sviluppo del progetto riduzionistico della Churchland, proiettando inoltre tale prospettiva nell'attuale dibattito concernente la natura delle "teorie della mente", diviso tra sostenitori di una *concezione teorica* del mentale (Theories of theory of mind) e sostenitori di una *concezione simulazionista* (Simulation theory of mind).

⁵ "Introspection, like all observations, is observation through the lens of theory" (Churchland, P., 1981, p. 1046).

riconosciuta alla consapevolezza in prima persona dei propri stati mentali (Churchland P., 1986, pp. 305-307). In forza del carattere teorico attribuito alla psicologia del senso comune, la Churchland giunge a delineare il processo di riduzione che dalla teoria mentalistica conduce al dominio delle neuroscienze (Churchland, P, 1982, p. 1044). Così:

If we see that folk psychology has no right to epistemological privilege, and no immunity to revision and correction, than we can begin to see that its generalization and categories can be corrected and improved upon.

Once it is recognized that folk psychology is not immune to scientific improvement, this reveals the possibility what will eventually reduce to neuroscience are generalizations of scientific psychology that have evolved a long way from the home "truths" of extant folk psychology.

(*Neurophilosophy*, 1986, pp. 311, 312)

Il percorso seguito dalla Churchland prosegue attraverso la ricerca di correlazioni osservative che mettano in grado di tradurre i termini di un *linguaggio mentalistico*⁶, tipico della psicologia ingenua, nel più rigoroso linguaggio delle neuroscienze. Si palesa in questo modo l'importanza che lo sviluppo delle tecniche sperimentali riveste nella definizione della proposta riduzionista. L'associazione tra specifici stati cerebrali e stati mentali, quali i desideri e le credenze, passa infatti attraverso la sempre più raffinata capacità d'indagine fisiologica. Proprio l'accumularsi delle evidenze che pongono in correlazione stati mentali e stati neuronali permette alla Churchland di compiere il passo successivo verso la riduzione della teoria psicologica del senso comune al campo delle neuroscienze. La comparsa di correlazioni richiede infatti la messa a punto di un *panorama esplicativo* che, solo l'accettazione di un livello d'*identificazione* tra istanze di stati mentali e istanze di meccanismi neuronali, sembrerebbe in grado di fornire (Churchland, P., 2002, p. 154).

Il ragionamento seguito dalla Churchland rappresenta un tipico caso di *inferenza alla miglior spiegazione*:

⁶ Con la nozione di linguaggio mentalistico intendo riferirmi a linguaggi che annoverano termini concernenti stati mentali distinti da termini concernenti stati fisiologici. Il linguaggio quotidiano ne è un tipico esempio.

1. Sono riscontrate correlazioni osservative tra l'utilizzo di termini mentali nell'ambito della psicologia del senso comune e la funzione di specifici meccanismi neuronali;
2. Se le asserzioni che costituiscono la teoria psicologica del senso comune venissero identificate con asserzioni relative alle conoscenze fisiologiche concernenti il funzionamento del sistema nervoso, *allora* ne conseguirebbe la possibilità di riscontrare correlazioni osservative tra linguaggio mentale e specifici meccanismi neuronali;
3. La riduzione della teoria psicologica del senso comune alle neuroscienze è la miglior spiegazione della comparsa di correlazioni osservative tra l'impiego di termini relativi a stati mentali e funzioni di specifici meccanismi neuronali;

In altre parole, secondo la Churchland, portando l'analisi sul piano delle prerogative esplicative, sembra potersi giustificare l'avvicendamento tra una concezione teorica mentalistica e un modello naturale delle mente. Proprio la ricerca di un livello d'identificazione della sfera mentale, coincidente con il piano dei meccanismi neurologici, renderebbe conto di un insieme di riscontri osservativi che la teoria del senso comune non appare in grado di spiegare. Le correlazioni sperimentali tra stati mentali e stati cerebrali, ma anche la co-evoluzione⁷ riscontrabile tra aspetti della psicologia scientifica e delle neuroscienze, appaiono fenomeni comprensibili a partire da un'impostazione *monistica* del rapporto mente-corpo, all'interno della quale l'unico dominio dotato di valore esplicativo risulta quello rappresentato dalle proprietà biologiche del sistema nervoso.

Introdotta il ragionamento abducente alla base della tesi riduzionista, il passo verso una forma di *materialismo eliminativista* è breve. Una volta accettata la possibilità che espressioni del linguaggio mentale siano associate a stati di natura cerebrale e una volta assegnata superiorità esplicativa ai resoconti delle neuroscienze, l'accantonamento di ogni teoria mentale appare inevitabile.

Seguendo lo schema introdotto dalla stessa Churchland, la nozione di materialismo eliminativista (*eliminative materialism*) si raccoglie attorno alle seguenti assunzioni:

- 1) That folk psychology is a theory;

⁷ Scrive la Churchland a proposito della co-evoluzione di neuroscienze e psicologia: "The coevolutionary development of neuroscience and psychology means that establishing points of reductive contact is more or less inevitable" (Neurophilosophy, p. 374). Si tratta per la Churchland di una possibilità vagliabile solo a posteriori, potenzialmente limitata dal progresso delle tecniche d'indagine o da successive scoperte empiriche.

- 2) That is a theory whose inadequacies entail that it must eventually be substantially revised or replaced outright (hence “eliminative”);
- 3) That what ultimately replace folk psychology will be the conceptual framework of a mature neuroscience (hence “materialism”);

(*Neurophilosophy*, 1986, p. 396)

Così come avvenuto per la teoria del flogisto in ambito chimico, o per la teoria vitalistica in ambito biologico, anche la concezione mentalistica, nonostante il livello delle conoscenze ancora non sia pienamente maturo, appare agli occhi della Churchland sbagliata, pertanto destinata a essere sostituita con una più adeguata concezione scientifica. Proprio l'accantonamento del linguaggio impiegato dalla teoria psicologica del senso comune, in quanto ritenuto inadatto a rendere conto delle regolarità osservative che caratterizzano il rapporto tra impiego di termini mentali e riscontro di specifici stati cerebrali, permette infine di oltrepassare una forma di riduzionismo metodologico, approdando a una più radicale forma di riduzionismo *ontologico*.

Ontological simplification – simplification concerning what kinds of things we think exist – is thus a typical feature of reduction [...]the simplification may be achieved either through reduction and identification or through elimination.

(*Neurophilosophy* , 1986, p. 280)

Prima di procedere nell'analisi di ulteriori aspetti che contraddistinguono la proposta di Patricia Churchland, è opportuno dissolvere sul nascere alcuni possibili fraintendimenti riguardo alla nozione di riduzionismo che si è fin qui delineata:

a) In primo luogo occorre osservare che le argomentazioni riduzioniste della Churchland non si prestano a un'interpretazione causale del rapporto tra mente e corpo. All'osservazione di correlazioni tra riferimenti mentali e stati fisiologici, la Churchland non fa conseguire alcuna diretta corrispondenza causale. La prospettiva riduzionista si basa piuttosto sul ricorso a un'argomentazione ipotetica di tipo esplicativo che conduce all'eliminazione del riferimento a termini mentali. Il tradizionale problema posto dall'*interazione causale* tra mente e materia viene in questo modo eluso.

The classical mind–body problem was how the nonphysical stuff that makes up the immaterial soul can causally interact with the material stuff that is the body [...]. But we can see now that interaction is a pseudo problem, like the problem of how the crystal spheres of the heavens daily rotate, or how the heart concocts animals spirits. Compelling evidence implies the extreme improbability that thinking, feeling, and experiencing are events in a nonphysical soul. Rather, they are events of the entirely physical brain. Because one party to the alleged interaction almost certainly does not exist, interactionism is a non-problem.

(*A neurophilosophical slant on consciousness research*, 2005, p. 285)

b) In secondo luogo il riduzionismo di Patricia Churchland si presenta come una tesi empirica basata sul riscontro di dati osservativi e falsificabile in linea di principio. Il riscontro di correlazioni sperimentali stabili tra stati mentali e stati fisiologici risulta essere una condizione irrinunciabile alla formulazione di una tesi riduzionista; né il riduzionismo, né l'anti-riduzionismo possono essere infatti concepiti, secondo la Churchland, come tesi valide a priori.

In saying that physicalism is an hypothesis, I mean to emphasize its status as an empirical matter. I do not assume that it is a question of conceptual analysis, a priori insight, or religious faith, though I appreciate that not all philosophers are at one with me on this point. [...] Whether science will finally succeed in reducing psychological phenomena to neurobiological phenomena is, needless to say, yet another empirical question. Adopting the reductionist strategy means trying to explain the macro levels (psychological properties) in terms of micro levels (neural network properties).

(*Can neurobiology teach us anything about consciousness?*, 1994, p. 23)

c) Infine un ruolo centrale è ricoperto dall'assunzione di criteri guida indispensabili alla formulazione del ragionamento abduittivo. Affinché una soluzione monistica, come quella proposta dalla Churchland, possa venir considerata la *miglior spiegazione* disponibile rispetto all'occorrenza di correlazioni sperimentali tra riferimenti a stati mentali e stati fisiologici, è necessario che si adottino, in via preliminare, *criteri di selezione* idonei a indicare *univocamente* un unico dominio esplicativo. Candidato a rivestire questa funzione può essere considerato il principio regolativo di *unità* della conoscenza. Accettare tale principio quale premessa dell'impresa conoscitiva significa infatti ammettere, in linea di principio, la possibilità che ogni conoscenza possa essere ricondotta a un unico e solo ambito disciplinare. Assumendo tale presupposto, non solo apparirebbe ragionevole accettare la tesi riduzionista, ma la stretta identificazione tra psicologia mentalistica e neuroscienze verrebbe a rappresentare la miglior soluzione esplicativa rispetto alle correlazioni osservative che contraddistinguono il rapporto mente-corpo.

The unity of science is advocated as a working hypothesis not for to sake of puritanical neatness or ideological hegemony or hold positivistic tub thumping, but because theoretical coherence is the "principal criterion of belief-worthiness for epistemic units of all sizes from sentences on up"

(*Neurophilosophy*, 1986, p. 376)

Una concezione pluralistica della spiegazione, la quale preveda ambiti diversi di comprensione dei fenomeni, o una concezione che ammetta più di un livello ontologico, non sembrano di fatto essere prese in considerazione dalla Churchland.

3. La coscienza

Accordingly, just it turn out that there was not such things as impetus, there may be no such things as awareness. (P. Churchland 1986)

L'aver messo da parte la teoria psicologica del senso comune, per mezzo di un progressivo processo di riduzione esplicativa, conduce la Churchland a porre in discussione anche le più tradizionali tassonomie concettuali riguardanti la nozione di coscienza. (Churchland, 2002, p. 129). L'indagine rivolta alla comprensione dei vissuti in prima persona diviene, all'interno del nuovo contesto riduzionista, un progetto a *lungo termine* volto a inserire la coscienza stessa nel dominio degli eventi naturali, indagabili per mezzo di procedure già impiegate con successo nella descrizione di altre proprietà dell'organismo umano.

In the long haul, of course, we want to understand consciousness at least as well as we understand reproduction or metabolism, but in the short haul it is wise to have realistic goals.

(Brain Wise, 2002, p. 134)

La ricerca di una soluzione naturalistica porta la Churchland a privilegiare un approccio sperimentale di tipo *indiretto*, all'interno del quale una teoria esplicativa delle correlazioni tra stati biologici e fenomeni coscienziali troverebbe spazio come stadio successivo rispetto allo sviluppo di una matura teoria delle funzioni cerebrali (Churchland, P., 2002, p. 156). Analogamente a quanto avvenuto per le teorie vitalistiche nel corso del XIX secolo, anche l'accantonamento della distinzione tra stati di coscienza e stati cerebrali attenderebbe, dunque, nient'altro che lo sviluppo di una conoscenza fisiologica adeguata, ma non ancora pienamente disponibile.

Argomenti volti a decretare l'impossibilità dei programmi di riduzione come quello della Churchland, basati per lo più su versioni del classico *argumentum ad ignorantiam*⁸ (Churchland, 2002, pp.174-176), così come sul ricorso a esperimenti mentali intesi a

⁸ Una versione canonica dell'*argumentum ad ignorantiam* è così definito dalla Churchland: "Science is largely ignorant about the nature of P, therefore we do know that P can never be explained, nothing science could ever discover could deepen our understanding of P, or P can never be explained in terms of properties of kind S" (Churchland, 2002, p. 174)

dimostrare la non identificabilità di stati fisici e stati della coscienza⁹, sono respinti dalla Churchland in forza del carattere preminentemente a posteriori del suo progetto (Churchland, 2002, p. 181). Del resto, il riduzionismo proposto dalla Churchland non mira a fondare la propria validità sul piano dell'analisi logica, quanto piuttosto sull'acquisizione di risultati sperimentali, con l'obiettivo di giungere alla formulazione di una teoria in grado di *spiegare* la correlazione esistente tra l'impiego di espressioni relative a stati qualitativi e specifici stati fisiologici del cervello.

In linea generale, nell'analisi di Patricia Churchland si osserva una riformulazione del classico problema della coscienza. Al centro dell'attenzione non si trova più la questione *a priori* concernente le misteriose possibilità d'interazione tra mente e materia, quanto piuttosto la questione *empirica* relativa a quali condizioni naturali permettano il realizzarsi delle diverse funzioni che contraddistinguono la comparsa di fenomeni come la coscienza umana.

L'approdo del percorso intrapreso dalla Churchland è rappresentato, per questo, dalla definizione di una teoria *scientifica della coscienza* che sia *predittiva, manipolativa e gerarchica*; ovvero in grado di anticipare il verificarsi di stati mentali muovendo dall'analisi dei corrispondenti stati cerebrali, di modificarne il decorso, e di descrivere correttamente i rapporti di riduzione che collegano ogni riferimento mentalistico al corrispondente stato cerebrale. Una teoria non solo in grado di fornire adeguati riscontri sperimentali, ma anche in grado di *spiegarne* la possibilità naturale attraverso la messa a punto di una concezione unificata delle proprietà mentali e corporee (Churchland, 2005).

4. La Neuroetica

Our moral nature is what it is because our brains are as they are [...] (P. Churchland, 2007)

Il superamento dell'impostazione tradizionale del problema mente-corpo permette una serie d'interessanti conseguenze in più di un settore della ricerca filosofica. Particolarmente interessanti sono le ricadute concettuali che la prospettiva riduzionista mostra di possedere rispetto a nozioni quali quella di responsabilità e di libero arbitrio, inserendosi per questo a pieno titolo nel contesto del dibattito etico.

The new research on the nature of ethics is located at the interface of philosophy, jurisprudence, and many sciences—neuroscience, evolutionary biology, molecular biology,

⁹ Ne è un esempio l'argomento per cui è sempre possibile immaginare che a due stati fisici identici possano essere associati rispettivamente stati della coscienza diversi. L'esperimento mentale dei "qualia invertiti" ne è una celebre formulazione (si veda N. Block, J. Fodor, 1972).

political science, anthropology, psychology, and ethology. These interdisciplinary inquiries will have profound, and rather unpredictable, social consequences, as people in general rethink their conventional ideas concerning the basis for moral standards and practices.

(*Neuroscience, ethics, agency, and the self*, 2005, p. 3)

L'esclusione di un ruolo epistemologico preminente riconosciuto a stati soggettivi della coscienza (§2), impone di fatto una riformulazione dei problemi tradizionali, particolarmente di quelli legati al contrasto tra analisi naturalistiche della mente e concezioni volte a salvaguardare lo spazio per la definizione di una libera possibilità di scelta basata su processi di introspezione (Churchland, P., 2002, p. 210). Consapevole del rischio rappresentato dalla classica fallacia naturalistica, la Churchland non intende proporre la derivazione di categorie morali dal riscontro di regolarità naturali, quanto piuttosto portare l'attenzione sul ruolo giocato dalle basi biologiche che sottostanno alla possibilità degli atti morali. Prendendo in considerazione i vincoli e le regolarità fisiologiche correlate alle decisioni che hanno a che fare con la sfera etica, la Churchland non esclude la presenza di altri fattori determinati come la storia personale di un soggetto, o la sua condizione sociale e culturale, concedendo rilievo soprattutto alla particolare costituzione naturale che contraddistingue i soggetti chiamati a compiere le scelte al centro dell'analisi etica (Churchland, P., 2005).

Le conoscenze neuro-biologiche mostrano con sempre maggior precisione come la possibilità di compiere giudizi e scelte comportamentali da parte di un soggetto risulti strettamente legata anche alla presenza di determinate circostanze neurofisiologiche. Tecniche d'indagine sperimentale mettono da tempo in evidenza il ruolo dei sistemi emozionali nella realizzazione di scelte e pratiche comportamentali; l'attivazione di più aree del sistema nervoso sembra dunque correlarsi strettamente alle capacità di giudizio e valutazione etica dimostrate da un soggetto.

Nel riconoscere il ruolo che gli aspetti fisiologici rivestono nell'impostazione di un problema tradizionalmente filosofico come quello del libero arbitrio, la Churchland non intende affermare a priori la presenza di un rigido determinismo comportamentale, quanto piuttosto la necessità di riformulare all'interno di un contesto naturalizzato le questioni riguardanti la possibilità di compiere scelte valoriali. Anche in questo caso non è il misterioso rapporto tra stati mentali e stati fisici a essere oggetto delle attenzioni della Churchland, concentrate invece a indagare le *condizioni naturali* che rendono possibile la selezione e la valutazione di un dato comportamento da parte di un soggetto.

La proposta della Churchland si concretizza nella messa a punto di un criterio di discriminazione tra scelte riconducibili ad atti volontari, posti *sotto il controllo* delle aree

cerebrali adibite alla valutazione e alla pianificazione comportamentale (in particolare tra queste le aree della corteccia pre-frontale, l'ipotalamo, l'amigdala, e il nucleus accumbens), e scelte che restano invece *estranee al controllo* di tali aree. Ciò permetterebbe lo sviluppo di criteri osservativi atti a consentire la distinzione tra comportamenti "normalmente controllati", dunque potenzialmente assoggettabili a responsabilità morali e comportamenti privi di tali caratteristiche. La nozione di libero arbitrio emergente da una simile prospettiva non rispecchierebbe, dunque, la possibilità da parte del soggetto di dar luogo ad atti privi di antecedente causale, quanto piuttosto la possibilità di dar luogo a un comportamento indotto dai naturali (non patologici e non deficitarii) meccanismi fisiologici correlati a stati intenzionali e volontari.

The important core of the idea of free will consists not in the notion of uncaused choice, whatever that might be, but in choices that are made deliberately, knowingly, and intentionally; where the agent is in control [...]. Because of developments in neuroscience and cognitive science, it is now possible to formulate a rough hypothesis concerning the neurobiology of 'in-control' brains, and the respects in which it differs from that of 'not-in-control' brains.

(*Neuroscience, ethics, agency, and the self*, 2005, p. 15)

Naturalizzare l'etica significa inoltre per la Churchland inserire parzialmente tale ambito all'interno della prospettiva evuzionistica, conferendole ulteriore sostegno esplicativo (Churchland, P., 2002, p. 254).

5. Verso una neuro-filosofia?

So it is that the brain investigates the brain, theorizing about what brains do when they theorize, finding out what brains do when they find out, and being changed forever by the knowledge. (P. Churchland, 1986)

Neurophilosophy si conclude con l'invito a liberarci dei paraocchi che ostacolano lo sviluppo di un atteggiamento in grado di tradurre i tradizionali problemi filosofici all'interno di un contesto naturalizzato. Non sorprende, dunque, che la Churchland si prodighi nella difesa di una forma di epistemologia all'interno della quale conoscere i processi alla base del funzionamento dei sistemi cognitivi rappresenta una condizione indispensabile per comprendere come i sistemi cognitivi stessi conoscono¹⁰:

¹⁰ Proponendo un'analisi delle condizioni biologiche correlate ai processi conoscitivi, è intenzione della Churchland porsi nel solco tracciato da W. V. Quine rispetto alla nozione di epistemologia naturalizzata. Si veda Churchland (1987).

As a bridge discipline, neuroepistemology is the study of how brains represent the world, how a brain's representational scheme can learn and what representations and information in nervous systems amount to anyhow.

(*Brain Wise*, 2002, p. 270)

Muovendo dalla definizione di epistemologia, fornita dalla stessa Churchland, quale "studio della natura della conoscenza" (Churchland, P., 2002, p. 241), appare però difficile comprendere come tale compito possa essere a sua volta ricondotto all'acquisizione di conoscenze specifiche, nella fattispecie biologiche. Se è pur vero che la consapevolezza delle condizioni naturali, che rendono possibile lo sviluppo di atteggiamenti conoscitivi, può gettare luce sulle modalità con cui la conoscenza viene normalmente acquisita e trasmessa, sembra indispensabile salvaguardare al contempo la legittimità di forme d'analisi il cui valore non appare direttamente riducibile alla definizione di specifici processi fisiologici. Oltre alla considerazione di condizioni al contorno, come il contesto culturale o sociale, entro cui si sviluppa il sapere, il compito dell'indagine epistemologica sembra infatti implicare anche l'analisi degli aspetti *qualitativi e fenomenologici* a partire dai quali si solleva e si struttura la conoscenza scientifica. Uno studio della natura della conoscenza, che non prenda in considerazione gli aspetti *immanenti* dell'esperire, rischia di escludere una componente fondamentale.

Risalendo a ritroso il percorso seguito in questo lavoro, una prima difficoltà presente nell'ambizioso progetto di naturalizzazione della filosofia riguarda la riduzione dei tradizionali problemi concernenti la sfera etica al campo di ricerca delle neuroscienze. Il tentativo di identificare la nozione di atto intenzionale-volontario, potenzialmente assoggettabile a responsabilità, con quella di stato fisiologico "sotto controllo" (§4) conduce infatti a una condizione di circolarità definitoria. Il riferimento a una particolare disposizione del sistema nervoso, intesa come stato fisiologico non patologico, correlato *di norma* ad atti intenzionali assoggettabili a responsabilità, richiede che sia possibile individuare quest'ultimi indipendentemente dai primi. Senza mettere in discussione il ruolo che le neuroscienze ricoprono nella diagnosi degli stati patologici e deficitari del sistema cognitivo, il tentativo di definire la volontarietà di un atto in funzione di correlazioni sperimentali con particolari stati cerebrali si riduce a una sorta di circolo vizioso. Come selezionare, infatti, gli stati fisiologici ai quali correlare atti volontari compiuti dal soggetto, se non attraverso una precedente definizione di cosa si debba intendere per atto volontario?

Il problema può apparire solo a prima vista risolvibile collocandosi all'interno di un'ottica eliminativista, entro la quale sia esclusa la possibilità di riferimento a teorie facenti ricorso

a stati mentali. La strategia esplicativa perseguita dalla Churchland prevede infatti che, al limite dello sviluppo delle conoscenze neuroscientifiche, sarà possibile definire una teoria completamente naturalizzata della mente e degli atti volontari. Radicalizzare una metodologia esplicativa, trasformandola in una condizione ontologica, presenta però limiti che la Churchland non sembra riconoscere. La definizione di correlazioni tra stati della coscienza e stati del cervello appare infatti non proseguire oltre il riscontro delle relazioni intercorrenti tra i resoconti dichiarati da un soggetto e certi specifici stati fisiologici, mentre il riscontro di correlazioni coinvolgenti stati qualitativi manterrebbe la tipica inaccessibilità in quanto esperibile solo in prima persona. Resta per tanto poco chiaro fino a che punto possa spingersi la tesi eliminativista. L'accantonamento di una teoria non fisiologica degli stati qualitativi non sembra infatti poter escludere a sua volta la presenza di una forma di *accesso* a certe condizioni fenomeniche esperite in prima persona. Ciò che è in discussione non è la scelta di un approccio metodologico che faccia ricorso o meno a riferimenti di tipo mentalistico, o qualitativo, quanto piuttosto la ripercussione sul dibattito ontologico che ciò comporta. Il ricorso all'impiego di soluzioni teoriche riduzioniste può certamente essere giustificato dagli scopi che si hanno di mira, senza essere associato necessariamente a un livello di analisi meta-teorico anch'esso riduzionista e monistico.

A ben vedere, è proprio l'impostazione esplicativa conferita dalla Churchland alla proposta riduzionista a limitarne la portata ontologica. Affinché possa affermarsi la validità di un ragionamento abducente, risulta infatti indispensabile riconoscere il ruolo di *assunzioni preliminari* che permettano la selezione della miglior ipotesi esplicativa disponibile. La scelta di criteri regolativi come quello di *unità* della conoscenza (§2) permette di sviluppare modelli monistici della spiegazione nei quali un solo livello descrittivo, nel caso della Churchland rappresentato dalle neuroscienze, viene selezionato per svolgere il ruolo esplicativo fondamentale. Come si può notare, la possibilità di variare i criteri di selezione in funzione dei problemi che s'intende risolvere, ad esempio abbandonando il principio di unità della conoscenza nel caso si necessiti di contesti di descrizione alternativi, permetterebbe di non limitare la scelta delle ipotesi esplicative a un solo dominio di riferimento. Il *monismo* riduzionista sembra così inquadrabile come *caso speciale* di un più ampio panorama, all'interno del quale trovano spazio forme descrittive plurali in grado di rispondere a necessità diverse. Al di là dell'esigenza di sviluppare concezioni del rapporto mente corpo *univoche e locali*, certamente efficaci ai fini della preziosa ricerca applicativa in campi diversi, come ad esempio quello diagnostico, o farmaceutico, possiedono una

loro legittimità anche forme di descrizione basate su esigenze di tipo *storico e narrativo*. Modelli teorici non riduzionisti potrebbero, infatti, fare al caso di richieste esplicative che mettano al primo posto l'esigenza di rendere conto di quegli aspetti immanenti dell'esperienza che, proprio la soluzione della Churchland, non sembra in grado di catturare.

Come si è visto, il riduzionismo della Churchland è inoltre basato sull'assegnazione di un carattere prettamente teorico alla psicologia del senso comune, condizione indispensabile questa affinché il linguaggio mentalistico in essa impiegato possa essere ricondotto entro l'alveo delle conoscenze neuroscientifiche. Oltre a dover riconoscere le difficoltà relative a un'esatta definizione di tutte le asserzioni che compongono una teoria per sua natura *vaga* come quella del senso comune, occorre considerare come alcuni recenti studi relativi ai processi cognitivi, coinvolti nel riconoscimento e nella comprensione dell'alterità, propendono verso soluzioni teoriche in grado di salvaguardare lo spazio per una dimensione *fenomenologica* della coscienza. Esempi di ricerche, provenienti proprio dalle neuroscienze, sembrano mettere in crisi il paradigma riduzionista basato sul carattere integralmente teorico della psicologia del senso comune. Il possesso di un sistema di neuroni specchio in certe aree del cervello, per esempio, determina *secondo alcune interpretazioni* la possibilità di delineare uno spazio d'azione condiviso, all'interno del quale ogni atto, proprio o altrui, apparirebbe immediatamente compreso senza che ciò richieda il ricorso ad alcuna operazione conoscitiva (Rizzolatti, Sinigaglia, 2006).

Ciò che rende particolarmente interessante tale dibattito è il fatto che, proprio a partire dal riscontro di dati sperimentali concernenti funzioni neurologiche, la prospettiva riduzionista incontra un possibile punto di arresto, avvalorando la richiesta di una diversa formulazione dei rapporti tra stati mentali e stati fisiologici.

Silvano Zipoli Caiani

Bibliografia

- Block, N., Fodor, J., What psychological states are not, *Philosophical review*, 1972, 81, pp. 159-181.
- Churchland, P. S., A Perspective on Mind-Brain Research, *Journal of Philosophy*, 1980, 77, 4, pp. 185-207.

- Churchland, P. S., Mind Brain Reduction - New Light from the Philosophy of Science, *Neuroscience*, 1982, 7, 5, pp. 1041-1047.
- Churchland, P. S., *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, 1986, MIT Press, Cambridge MA.
- Churchland, P. S., Epistemology in the Age of Neuroscience, *Journal of Philosophy*, 1987, 84, 10, pp. 544-553.
- Churchland, P. S., Is Neuroscience Relevant to Philosophy?, *Canadian Journal of Philosophy*, 1990, pp. 323-341.
- Churchland, P.S., Can Neurobiology Teach Us Anything about Consciousness?, *Proceedings and Addresses of the American Philosophical Association*, 1994, 67, 4, pp. 23-40.
- Churchland, P. S., *Brain-Wise: Studies in Neurophilosophy*, 2002, Bradford Book/MIT Press, Cambridge MA.
- Churchland, P. S. A neurophilosophical slant on consciousness research, *Progress in Brain Research*, 2005, 149, pp. 285-293.
- Churchland, P. S., Moral decision-making and the brain, in *Neuroethics: Defining the Issues in Theory, Practice and Policy* , edited by J. Illes., 2005, Oxford University Press, New York.
- Rizzolatti G., Sinigaglia, C., *So Quel che fai. Il cervello che agisce e i neuroni specchio*. Cortina, Milano, 2006.

Intervista **LUCIO LUZZATTO** di Matteo Borri

All'interno della manifestazione culturale Pianeta Galileo si è tenuto a Grosseto il 23 ottobre 2007 un incontro su La ricerca per alleviare il Cancro. Tra i relatori vi era il professor Lucio Luzzatto - oncologo di fama mondiale - che ha recentemente pubblicato il libro *Capire il cancro. Conoscerlo, curarlo, guarire* (Rizzoli, 2006). Al termine dell'incontro, il professor Lucio Luzzatto si è concesso per una breve intervista.

<http://www.consiglio.regione.toscana.it/news-ed-eventi/pianeta-galileo/pagine/default.asp?anno=2007&provincia=3>

1. Le cause del formarsi del cancro sono generalmente di tipo ereditario, o esiste anche una derivazione da fattori ambientali e sociali? Si può definire una relazione tra i due casi?

La domanda contiene in sé già una parte della risposta. È certo che giocano un ruolo sia fattori ambientali sia fattori ereditari. Sappiamo ad esempio che ci sono alcune famiglie che apparentemente pagano uno scotto troppo alto. In alcuni casi riusciamo addirittura a identificare il gene responsabile, in altri casi no; ma complessivamente ci sono almeno una cinquantina di geni già noti che influenzano fortemente il rischio di cancro, e se ne stanno cercando altri che influiscono più debolmente.

Per quanto riguarda i fattori ambientali, il fumo delle sigarette è pericoloso, è scritto su ogni pacchetto e questo significa qualcosa, cioè che l'effetto è molto grave, soprattutto per il cancro al polmone ma non soltanto. Altri fattori ambientali certamente ci sono. Purtroppo esistono degli inquinanti industriali e talvolta ambientali come il famoso amianto che causa il mesotelioma.

Oggi si parla molto di stili di vita; un fattore di rischio – non molto noto ma importante – è il soprappeso. Da questo risulta evidente che forse non è soltanto – come solitamente viene detto – il tipo di dieta ma anche la quantità della dieta. Se ci moderiamo è meglio: e ancora più utile è fare esercizio fisico per smaltire calorie.

2. Generalmente chi si occupa di comunicare – e a chi in prima istanza (parenti, diretto interessato...) – il riscontro di una malattia del genere? Esiste una figura ben definita che ha questo compito?

Chi se ne occupa, solitamente, è chi fa la diagnosi. All'atto pratico oggi l'indice di sospetto è alto, e viene praticata qualche indagine. Ci sono dei tumori (cutanei ad esempio) in cui un dermatologo può fare la diagnosi solo guardando. Nella maggior parte dei casi invece

occorrono o una procedura radiologica o una biopsia che portano alla diagnosi e allora di solito chi ha richiesto l'analisi è anche quello che poi dà la notizia al paziente.

E secondo me è bene che sia così: il paziente si aspetta una risposta dalla persona alla quale è andato a porre la domanda. Nello stesso tempo ci vorrebbe però anche un sostegno psicologico quando si apprende una notizia del genere.

3. Esiste un criterio standard dal punto di vista della divulgazione in ambiente extra-scientifico?

Se devo rispondere concisamente (da anglosassone) la risposta è no. Forse però è meglio rispondere in modo più articolato. Penso che sia un po' difficile standardizzare la divulgazione. Mi domanderei piuttosto quanto gli scienziati e quanto gli esperti di media in generale partecipino alla divulgazione e alla percezione sociale della malattia.

Quando ero giovane gli scienziati non comparivano quasi mai sui media. Prendiamo ad esempio il 1953, quando Watson e Crick hanno scoperto il modello del DNA: nei media non apparve niente. Cinquanta anni dopo, per l'anniversario della scoperta, non si parlava d'altro.

E ora il Signor Venter fa notizia perché ha annunciato un esperimento che è sì molto interessante ma nemmeno lontanamente paragonabile all'evento di cui ho appena parlato.

Non c'è standard dunque. Personalmente sono dell'idea che gli scienziati non siano affatto obbligati a fare divulgazione scientifica ma quelli che lo fanno svolgono una funzione utile. È anche per questo che ho scritto il libro *Capire il cancro*: per comunicare con il pubblico, appunto.

4. Esiste un criterio standard relativo alla comunicazione ed al rapporto con il paziente?

Ho una convinzione molto radicata, sebbene non possa dire che sia condivisa da tutti.

Ritengo che faccia parte del lavoro del medico non solo fare la diagnosi ma anche comunicarla in modo corretto al paziente. Che poi a fare la comunicazione sia il medico che per primo ha visto il paziente, o il team insieme, o un tutor designato ad hoc, è da discutere: ma qualcuno lo deve fare. Secondo me va fatto senza dire bugie. Uso volutamente questo termine, che può sembrare provocatorio, perché purtroppo succede ancora – in forma larvata e per la mia esperienza specialmente in Italia – di leggere referti in cui appaiono termini criptici come per esempio “lesione di presumibile natura

cariocinetica”. La cariocinesi è la divisione cellulare; questo dunque dovrebbe essere un termine eufemistico per dire “una lesione probabilmente dovuta a cellule che si dividono”, in altre parole una lesione probabilmente tumorale. A chi si vuole nascondere la corretta comunicazione? Probabilmente non al collega ma al paziente.

Sinceramente sono molto perplesso di fronte ad una possibilità del genere, per due motivi. In primo luogo perché nascondendo qualcosa è come se volessi saperne di più del paziente, che però è il soggetto della malattia. Si tratta del suo corpo, non del mio: dunque io non penso di avere questo diritto.

Il secondo motivo riveste una importanza pratica: dovranno essere fatte delle scelte terapeutiche. A mio modo di vedere, tali scelte devono essere concordate tra il medico (o l’equipe) e il paziente, soprattutto nel campo dei tumori; spesso c’è più di una opzione. Se il paziente non è pienamente informato sulla situazione, è chiaro che non può partecipare a queste scelte o – peggio – non può farlo in modo ottimale dal suo punto di vista. Spesso si tratta di scelte molto personali. Di fronte ad un tumore avanzato, ad esempio, un paziente potrebbe voler lottare per qualche settimana in più anche a prezzo di tossicità cioè di interventi “pesanti”; in condizioni simili un altro paziente potrebbe chiedere invece solo di essere lasciato in pace. A me sembra chiaro che di questi due atteggiamenti non ce ne è uno giusto e uno sbagliato: si tratta di due scelte diverse, ugualmente rispettabili. Noi dobbiamo creare il terreno sul quale la scelta possa essere fatta, e a me sembra che la *full disclosure* sia un presupposto. È un compito non facile, che mette a prova non solo la nostra professionalità, ma anche la nostra umanità.