

# FILOSOFIA DELLA MENTE

## Lineamenti storici fondamentali

- I. *La filosofia della mente. Premesse epistemologiche*
- II. *Sviluppo storico*
  1. *La concezione classica*
  2. *Cartesio e i moderni*
  3. *Le indagini sperimentali moderne e il loro ruolo sul problema mente-corpo*
  4. *Il comportamentismo*
  5. *Teoria dell'identità (mente=cervello)*
  6. *Cognitivismo e scienze cognitive*
  7. *Il funzionalismo computazionale (mente=computer)*
  8. *Il funzionalismo rappresentazionale*
  9. *Altri tipi di funzionalismo*
  10. *Altre posizioni*
  11. *Connessionismo*
  12. *Approcci neurofisiologici recenti*

### I. LA FILOSOFIA DELLA MENTE. PREMESSE EPISTEMOLOGICHE

*Scopo.* Intendo presentare in questo corso da una prospettiva tomistica alcuni elementi della problematica oggi dibattuta sotto il nome di *filosofia della mente*. Quest'approccio è insolito, se facciamo eccezione dei lavori di G. Basti, dal momento che la filosofia della mente è un campo relativamente nuovo, circoscritto originariamente all'area del pensiero filosofico anglosassone, nel quale non si è mai fatto un confronto con le dottrine classiche. Il motivo è che il tipo di filosofia elaborato in quell'area eredita la consuetudine dell'approccio analitico di compiere analisi senza una prospettiva storica. Le loro proposte filosofiche sono presentate in modo non-storico, seguendo il modello scientifico, anzi in vicinanza alle ricerche delle scienze ritenute rilevanti.

Ma neanche io elaborerò i temi che mi accingo ad esporre come una sorta di confronto storiografico con il pensiero di Tommaso d'Aquino. Il mio metodo sarà alquanto diverso. D'una parte, bisogna esporre le tematiche della *Philosophy of Mind* con un certo approccio storico interno, in quanto tale filosofia, pur nei rapporti interdisciplinari, gode già di una tradizione propria di cui i suoi esponenti diventano sempre più consapevoli. D'altra parte, pretendo reinterpretare gli elementi offerti dalla FM -filosofia della mente- alla luce di principi o concetti appartenenti alla visione

tomistica della conoscenza e dell'uomo, ma non in un modo storico, bensì sotto la mia responsabilità e quindi con una certa libertà. Perché questo metodo?

I motivi sono parecchi e mi stupisce che un simile lavoro non sia ancora stato fatto in un modo massiccio. Il tomismo, a causa del suo sfondo aristotelico, di per sé è una filosofia aperta allo sviluppo delle scienze naturali e, nel caso della conoscenza, alla psicologia e alla biologia. Il programma di ricerca di una filosofia ispirata a Tommaso d'Aquino non può evitare di dare importanza alle scienze, purché queste ultime ricevano almeno un minimo d'interpretazione realistica. Ora, la FM si nutre costantemente degli studi di linguistica, psicologia cognitiva, neuroscienze e scienze computazionali. *La FM è essenzialmente interdisciplinare* e in qualche modo sta sempre confrontandosi con le cosiddette scienze cognitive, anzi pretende di essere la loro interpretazione filosofica.

Inoltre, *la FM ha un solo avversario filosofico*, col quale sta sempre in situazione non direi di vera polemica, ma di esclusione parecchie volte riconfermata: *il dualismo cartesiano* (forse anche la fenomenologia in senso stretto, anche se la FM cerca di reinterpretarla). Non si polemizza veramente contro il cartesianesimo perché non è questa una dottrina viva, considerata come una possibile alternativa. Piuttosto essa viene considerata come quell'elemento negativo la cui invalidità trova tutti sempre consenzienti, al di là delle polemiche. Questo fatto obbedisce a un motivo di fondo: nella stragrande maggioranza, gli esponenti della FM sostengono una visione materialista dell'uomo e della conoscenza, detta anche "naturalista", con diversi gradi di radicalità secondo i casi (forse la "scuola australiana" di FM è la più impegnata nei dogmi materialisti). Tale tesi non risponde quasi mai ad un'analisi approfondita. Piuttosto è la premessa iniziale. Il nome stesso di *materialismo* è usato in un modo normale dai suoi sostenitori, convinti che soltanto una filosofia materialistica può essere in consonanza con le scienze naturali, dal momento che postulare un elemento superiore alla materia sarebbe inaccessibile alle scienze e quindi non potrebbe avere una reale esistenza ontologica.

*Da questo punto si evince che la FM appare come una filosofia dura.* Ed effettivamente lo è. "Dura" suggerisce nel mio contesto una filosofia che persegue le sue tesi entro una visione molto ristretta, con una metodologia vicina alle scienze e con un supporto apparente nelle ricerche scientifiche. In qualche modo, tale filosofia è la continuazione dell'empirismo logico, in un quadro però nuovo, fondamentalmente psicologico e con tendenza a risolversi nelle neuroscienze. Questa speciale "durezza" rende la FM piuttosto estranea agli indirizzi "continentali", come l'ermeneutica o altri richiami di tipo umanistico-storico. La FM tende piuttosto a presentare una visione poco

umanistica, il cui baricentro è la riduzione-dissoluzione dell'io o della persona nel cervello o nelle macchine intelligenti. All'interno della FM non mancano posizioni contrarie a queste tendenze, che comunque si sono limitate ad alimentare le polemiche, non prive di interesse, sempre però nell'orizzonte ristretto caratteristico della FM, come avremo occasione di vedere.

*Le problematiche abituali dei dibattiti della FM*, infatti, sono di solito: è la mente umana simile a un computer? Possono pensare le macchine? È la mente umana niente altro che il suo cervello e le sue funzioni? È il cervello umano una sorta di computer? Supererà l'intelligenza artificiale le nostre intelligenze naturali? I non riduzionisti rispondono negativamente, ma l'orizzonte di comprensione della FM è sempre questo. Può darsi che le discussioni al riguardo ormai siano arrivate a un punto di stagnazione e che forse la FM, almeno così come è stata concepita, si trovi in uno stadio finale, allo stesso modo che la stagione dell'empirismo logico a un certo punto si fermò. Prevedo, comunque, che essa sarà riattivata ogni volta che ci saranno nuovi progressi nel campo neurologico ed informatico, progressi che certamente arriveranno e che ci stupiranno.

A questo punto, *l'intervento in questa tematica di un'impostazione di stampo aristotelico e tomistico*, senza preoccupazioni storiografiche fuori posto nel nostro contesto, *può risultare opportuno*, dal momento che un filosofo che condivide certe tesi di fondo tomistiche è in grado di comprendere l'importanza delle acquisizioni delle scienze cognitive e anche di assimilarle alla propria concezione. Non pretendo però di stabilire una polemica con gli autori delle FM, nei quali non è prevedibile un'apertura mentale verso orizzonti personalistici e spiritualistici. Di conseguenza, non ho la pretesa di convincere alcun autore impegnato con una filosofia materialistica. Del resto, la componente ideologica in queste posizioni di fondo normalmente sta lì, occulta ma ferma. Io sono di coloro che credono che le ultime posizioni filosofiche e morali spesso sono sostenute *perché si è decisi* di sostenerle e basta. Il resto si organizza abilmente in funzione di tale scelta. Non sempre è così, e non dovrebbe esserlo in teoria, ma purtroppo l'uomo di carne ed ossa si mostra più condizionato dai pregiudizi da quanto egli crede, pregiudizi che spesso sono più sociologici che personali. La mia esposizione sarà, tuttavia, aperta alle scoperte delle scienze cognitive, anche se tenterà di mostrare che *l'interpretazione materialistica non è necessaria*.

Come farlo? Innanzi tutto, cercando di eliminare un equivoco primario: *il naturalismo come "a priori"*. Per i sostenitori del materialismo -un materialismo che non va certo addebitato a tutti gli autori della FM-, le tesi fondamentali della filosofia debbono essere accettabili dalla scienza fisica (biologia, neurologia e altre simili). Questa restrizione dei contenuti filosofici "seri" ("scientifici")

al campo semantico della fisica decide *a priori* la sorte della filosofia. Siamo davanti a una presa di posizione che si afferma come una suprema tautologia. In parole più semplici -anche la FM spesso vuole essere molto semplice nelle sue affermazioni-, per gli autori indicati, l'unica filosofia accettabile per la scienza fisica, che si occupa soltanto della materia, sarà una filosofia materialistica, e l'unica fonte della conoscenza per una tale filosofia sarà la scienza fisica. Ecco un circolo chiuso. La scienza naturale viene presa in questo senso in una maniera chiusa, nel senso che si suppone copra tutto l'orizzonte di ciò che è. Un discorso che parli di cose fisicamente non controllabili o non comprensibili con categorie fisiche non potrebbe pretendere di avere un'attendibilità filosofica al livello della scienza. Potrà essere considerato, al massimo, come un tipo di discorso popolare, più o meno utile nella vita di tutti i giorni, ma non come un punto di partenza per un'autentica ontologia. Le categorie ontologiche saranno semmai una generalizzazione del tipo di categorie usate nelle scienze naturali.

*Scienza, filosofia, senso comune.* Arriviamo così al punto metodologico fondamentale. Questo corso è di filosofia del pensiero, un approccio diverso da quello della psicologia empirica e delle scienze computazionali. Questo scopo presuppone una distinzione tra la filosofia e le scienze e una certa concezione dei loro rapporti vicendevoli.

La *scienza* studia la struttura e i vincoli causali generali tra *aspetti* determinati di una realtà. Nella misura in cui ci sono molti possibili aspetti nelle cose, vi sono in corrispondenza molte possibili scienze. Gli aspetti si presentano come tali solo quando l'uomo riesce a focalizzare un ambito della realtà in base a una certa astrazione o "punto di vista" analitico, e la scienza nasce quando tale focalizzazione consente un discorso sistematico di tipo razionale. La *filosofia*, invece, intende conoscere *ciò che è più essenziale* delle cose. Il suo approccio, quindi, tende a considerare la totalità nei suoi principi ultimi. Per la prospettiva filosofica, il compito delle scienze è parziale e incompleto. Ma se la scienza crede di afferrare qualcosa di totale e di definitivo nelle sue analisi, allora si pone come filosofia. Risulta "definitivo", ad esempio, pensare che le scienze ci parlano su ciò che è più importante e più vero sul mondo, quindi tale posizione non appartiene ad alcuna scienza in particolare, ma è un tesi di filosofia (in questo caso, filosofia della scienza).

La "*visione personale*". Lo sguardo verso la totalità nei suoi aspetti essenziali e nei suoi principi ultimi non è facile. Spesso passa inosservato, come un presupposto implicito di quanto facciamo o diciamo. Possiamo averlo riguardo a molte cose in un modo naturale, con certi limiti, ma spesso vi siamo portati, in un modo inconsapevole, tramite le valutazioni sociali, istituzionali o religiose. In questo senso, ad esempio, consideriamo che vale la pena sacrificarsi per gli amici o che

non dobbiamo uccidere le altre persone. Le filosofie elaborate perseguono tale tipo di comprensione con metodi di riflessione sistematica. A questo punto emerge la rilevanza della nostra *comprensione personale* delle cose. Siamo noi personalmente a dover fare la valutazione essenziale delle cose, utilizzando le nostre risorse intellettuali. Se lasciamo che altri più sapienti svolgano per noi questo compito, con affidamento di fede, ciò potrà essere una buona scelta in certi casi, ma non sarà un atteggiamento filosofico.

Quindi sembra che, per la filosofia, “l’ultima istanza” del giudizio essenziale e valutativo siamo noi stessi o, in altre parole, la nostra intelligenza o ragione. Questa posizione non comporta una chiusura egoistica nelle nostre idee. Possiamo -anzi dobbiamo- tener conto del pensiero altrui, ma alla fine siamo noi, con la nostra responsabilità, a dover giudicare. In qualche modo non possiamo evitare di farlo implicitamente, poiché le nostre scelte e la stessa vita ci costringono in questa linea. In un certo senso non si può evitare di essere “filosofi”. Non possiamo non prendere un atteggiamento di fronte alla realtà globale (solo possiamo essere indifferenti o “neutrali” nei confronti di cose particolari, che si possono ignorare o evitare).

*La critica contro il senso comune (eliminativismo).* Perché sto dicendo queste cose nel corso che mi accingo a esporre? Perché una delle posizioni materialistiche sopra accennate affronta la conoscenza umana squalificando sin dal principio la nostra capacità personale di comprendere noi stessi. Tale capacità è considerata come un modo di pensare non scientifico e quindi non affidabile. Di solito è chiamata *visione del senso comune* o *popolare* (conoscenza volgare, prescientifica). Così come con le sole risorse del “senso comune” (conoscenza ordinaria, fenomenologica) tendiamo a pensare che la terra è immobile, e che solo la fisica, con i suoi metodi, ci fa scoprire che in realtà essa si muove, analogamente, argomentano i riduzionisti, non sono attendibili le nostre convinzioni personali ad esempio sul fatto che siamo coscienti, che soffriamo o amiamo. Le nostre idee ordinarie su ciò che siamo vengono denominate, in questo senso, *psicologia popolare*. La psicologia popolare sarebbe come la fisica o la chimica popolare: una serie di idee non scientifiche, prive di valore, o forse un tipo di teoria che posteriormente potrebbe rivelarsi falsa, così come una persona seguendo la sua “teoria fisica popolare”, cioè attendendo alle presentazioni antropomorfe della realtà, potrebbe ritenere che “il Sole si alza ogni mattina”.

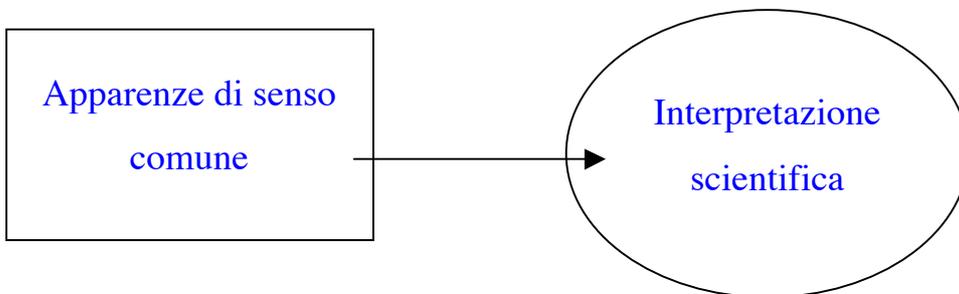
*Le “ermeneutiche forti”.* Con queste premesse, naturalmente, le nostre convinzioni sul fatto che siamo liberi, che pensiamo o amiamo, o le nostre idee morali, non contano niente. Esse saranno soltanto un primo materiale fenomenologico da interpretare ulteriormente secondo i canoni scientifici. Siamo davanti a una struttura epistemologica particolare. La chiamerei *ermeneutica in*

*sensu forte*. Secondo tale struttura, le convinzioni di senso comune, ritenute solo fenomenologiche - nel senso di meramente fenomeniche, superficiali, apparenti e antropomorfe - vanno interpretate da un'istanza considerata superiore. Tale istanza non solo le approfondisce, ma può arrivare anche a negarle, e in qualsiasi caso le reinterpreterà in maniera "profonda". Potrebbe risultare, ad esempio, che in realtà io *non sono libero*, grazie a tale interpretazione, benché io abbia la pseudo-convinzione di essere libero.

La struttura ermeneutica forte è stata spesso rivendicata dai filosofi. L'idealismo, ad esempio, talvolta ha ritenuto che le convinzioni degli uomini volgari sono pensieri pieni di ombre, immaginazioni, opinioni, per cui la verità certa e necessaria si mostrerebbe soltanto al sapiente filosofo e a chi lo segue. Il movimento, secondo questa linea, potrebbe essere un'illusione della gente. Solo al pensatore profondo apparirebbe la verità dell'immutabilità di tutte le cose (così si può giustificare, ad esempio, una forma di parmenidismo). Una peculiarità di questa impostazione epistemologica è che essa lascia qualsiasi persona senza la possibilità di argomentare, dal momento che qualsivoglia replica della persona contro tale teoria verrà automaticamente interpretata dal teorico "superiore". Il filosofo, lo scienziato superiore, non può dialogare con il volgo. Lo può soltanto interpretare, o può insegnare l'incolto ad elevarsi al suo punto di vista superiore ("superiore" qui significa la posizione di un'istanza che si ritiene in grado di giudicare e di non essere giudicata da altre istanze che allora sono "inferiori").

Spesso si ritiene che il freudismo, il marxismo o altre "filosofie del sospetto" abbiano una struttura ermeneutica forte, nella misura in cui queste dottrine interpretano le manifestazioni di superficie del linguaggio, la condotta e la vita sociale secondo canoni profondi di cui i soggetti sottoposti non sono consapevoli (qui "profondo" significa nascosto, a titolo di struttura di fondo che si ritiene spieghi ciò che sta invece in superficie). Le ermeneutiche forti, come si vede, mantengono un atteggiamento critico nei confronti della coscienza, e quindi sono costrette ad insistere sul ruolo degli elementi "profondi e inconsci", a partire dai quali sono giudicati i fenomeni più superficiali della coscienza fenomenica. Di conseguenza, nessun argomento elaborato in base a ciò che può apparire alla coscienza risulta valido per tale atteggiamento. È meno noto che questa stessa struttura ermeneutica forte viene ripresa dalle filosofie "profonde" della mente, ad esempio neurologiche. La negazione di attendibilità delle proprie esperienze coscienti sembra un'eredità del comportamentismo. Ma ora, quando si dispone di una teoria esplicativa profonda delle nostre esperienze (mentre il comportamentismo voleva rimanere a livello descrittivo), siamo di nuovo caduti nell'interpretazionismo radicale.

*Intendo rilevare il carattere contraddittorio di questo metodo.* Questo punto si potrebbe evidenziare in mille modi. “Io adesso provo fame”, ad esempio, è un’affermazione di cui sono cosciente. Se questa affermazione ha un senso per me e per altri, e può essere vera, è perché io e gli altri capiamo il significato dei termini in essa impiegati, cioè “io”, “provare o sentire fame”. Il fatto di capirlo non significa avere a disposizione una teoria filosofica o scientifica sull’io o sulla sensazione di fame. Tale teoria potrebbe *spiegare* in un modo veramente più profondo, anche causale, il fenomeno della fame e il suo rapporto con l’io. Ma una teoria riduttivista, nel senso prima menzionato della riduzione alla fisica, *elimina* il valore della stessa descrizione, in quanto essa non sarebbe scientificamente attendibile. Si potrà arrivare a dire, ad esempio, che l’io è un’illusione, e lo stesso vale per la fame. Poi, tali apparenze verrebbero *ridotte* ad altre realtà, in base alla teoria materialistica adoperata. Tale riduzione intende collocarsi al di fuori di ogni possibile argomento basato sulle mie pretese evidenze, le quali non hanno alcun valore, perché potrebbero essere mere apparenze. La teoria giudicante, allora, si è auto-collocata come inconfutabile, in quanto non può essere giudicata da nessun’altra istanza superiore alla teoria stessa. Tale teoria giudicante ammette soltanto argomenti interni ad essa. Il circolo si è chiuso. Eppure, non si evita la contraddizione, poiché la teoria materialistica giudicante viene giustificata e proposta sul piano delle apparenze antropomorfe, per esempio quando si pretende di dimostrare che “l’io non è che un’illusione”, una frase enunciata con la pretesa di essere vera e fondata.

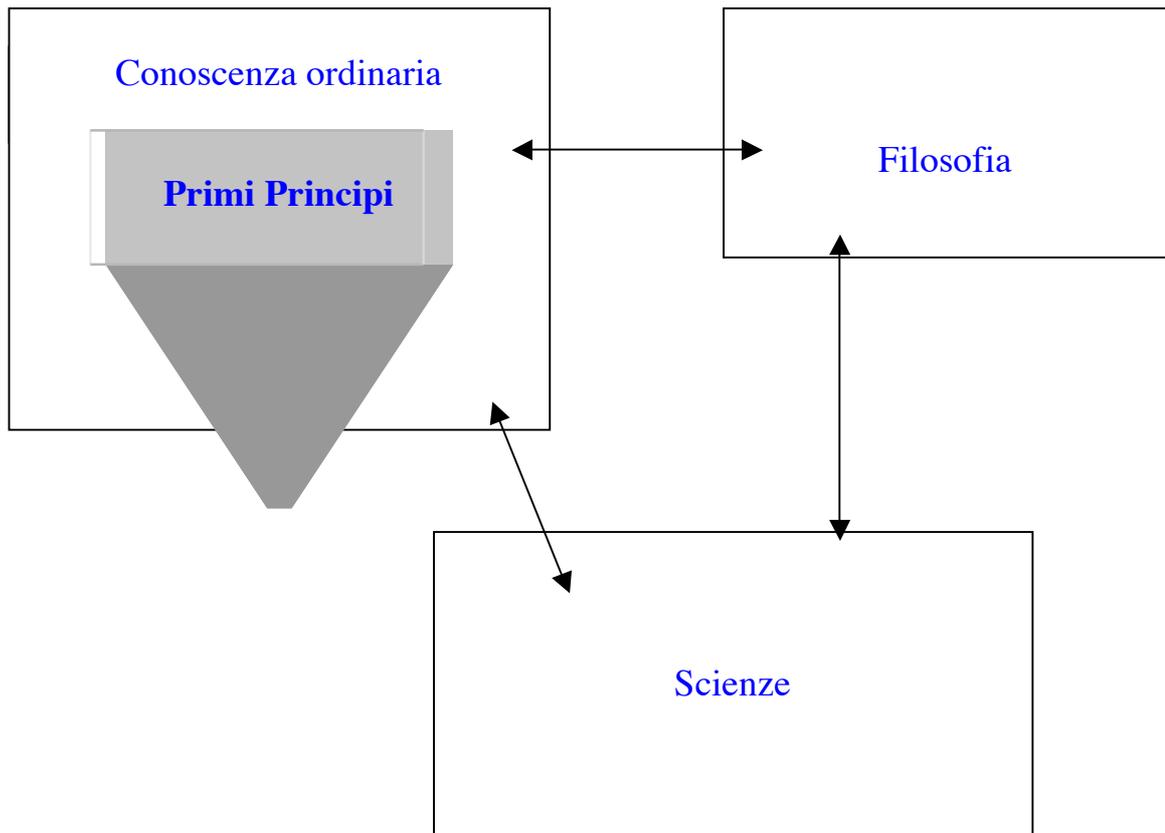


Secondo questo schema, le apparenze di senso comune sono interpretate da una teoria scientifica -nel senso di scienza che abbiamo definito prima, cioè una visione sistematica e causale di un aspetto della realtà-, ma tale teoria non può essere confutata dal senso comune, eppure la riduzione viene fatta con le argomentazioni e il linguaggio del senso comune (in questo senso, contraddittoriamente, una persona dovrebbe arrivare a credere che “lei non è niente”, contro tutte le sue semplici “apparenze”).

*Valore ineliminabile delle evidenze primarie.* Contro questa ermeneutica forte, occorre un’epistemologia che riconosca un valore assoluto a certe evidenze primarie, invece di assegnarlo a una teoria che di per sé non è evidente, e che al limite, se deve essere accettata in qualche senso, lo

sarà sempre in base ad argomentazioni che in ultima istanza procedono dalle evidenze primarie. Il campo delle evidenze primarie in fondo non è altro che la nostra capacità intellettuale personale. Infatti, se il riduttivista è convinto della potenza di spiegazione della sua teoria riduttiva, è perché lui *vede* (o *crede di vedere*) che è così. Semplificherei questo punto con la semplice frase: *per vedere, non abbiamo altro che i nostri occhi*. Di conseguenza, se qualcuno, quando vedo un albero, mi dice che in realtà io sto vedendo “inconsapevolmente” un elefante, non gli crederò. O se per un attimo di debolezza gli credo, la mia assurda credenza non durerà a lungo. Non è possibile opporsi a lungo andare alle nostre evidenze primarie. Queste evidenze giocano un ruolo correttivo e neutralizzante di tutte le teorie assurde che ci possono essere, sia scientifiche che filosofiche. Se mi è consentito di esprimermi così, esprimerei questo fatto con la frase: *l'uomo non è stupido*. Nessuna teoria lo convince fino in fondo, se tale teoria si oppone alle sue convinzioni primarie, basate su certe evidenze fondamentali.

È vero che non tutto ciò che ci risulta apparente è reale. La nostra esperienza stessa ci attesta che non qualsiasi apparenza soggettiva è vera. Le interpretazioni, quindi, sono necessarie, dal momento che molte apparenze, conoscenze incomplete e discorsi di ogni tipo contengono delle ambiguità e hanno bisogno di una chiarificazione, riguardo alla quale le nostre risorse intellettuali ordinarie possono essere insufficienti. La scienza e la filosofia sono necessarie. Ma anche in questo caso il rimando alla visione intellettuale personale è ineliminabile. Le spiegazioni filosofiche e scientifiche ci convincono nella misura in cui sono collegate a certi elementi di base delle nostre evidenze primarie assolute (ad esempio, evidenza sull'invalidità della contraddizione, sull'esistenza del mondo, sul valore costringente delle prove, sul senso di verità delle proposizioni).



*I primi principi.* Lo schema segnala tre elementi fondamentali che configurano la nostra conoscenza. D'una parte abbiamo la conoscenza ordinaria, nella quale, insieme a numerose e svariate notizie sulla realtà, acquisite tramite l'esperienza, l'insegnamento, le tradizioni sociali, esiste un nucleo forte di conoscenze primarie di carattere ontologico. Tale nucleo è ciò che Aristotele denominava *primi principi*, ma in un senso più ampio. Non soltanto includiamo in questo nucleo gli assiomi metafisici fondamentali, come il principio di non contraddizione e certe regole e nozioni morali di base -ad esempio, la distinzione tra giustizia ed ingiustizia, la nozione di menzogna-, ma anche alcune conoscenze esistenziali da cui non si può prescindere, come ad esempio la percezione dell'esistenza del mondo, di noi stessi e delle altre persone, o il fatto di sapere che siamo razionali e che parliamo un linguaggio. Queste conoscenze sono naturali, comuni a ogni uomo, implicite e incrollabili finché si mantiene l'uso ordinario della ragione. In quanto esse sono ciò che l'uomo comprende e vede continuamente, in modo preverbale, il loro raggio d'azione costituisce l'*intellectus* o il *noûs* di cui parlavano Aristotele e Tommaso. Non sono propriamente intuizioni, ma piuttosto conoscenze preve ad ogni concettualizzazione, anche se lo sviluppo del linguaggio può renderle oggettive, in forma concettuale, fino a un certo punto.

La nostra tesi sul nucleo forte dei primi principi non è gratuita e non nasce da un pregiudizio filosofico. Direi che esso si fonda sull'esperienza che abbiamo di noi stessi, e che si può argomentare indirettamente. Come riesco a sapere, ad esempio, che io sono diverso dagli altri e dal mondo (autopercezione della mia identità)? *Notiamo le seguenti caratteristiche* (si potrebbe fare qualsiasi altro esempio di natura ontologica o antropologica primaria):

- so con certezza che io veramente non sono le altre cose che percepisco, senza bisogno di avere una teoria filosofica o scientifica elaborata su che cosa significa in profondità *io, cosa, mondo, essere, verità, distinzione*;

- normalmente non dispongo di queste conoscenze come di un oggetto "che si sa". Esse sono piuttosto il presupposto di tutto ciò che faccio o che capisco, e anche della vita sociale. Io sono sempre nel mondo, intervengo in esso con azioni, conosco e interagisco con altre cose, converso con altre persone. Ma non saprei spiegare bene, se sono domandato, che cosa sono io, la verità, il mondo, o l'esistenza di una persona nel mondo. Queste sono appunto le tematiche su cui i filosofi indagano a non finire;

- non posso perdere queste conoscenze abituali e preoggettive in base a nessun'altra conoscenza. Anche se imparo teorie filosofiche che sostengano, ad esempio, l'inesistenza del mondo o di me stesso, non riuscirò a credere in tali teorie in un senso autentico e vitale, ma solo accademico;

- se qualcuno perde il senso di tali conoscenze, consideriamo che ha perso l'uso della ragione e quindi è afflitto da una anomalia psichica (delirio, allucinazione, ecc.);

- le conoscenze primarie sono argomentabili indirettamente, poiché qualsiasi attività dell'uomo, come le scienze, la filosofia, la religione, la politica, le presuppone. D'altra parte, nessuna scienza parte mai dal nulla, ma presuppone certe conoscenze primarie. Quindi non è possibile una scienza, nemmeno una filosofia, che auto-chiarifichi completamente la conoscenza umana. Le scienze si limitano a indagare entro un quadro dato di presupposti;

- le conoscenze ontologiche di base sono conoscenze oscure in quanto al senso, eppure sono indubitabili nel loro valore di verità. Nessuna persona può dubitare seriamente se esiste o se è già morta, ma ciò non significa che abbia una comprensione profonda dell'esistenza o della morte. Uno dei compiti della filosofia è il chiarimento del senso dei primi principi. Delle filosofie si può dubitare, non dei primi principi.

*Le frecce dello schema presentato indicano i rapporti interdipendenti tra la conoscenza comune, la filosofia e la scienze.* Ogni livello può arricchire gli altri da un certo punto di vista. La conoscenza filosofica rimane libera dalle restrizioni metodologiche definitorie delle scienze particolari e cerca di arrivare a una visione essenziale ed integrata dell'insieme della realtà, o almeno di una sua sezione abbastanza ampia, come può essere la natura, l'uomo, la società (nascono così le filosofie della natura, dell'uomo, della storia, e altre simili). *Esiste sempre la tentazione di prendere un punto di vista particolare e di elevarlo a criterio interpretativo ultimo o essenziale* (pur non impiegando questi termini). La storia del pensiero registra, in questo senso, il caso delle filosofie "riduttivistiche", come il logicismo, psicologismo, fisicismo e altre simili. Le filosofie riduttivistiche sono però contraddittorie, poiché il tentativo di trasformare una visione scientifica in una visione completa e totale non si può giustificare con i canoni scientifici di tale ramo del sapere. La tesi secondo cui la fisica è il punto di vista esclusivo non è giustificabile tramite la fisica. Tale tesi si sostiene soltanto a partire da un livello di pensiero più alto, meta-scientifico, livello che a buon diritto bisogna chiamare *filosofico*. Lo scienziato particolare, quindi, è costretto a riconoscere la parzialità del suo approccio, anche se egli, se non è abituato a considerare con serietà altri ambiti di pensiero, facilmente prova la tentazione di credere che la sua scienza è l'impostazione fondamentale o totale (si tratta di una vera "tentazione filosofica", molto naturale nell'uomo).

*Nozione di conoscenza comune.* Ritengo necessario riconoscere una certa autonomia a quello che denomino "conoscenza comune", pur ammettendo che non ci sono limiti precisi tra essa e la filosofia o le scienze. Il nome "comune" indica una certa universalità. La conoscenza *comune* -si potrebbe anche chiamare *personale*- è condivisa più o meno da tutti, almeno in una certa area culturale dove le persone sono comunicate. Tale tipo di conoscenza sarebbe, quindi, il sapere che possiamo trovare in ogni persona di una cultura. Ovviamente gli individui non hanno tutti lo stesso livello di educazione, informazione, esperienza e sapere. La "conoscenza comune" è variabile e dinamica. Si nutre da tutto, cioè dalle scienze cui si accede attraverso lo studio, dalla divulgazione, dalle semplici letture, dalle notizie e dati procedenti di Internet o dai mezzi di comunicazione. Si nutre pure da idee filosofiche acquistate in letture, conversazioni, conferenze, studi accademici, incontri. Tale conoscenza comune ha strati stabili e fermi e altri in evoluzione, secondo le attività intellettive delle persone. Può diventare più specialistica in qualche senso, ad esempio, perché si acquista molta esperienza in un certo campo pratico.

Il punto interessante è che tale conoscenza ordinaria è un luogo di *integrazione concreta* dei diversi saperi astratti o di altre specializzazioni. Un fisico può sapere ed esercitare la scienza fisica quando entra nell'ambito della sua rigorosa metodologia. Nella misura in cui lo fa, egli è situato in

una sfera astratta del pensiero, necessaria per far progredire la conoscenza nel senso in cui la scienza pretende di farlo. Tuttavia il fisico, come ogni altro scienziato, non conosce unicamente attraverso il suo “modulo mentale” di fisico. Egli è in contatto con altre scienze, con molte persone, e ricorda tutto quanto ha studiato nella vita, forse non nei libri, ma ascoltando maestri, esperti e cose del genere. Nella sua testa egli sta costantemente integrando questi flussi di conoscenza che occupano tutta la sua vita come persona umana, sin dall’infanzia. Questa conoscenza concreta ed integrata, dove in una forma difficile da definire si stanno costantemente integrando, e illuminando a vicenda, tutte le diverse conoscenze teoriche e pratiche, è quello che io chiamo, in mancanza di un nome migliore, conoscenza ordinaria, comune o personale.

Tale conoscenza contiene nel suo nucleo più intimo ed inderogabile, come ho detto prima, i primi principi, nel senso in cui li ho intesi sopra. Ma contiene molte altre cose, alcune in modo implicito, perché non ci ricordiamo in atto di tutto quanto sappiamo, o anche perché molte cose si possono sapere facilmente se compiamo certi atti. La visione comune, quindi, è attiva e non passiva. Non si riduce, come talvolta si è pensato, alle idee sociali ereditate dalla tradizione. Queste idee appartengono alla conoscenza comune solo come un ingrediente. Vi si aggiunge il sapere acquistato dall’esperienza personale, non inquadrabile nelle astrazioni scientifiche. Inoltre ad essa appartiene pure quell’ambito di riflessioni personali tramite i quali alcuni, se non tutti, si formano una certa “idea” della vita, il che potrebbe essere *in nuce* una visione prefilosofica (idee, ad esempio, sulla moralità, su Dio, sulla politica). La conoscenza personale corregge, anche inconsapevolmente, tutto quanto non è convincente nei saperi astratti o nelle idee filosofiche, anche perché cerca di integrare o di armonizzare, fino a un certo punto, le aree diverse del sapere acquisito tramite mezzi scientifici, retorici, estetici, religiosi, o altro.

Quanto ho detto può gettare una luce su tematiche quali i rapporti scienza e fede, scienza e filosofia, scienza e cultura. Spesso si dice che le prospettive filosofiche e scientifiche restano due ambiti diversi del sapere, che non vanno mescolati se non si vogliono creare confusioni metodologiche. Questo punto è vero. Ma è altrettanto chiaro che queste forme del sapere sono in interazione dinamica, spesso in modo non esplicito o riflesso, *nelle persone concrete*. La nostra testa può tenere separati ambiti conoscitivi diversi solo fino a un certo punto. Alcuni sono predominanti, altri si mantengono in periferia, alcuni restano intatti, altri sono alterati con una maggiore o minore radicalità. L’interazione tra i nostri saperi teorici e pratici forse si può capire meglio se li vediamo come *abiti*, cioè come possessi abituali non coscienti e dinamici, di cui si può disporre in vario modo a seconda dei casi (così come il possesso dell’abito linguistico dell’italiano e dell’inglese ci consente di formare frasi concrete in tali lingue). *L’intreccio di tutti gli abiti*

*conoscitivi*, di cui il più radicale è l'abito dei primi principi, è quello che ho chiamato *conoscenza personale*. Si può concludere, in questo senso, che la conoscenza comune o personale non è un bagaglio di conoscenze oggettive, o una sorta di magazzino cognitivo statico, ma è *un insieme di abiti in costante interazione, abiti molteplici ma dotati anche di una certa unità*, a causa dell'unità della persona che li possiede.

Vista la densità delle mie considerazioni precedenti in questo paragrafo introduttivo, presento in seguito una *schema riassuntivo*. Il lettore avrà capito quanto ho esposto se si sente in grado di commentare le frasi che seguono:

\* *La filosofia della mente* (FM) opera in un ambito interdisciplinare (scienze cognitive). Di conseguenza, occorrono prima di tutto alcuni chiarimenti epistemologici, propri della filosofia della scienza e della gnoseologia.

\* Questi chiarimenti sono specialmente necessari, dal momento che una parte importante della FM sostiene un *riduzionismo fisicalistico* (riduzione forte alla scienza fisica e agli atti fisici, in nome dell'anticartesanesimo o del dualismo).

\* Le *scienze* studiano sistematicamente aspetti particolari della realtà (astrazione). Esiste una pluralità di aspetti nelle cose e quindi il tentativo di riduzione del sapere ad un'unica scienza è impossibile.

\* *La filosofia* studia ciò che è ritenuto essenziale e fondamentale nelle cose (guarda la totalità nei suoi principi ultimi). Tiene conto delle scienze e si colloca in una prospettiva meta-scientifica. Ogni pretesa di giudizi essenziali (anche in modo negativo, o implicito) è filosofica.

\* *La riduzione forte della conoscenza a una scienza particolare* è un'operazione filosofica. Quindi è contraddittoria. I canoni delle scienze non giustificano mai le affermazioni filosofiche (cioè totali).

\* Anche se uno scienziato ritenesse poco chiara l'impresa filosofica, egli comunque dovrebbe essere consapevole della parzialità delle tesi scientifiche (la scienza non copre tutta la realtà, soprattutto non copre le realtà più importanti). Ma *la posizione filosofica*, almeno a livello di conoscenza personale, è *inevitabile*. Ogni giudizio valutativo globale -ad esempio, "la scienza è importante e vale la pena"- è filosofico.

\* *La conoscenza personale*, con le sue evidenze, è l'ultima istanza giudicativa, nei confronti di ogni forma di sapere, scientifico o filosofico. Essa contiene una filosofia implicita.

\* La squalifica radicale della conoscenza personale come “visione popolare” è *contraddittoria* (si fa usando la conoscenza personale).

\* Le filosofie o scienze che squalificano in modo radicale la conoscenza comune, ordinaria o personale, viste come “conoscenza prescientifica, volgare, popolare”, fanno parte *dell’ermeneutica forte*. Nell’ermeneutica forte il sapere personale è visto come superficiale e fenomenico, mentre l’istanza teoretica superiore che lo squalifica è vista come un livello profondo e reale. Ogni ermeneutica forte è contraddittoria.

\* *La conoscenza personale contiene* un nucleo forte di conoscenze ontologiche primarie, le esperienze acquisite, le tradizioni ricevute, il bagaglio culturale, l’applicazione concreta di tutte le scienze o idee filosofiche acquisite. La conoscenza personale coordina ed integra, non sempre con successo, tutti i saperi e abilità acquisite della persona, e li usa *ad casum*. Essa possiede un peculiare dinamismo, almeno in alcune delle sue parti. *È un abito, o un insieme di abiti*.

\* *La conoscenza ontologica primaria* (primi principi, certezze esistenziali fondamentali) è *continua, preverbale, preoggettiva o preconcettuale, vitale, abituale, implicita, operante per dare senso a ogni forma oggettiva di conoscenza*. Non è eliminabile, tranne una malattia mentale. È interpretata e approfondita dalle idee filosofiche, ma resta indipendente, in quanto la filosofia potrebbe ridurla o degradarla. Non è dimostrabile, ma è argomentabile indirettamente (per riduzione). Le virtù intellettuali danno più forza a questa conoscenza primaria. I vizi intellettuali la restringono, ma non possono eliminarla del tutto.

## II. SVILUPPO STORICO

Due strade sono possibili per presentare la tematica filosofica del rapporto tra il pensiero e il cervello, con tutti i suoi problemi annessi. La prima sarebbe iniziare con una panoramica delle scienze cognitive attuali, includendo gli aspetti neurofisiologici, per poi arrivare alle conclusioni filosofiche pertinenti, tenendo conto della varietà di interpretazioni. La seconda strada è cominciare con l’esposizione storica delle opinioni filosofiche sulla tematica, nate comunque in rapporto con l’andamento storico delle scienze cognitive, particolarmente la psicologia e la neurofisiologia. Ritengo preferibile, a scopo didattico, questa seconda procedura. Seguirò in una prospettiva storica l’intreccio tra la storia della filosofia della mente (anche se in altre epoche tale tematica non era vista sotto questa categoria) e la storia delle scienze umane pertinenti (quelle che oggi si dicono “scienze cognitive”). Il mio interesse è filosofico e non scientifico. Cercherò di enucleare nelle

brevi esposizioni scientifico-storiche soltanto ciò che è essenziale per la comprensione di idee filosofiche che non possono evitare di fare un riferimento allo stato delle scienze nel momento in cui tali idee appaiono.

Cornelio Fabro seguì invece la prima strada, quando egli studiò negli anni 40 la psicologia contemporanea per poi darne un'interpretazione tomistica alla luce della dottrina della cogitativa. Infatti, egli scrisse nel 1941 *La fenomenologia della percezione*, opera prevalentemente psicologica, e successivamente nello stesso anno pubblicò *Percezione e pensiero*, di carattere più speculativo. Però la situazione è diversa nel nostro caso, a causa della varietà e ricchezza delle scuole filosofiche sul problema che ci occupa. Il dibattito, inoltre, ha avuto altissima vivacità nella seconda metà del XX secolo e appartiene chiaramente al novero delle questioni più rilevanti che hanno agitato il pensiero filosofico contemporaneo.

Ritengo inoltre che l'approccio storico, sia scientifico che filosofico, sia il miglior modo di introdurci in modo efficace in problemi così ardui, ricchi e complessi come quelli che ci accingiamo a vedere. Se le opinioni, tendenze e scoperte non sono situate storicamente, il non esperto in materia rischia di disorientarsi nella giungla delle esposizioni attuali sulla filosofia della mente, le quali fanno riferimenti storici continui, non sistematici ma brevi e schematici, al solo scopo di presentare certe idee. Ma se il lettore profano non sa che cosa è veramente attuale e che cosa magari è superata, e fino a che punto, la riflessione filosofica personale diventa impossibile. Essa si può fare con successo soltanto una volta che si padroneggia la situazione storica. D'altra parte, anche nel modo di fare storia delle idee e delle scoperte scientifiche non di rado si presentano o al meno si preparano subdolamente le proprie tesi filosofiche, col rischio di fare ideologia anziché vera filosofia. E allora solo chi conosce la storia può discriminare sulla vera rilevanza di quanto si afferma.

Quanto ho appena detto è sempre vero per qualsiasi tematica filosofica. Ma riguardo alla situazione attuale della filosofia della mente direi che è *assolutamente indispensabile*. La complessità e la sofisticazione che queste materie hanno raggiunto negli ultimi anni scoraggia qualsiasi lettore impreparato. In questo campo la divulgazione è pericolosa, in quanto facilmente è ingannevole, data la moda di divulgare semplificando e introducendo idee in forme retoriche, sotto l'apparenza di oggettività scientifica. Il nostro argomento merita queste cautele. Vi è in gioco tutta una costellazione di valori e l'intera concezione antropologica.

## 1. La concezione classica

Diamo un breve sguardo alla filosofia delle mente dei classici greci e medievali. Il ricordo delle loro dottrine ci aiuterà a situare il problema contemporaneo. Non deve stupirci di trovare molto presto presso i classici la distinzione tra corpo ed anima o spirito, indipendentemente dalle teorie ontologiche. D'altra parte, anche nella letteratura religiosa e mitologica si profila, benché in modo oscuro, la distinzione interna all'uomo tra corpo sensibile e materiale e qualcosa di altro, che sopravvive ed è sede delle virtù.

Per quasi tutti gli antichi filosofi greci troviamo nell'uomo una dimensione non corporea, alla quale appartengono i pensieri e la razionalità (*lógos*). Questa parte dell'uomo è chiamata talvolta *pneuma* (spirito), talvolta *psiché* (anima), oppure *noús* (intelletto) o *lógos* (ragione). Non è corporea, almeno nel senso ordinario del termine corpo (ciò che è tangibile, visibile, spostabile, dimensionale, localizzato, soggetto a nascita, formazione e corruzione). Eppure si trova insediata nel corpo e in qualche modo lo dirige. I platonici parlano addirittura di una pluralità di anime nell'uomo, vedendo le loro funzioni svariate e la mancanza di unità tra di esse (funzioni istintive, vegetative, sensitive, intellettive). La visione delle molte-anime di solito è gerarchica. Esiste una parte superiore nell'uomo chiamata a predominare, affinché l'uomo non viva come una belva. Essa è normalmente la ragione o il *noús*, e nel platonismo, nonché nell'aristotelismo, è vista come costituita da una sorta di "sostanza" divina, cioè è qualcosa che ci accomuna agli esseri superiori e immortali, gli dei o la Divinità. Tale anima superiore (negli stoici è anche chiamata *l'hegemonikón*) contiene tutto ciò che c'è di più nobile nell'uomo. Nello stesso tempo, i rapporti delle anime con il corpo si intravedono in qualche modo, specialmente nel campo delle sensazioni, che incidono sugli organi dei sensi e influiscono sul pensiero, e nell'ambito delle emozioni o passioni, le quali sembrano avere la loro sede in certe parti del corpo (cuore, reni, organi sessuali) e hanno a che vedere con gli umori corporei (secondo l'antica teoria degli umori, che influiscono sul carattere). Già il cervello fu preso in considerazione dai medici greci quale organo centralizzato dei sensi e della memoria (opinione seguita da Platone, mentre Aristotele, stranamente, vede il cervello soltanto come dotato di una funzione di refrigerazione del corpo). La locomozione e la funzione parlante sono per quasi tutti guidate dalla ragione o dalla volontà. Il dominio sui movimenti del corpo appartiene all'esperienza immediata della libertà. Aristotele, seguendo Platone, avvertì che non dominiamo nello stesso senso tutti i nostri moti corporei. I moti provocati dalle emozioni sono naturali, e sono dominati indirettamente, in modo "politico" e non "despotico", nella misura in cui l'uomo può agire sulle rappresentazioni che attivano i desideri e altre passioni. Gli antichi

credevano anche facilmente all'influsso astrale sulla dimensione psicofisica delle persone, influsso considerato come sistematico e specifico dagli antichi astrologi.

Questo quadro ci dimostra l'esistenza presso gli antichi filosofi precristiani di una sorta di "dualismo funzionale" di tipo naturalistico e cosmologico, prima di essere elaborato in modo ontologico o di acquistare una certa coloritura secondo certi indirizzi filosofici e spirituali. Platone, ad esempio, non è un dualista forte, nel senso di svalutare il corpo umano, anche se le sue valutazioni in questo senso cambiano nel contesto delle sue opere e lungo la sua vita come filosofo. Platone vede le anime con una funzione primordialmente motrice del corpo. In questo senso è dualista. Con una concezione troppo passiva del corpo, visto quasi come inerte, l'anima si presenta in lui come la fonte del movimento autonomo, per cui doterà anche agli astri, come farà pure Aristotele, di motori-anime. Nel campo educativo ed estetico, Platone invece sostenne la necessità di mantenere in armonia dinamica l'anima e il corpo, attraverso gli esercizi e le funzioni simboliche.

Gli stoici sono meno dualisti di Platone e tendono ad un'ontologia monistica, a causa del loro concetto ampio di materia. Tutto il cosmo è dominato dal principio rettore, la Ragione universale, di cui l'uomo partecipa in modo speciale con la sua razionalità e la sua capacità linguistica. Non accentuano la necessità di separarsi dal corpo, bensì di dominarlo razionalmente, specialmente nell'orbita delle passioni, le quali in alcuni degli stoici sarebbero una forma diminuita di razionalità, una sorta di degradazione o malattia dell'anima. L'ideale stoico è arrivare ad una piena razionalizzazione del cosmo, poiché la ragione è il principio fondamentale della natura.

I platonici medi e ulteriormente il neoplatonismo, invece, si sono radicalizzati sul concetto di separazione. L'anima appartiene a un altro mondo, il mondo superiore, e in questa terra si sente punita ed in esilio. Le funzioni corporee si vedono quindi come un ostacolo, un ritardo o un fattore di pesantezza. La via religioso-filosofica viene impostata come un progressivo allontanamento dall'elemento corporeo. Però se nel manicheismo e la gnosi il mondo fisico appartiene al regno del male, nel neoplatonismo la natura fisica non è del tutto svalutata, ma è configurata piuttosto come l'ultima emanazione degli ambiti superiori, una specie di "spirito degradato" che pure può richiamarsi in maniera simbolica alla vita più alta.

In Aristotele troviamo una posizione peculiare, che sottolinea l'unità ontologica e funzionale dell'anima-corpo, grazie al contesto cosmologico universale dell'ilemorfismo, e allo stesso tempo non abbandona il principio platonico della separazione dell'anima intelligente rispetto al corpo. Lontana dal pansichismo o dall'ilozoismo, l'ontologia aristotelica è in grado di distinguere tra

mondo inanimato, mondo animato e addirittura senziente e ambito umano grazie al suo concetto flessibile di forma. La nozione di forma come atto, una sorta di principio attivo di organizzazione della materia e quindi della sua intelligibilità in quanto informata, gli consente di superare il materialismo di coloro che si ritenevano soddisfatti con la spiegazione delle cose in funzione della loro materia costitutiva. Questa nozione di materia come potenza serve anche per dare una definizione del materialismo. Ma in Aristotele la materia e la forma non sono visti come opposti, bensì come due principi complementari e non separabili nel mondo (tranne l'intelletto). La forma è l'elemento non-materiale del mondo fisico. Le forme possono essere più attive nella loro funzione ontologica e così si apre la via alla comprensione del concetto di anima come atto di un corpo funzionale ed organico, destinato alle funzioni vitali. La conoscenza sarà spiegata dallo Stagirita in funzione dell'immaterialità, un'immaterialità parziale che in realtà comporta un'elevazione del corpo a una forma di esistenza più alta (corpo vivente, corpo che sente le sue funzioni, e quindi diventa in qualche modo più singolare, più rivolto su se stesso).

Nel caso dell'intelletto, invece, Aristotele intravede un superamento assoluto della sfera materiale e quindi un principio trascendente e universale, candidato alla divinità e ad una funzione cosmica. Per Aristotele questo problema era fondamentale. Mentre in Platone la trascendenza al di sopra del mondo sensibile corrisponde alle essenze ideali contemplate dalla mente umana, in Aristotele, al di sopra dell'ambito fisico delle sostanze "normali", sostanze ilemorfiche, esiste un ambito di "sostanze separate", cioè un ambito in cui la forma sussiste per sé stessa, quindi è sostanza, il che appartiene di diritto al *noús*. Così arriva Aristotele a Dio. L'attività del *noús* è contemplativa, ma è anche motrice, in un senso non meccanico, che lo Stagirita riesce a concepire prevalentemente come finalistico (creando così una complicata cosmologia, pur non priva di una certa bellezza).

Il problema dell'uomo è che sembra di avere qualcosa di tale *noús*, in quanto egli è capace di comprensione intellettiva. Ma a questo punto Aristotele diventa oscuro e non prosegue. Per alcuni filosofi arabi, come sappiamo, la parte immortale e "separabile" dell'uomo non sarebbe altro che una forma di essere comunicata con un Intelletto separato. Il rapporto anima-corpo in Aristotele, quindi, va visto nella linea del concetto di *informazione*, e dal punto di vista operativo come necessità del pensiero astratto di attivarsi solo in rapporto ad immagini ed esperienze. Ricordiamo che Aristotele in qualche senso è il fondatore della psicologia. Così egli ha delineato in modo notevole, per la prima volta nella storia, le funzioni sensitive, la memoria, l'immaginazione, le leggi associative della memoria, la coscienza, l'istinto, l'intelligenza e la volontà.

Nel periodo cristiano, la concezione sui rapporti anima-corpo in un primo periodo segue le orme del platonismo e poi, con Tommaso d'Aquino, dell'aristotelismo. La novità introdotta dal Cristianesimo è la rilevanza della persona, non solo come concetto ontologico, ma anche come nucleo assiologico ed esistenziale. L'Incarnazione di Dio e il dogma della risurrezione dei corpi sottolineano l'importanza del corpo e impediscono il sopravvento dell'estremo spiritualismo platonico. In Agostino l'anima viene vista pure in modo "fenomenologico", come un'interiorità inquieta, con una dimensione temporale biografica in cui conta molto la memoria, la coscienza e l'amore, con una ferma consapevolezza dell'autonomia della volontà. L'anima è in questo senso *imago Trinitatis* e sta costantemente alla ricerca di Dio, della verità eterna. Questa percezione dell'anima come autocoscienza libera si colloca al di sopra del naturalismo greco e della visione psicologico-cosmologica plotiniana. L'anima, nel vocabolario cristiano, diventa l'io interiore ed esistente ed è anche la persona. Nondimeno per Agostino il corpo è essenziale alla costituzione dell'uomo.

In Tommaso d'Aquino, la ripresa di Aristotele gli consente il passaggio ad un'ontologia dell'unione monosostanziale tra l'anima spirituale e il corpo umano, in continuità con una psicologia dove il pensiero astratto è collegato alla base sensitiva. Per l'Aquinate non c'è incompatibilità tra l'immanenza informativa sostanziale dell'anima nel corpo organico umano e la sua trascendenza nelle sue funzioni superiori (intelletto e volontà). Le forme più alte sono più immateriali, e ciononostante organizzano meglio il corpo, per cui si può dire che i corpi più ricchi e complessi nell'universo sono quelli informati da atti o forme più alti. Allora non risulta strano che l'anima umana sia atto del corpo e che al contempo sia relativamente indipendente dal corpo.

Tale indipendenza Tommaso la vede in primo luogo funzionalmente. Il corpo umano non si esaurisce nelle sue operazioni come corpo, di tipo vegetativo, anche se elevato all'ordine spirituale (il "mangiare" umano è anche un atto con una dimensione spirituale). Esso svolge pure funzioni al servizio dello spirito (contemplare la realtà visibile, leggere, parlare). Ciò consente Tommaso di pensare al corpo umano nello stato di gloria in termini di uno statuto ontologico alquanto diverso dal nostro corpo. Un corpo glorioso è un corpo in cui risplende meglio l'anima e le sue attività superiori, come l'amore e la contemplazione. Le penalità corporee sono disarmonie nate dal peccato originale e non semplici limitazioni fisiche. Non nascono dai difetti corporei, ma da una sorta di mancata informazione dell'anima umana sul corpo.

Inoltre, la *commisuratio* tra anima e corpo umano, in quanto l'uno è fatto per l'altro e viceversa (cfr. *C. G.*, II, c. 81, n. 1621), rende difficile a Tommaso spiegare la situazione dell'anima

separata, nella quale persiste il rapporto essenziale al proprio corpo e non ad un altro, nonché la necessità di rivolgersi alla parte sensitiva per dare concretezza al pensiero. L'anima umana può sopravvivere, essendo immortale, ma il suo stato privo di corpo cui informare non è naturale. L'idea di *corpo proprio* comporta la sua importanza ontologica (l'anima umana non può attualizzare nessun altro corpo, se non il proprio), anche se rende più difficile la spiegazione della modalità della risurrezione del corpo.

L'unità tra corpo ed anima non è l'unità estrinseca, relazionale (o strumentale) tra due cose o sostanze. L'anima è atto del corpo e quindi si trova semplicemente in tutto il corpo animato, non stando presente in un senso locale, come conviene alle realtà corporee. Per Tommaso gli organi sono pure animati da particolari "energie" dell'anima e quindi ogni organo è un tutt'uno con la facoltà o energia animatrice corrispondente. Questo principio gli consente di concepire il cervello, seguendo principi della medicina araba, come sede ed organo delle funzioni sensitive superiori, perfino la memoria e la "cogitativa" o razionalità particolare (cfr. *S. Th.*, I, q. 78, a.4). "Il funzionamento efficace delle potenze sensitive interiori, quali l'immaginazione, la memoria e la cogitativa, richiede una buona disposizione del cervello (*bona dispositio cerebri*). Perciò l'uomo possiede un cervello maggiore tra gli altri animali, secondo la proporzione della sua quantità" (*De Anima*, q. un., a. 8, c.). Per cui, "la lesione di alcuni organi impedisce all'anima di comprendere direttamente se stessa e le altre cose, come quando si verifica una lesione al cervello" (*De Spiritualibus Creaturis*, q. un., a. 2, ad 7; cfr. *S. Th.*, I, q. 84, a. 7, dove si fa riferimento a malattie mentali -"frenetici", "letargici", causate da lesioni fisiche, che impediscono l'uso dell'intelligenza).

L'unità ontologica anima-corpo, organo-funzione animatrice (facoltà) si traduce nell'unità di azione: l'atto organico non va attribuito né all'anima né al corpo, ma al composto o sintesi di anima-corpo, organo-funzione. L'atto sensitivo o passionale è del corpo animato. Invece la trascendenza delle funzioni strettamente spirituali esclude l'attribuzione formale delle sue operazioni al corpo. Per Tommaso non esiste un organo fisico del pensiero o della volontà, anche se l'attuazione di queste facoltà richiede l'attività sensitiva superiore. In altre parole, in mancanza di un'adeguata attivazione corporeo-sensitiva, la parte spirituale dell'anima non può passare ad operare e rimane solo come una potenzialità nascosta.

## 2. Cartesio e i moderni

La perdita della filosofia della natura e della vita in Cartesio (1596-1650) e la sua impostazione gnoseologica dell'anima produce un nuovo modo di impostare la dualità anima-corpo. In Cartesio l'anima concepita come forma del corpo scompare. Perciò, anche se egli continua

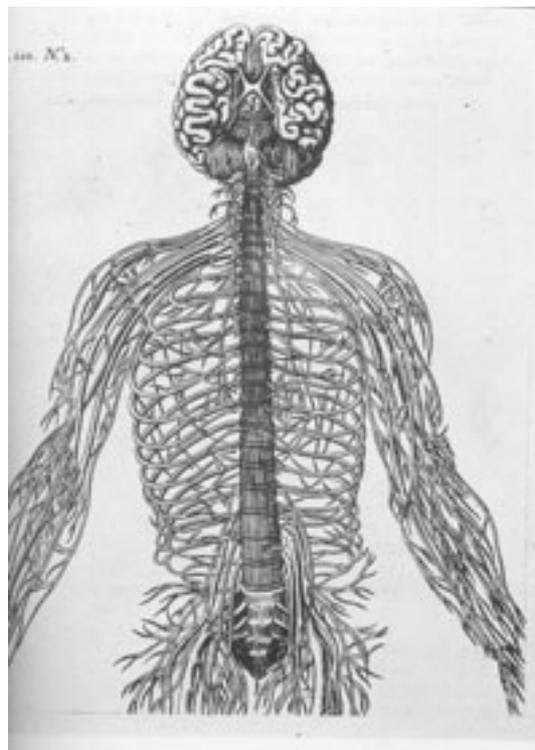
a parlare di anima, il termine più adatto è quello che diventerà usuale più tardi, *mente*, dal momento che l'anima è ridotta a un *locus* dove si contengono le rappresentazioni cognitive (pensieri, idee, rappresentazioni). Infatti, l'attuale filosofia *della mente* eredita da Cartesio questa terminologia, pur nel suo rifiuto della visione cognitiva del filosofo francese.

L'impostazione cartesiana rende la convivenza di mente e corpo *dualistica*, nel senso di un dualismo di due sostanze che non hanno niente in comune, se non il fatto di essere in comunicazione causale contingente (dipende dal volere di Dio) e di integrare la natura dell'uomo. L'organismo umano e anche quello degli animali viene concepito soltanto secondo i parametri della nuova fisiologia meccanica e quindi non è descritto né analizzato in termini di qualità sensibili, sensazioni, potenze o anima. La fonte del moto autonomo naturale non è più l'anima, come si pensava nell'antica meccanica platonico-peripatetica, ma va ricondotto al principio di inerzia e alle nozioni di forza e di quantità di movimento, quindi alle leggi della nuova meccanica (in Aristotele, invece, l'automovimento era segno della vita). I movimenti organici -vita fisica- sono spiegati, inoltre, in termini di calore. Lo spirito umano non si conosce a causa della necessità di arrivare a un principio ultimo dell'organizzazione del corpo umano, ma si pone solo in quanto appare come pensiero. Anzi in questo senso lo spirito pensante è la cosa più certa, nella sua autotrasparenza che non ammette dubbi (*cogito, ergo sum*). Il corpo umano è pure una nostra realtà, sì, ma problematica. Ciò che non è problematico è la *coscienza*: penso, sento, immagino. In Cartesio, lo spirito è ridotto alla coscienza, mentre cade in un certo senso l'importanza della distinzione tra sensazione o pensiero, visto che in fin dei conti tutte e due sono rappresentazioni coscienti. Perfino la sensazione dolorosa è una "rappresentazione" -legata al corpo- e non invece, come risulta più palese, un semplice atto del corpo dell'animale. Ma se noi ammettiamo come unica descrizione del corpo animale quella che ci offre la fisica -e la fisiologia cartesiana vuole essere pura fisica, il che comunque favorì lo sviluppo autonomo della moderna biologia-, allora il dubbio sullo statuto delle sensazioni si pone facilmente, come vedremo nelle discussioni del XX secolo.

D'altra parte, occorre rapportare la mente pensante alla fisiologia del corpo. Cartesio è considerato un *interazionista*, in quanto riteneva che la mente poteva interagire col corpo attraverso la *ghiandola pineale* (l'epifisi) del cervello. Non bisogna vedere in questa spiegazione una stranezza eccessiva. Il dettaglio fisico, benché antiquato, non è significativo. Ciò che conta è la necessità sentita da Cartesio di spiegare in che modo lo spirito umano, sia il pensiero che la volontà libera, è in grado di subire un influsso dal corpo e di muoverlo, il che d'altra parte ci risulta abbastanza ovvio come fatto. La sua teoria si può trovare specialmente in *L'homme*, scritto scientifico di anatomia e fisiologia umana, elaborato tra il 1632 e il 1633 (pubblicato in latino nel 1662 e in

francese nel 1664). Un'altra opera importante per la questione mente-corpo in Cartesio è il suo trattato *Passions de l'âme*, pubblicato nel 1649, un anno prima della sua morte.

Forse colpisce al lettore l'apparente modernità dell'impostazione di Cartesio di questi trattati, pur nei limiti scientifici legati allo stato del sapere del suo tempo. In verità i suoi trattati biologici riflettono le conoscenze anatomiche, fisiologiche e mediche diffuse del Seicento, ormai molto avanzate sin dal Rinascimento, in continuità con la scoperta europea della fisiologia greca (Ippocrate, Galeno) e al contempo in fase di rapido superamento. Il filosofo francese ha un'idea abbastanza globale e precisa sul modo in cui si producono nell'uomo le sensazioni a partire dagli organi periferici fino alla sede centrale del cervello, da cui escono i comandi che muovono i muscoli e così rendono possibili i movimenti delle membra del corpo. In sintonia con le pratiche del suo tempo, egli ama fare disegni didattici molto chiari, che oggi chiameremmo "modelli spiegativi". Ecco la rappresentazione del "sistema nervoso" apparsa nella pubblicazione latina della sua opera:

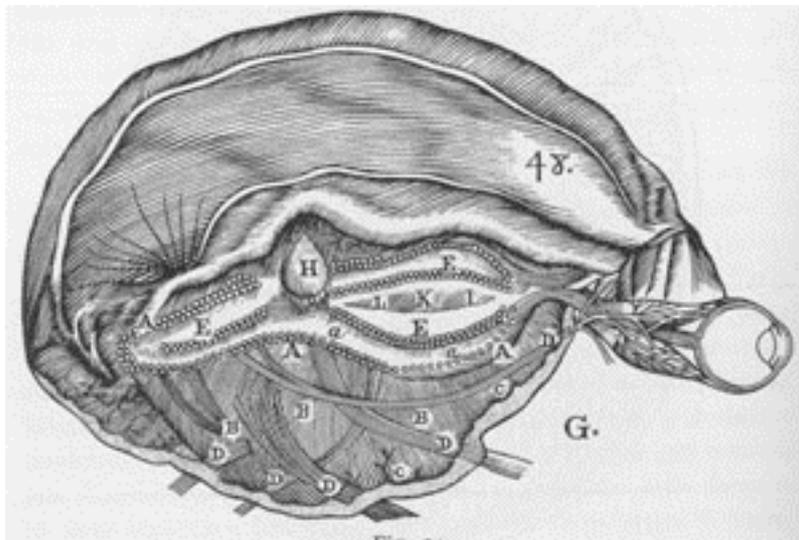


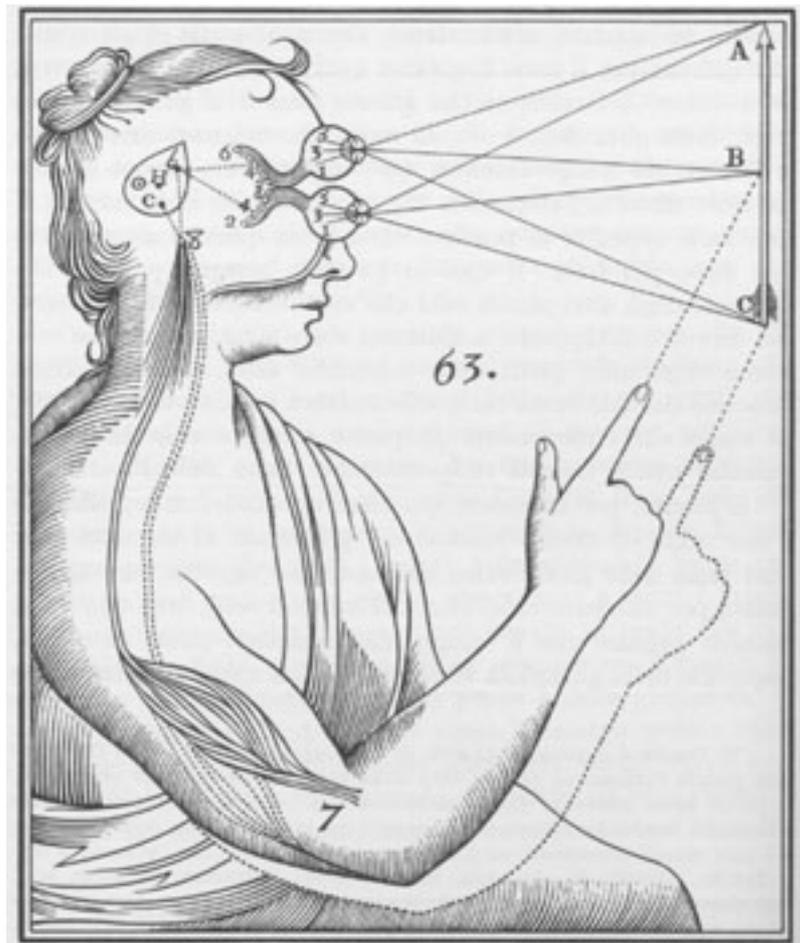
Nelle opere indicate egli tenta, con metodo ipotetico e non sperimentale, e nondimeno basato su certe esperienze note, una ricostruzione abbastanza completa della base nervosa delle sensazioni esterne ed interne (memoria ed immaginazione), nonché delle passioni sensibili. Possiamo considerare *psicofisica* la spiegazione cartesiana. Egli presuppone l'anima umana, intesa come puro spirito, o come la parte superiore dell'uomo che riceve a modo suo gli influssi del corpo e li domina attraverso la volontà e le rappresentazioni. Cartesio presuppone a parole, non senza una certa incoerenza con il resto della sua dottrina, la tesi tradizionale dell'unità ontologica dell'anima umana

con il corpo, e la presenza di tutta l'anima in tutto il corpo. Il dualismo cartesiano non è un “dualismo puro”: non bisogna semplificare la sua dottrina, nella quale egli innova, ma al contempo mantiene una continuità con la tradizione medica e filosofica.

Un punto fondamentale della metodologia cartesiana è il suo ricorso alla spiegazione “dal basso”, cioè secondo le strutture fisiologiche. Anziché spiegare a partire dall'anima, egli preferisce andare agli ipotetici dinamismi fisici, dimostrando in questo punto la nuova mentalità scientifica. Se l'uomo muore, non è perché l'anima abbia abbandonato il corpo, ma perché il corpo “si è rotto” e quindi non è più in grado di contenere l'anima. La sua visione del corpo è meccanica, nel senso che egli concepisce l'organismo come se fosse una macchina, e non vede un'eterogeneità assoluta tra le macchine fabbricate dall'uomo e gli organismi vitali. Molti esempi e modelli per immaginare come potrebbero essere le strutture fisiologiche sono prese da meccanismi artificiali. Il “meccanicismo” consiste nel concepire il corpo come un insieme di parti fisiche coordinate in modo locale e in rapporto reciproco in base sia a spostamenti tra i suoi componenti che a “colpi” (spinte, pressioni).

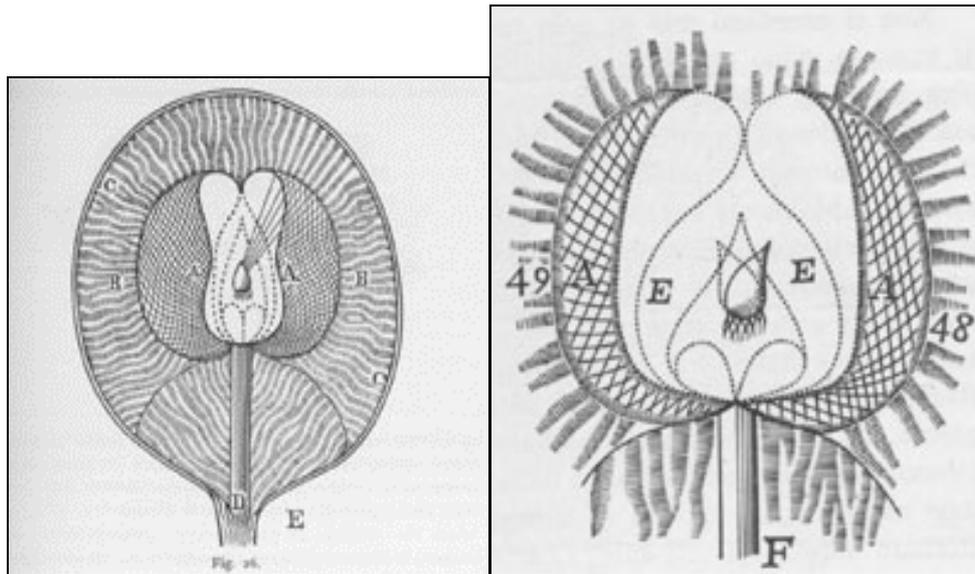
Ecco il suo modello di cervello: si osservi a destra l'occhio, e al centro, col nome di H, la ghiandola pineale, e più in giù un disegno delle vie visive.





I. *Entriamo in qualche dettaglio della teoria cartesiana.* La comunicazione intracorporea, secondo la visione antica, presa dai classici e notevolmente approfondita nella scienza europea sin dal Rinascimento, si produce attraverso i nervi, il sangue e i muscoli. I nervi sono come fili conduttori che comunicano le diverse membra sensoriali e motorie al cervello, facendosi “notare” ad esso tramite i loro movimenti meccanici. I nervi sono costituiti da filamenti e da piccoli tubi che li racchiudono, e da “spiriti animali” (nozione di derivazione galenica) o particelle sottili e veloci - quasi come particelle in uno stato molto dinamico della materia, intermedio tra l’aria o vento e il fuoco- derivate dal flusso del sangue, i quali si formano nel loro arrivo dal cuore alla ghiandola pineale, che sta al centro della concavità cerebrale (cfr. *L’uomo*, in R. Descartes, *Opere scientifiche*, vol. 1, a cura di G. Micheli, Utet, Torino 1998, pp. 70-72). Gli spiriti animali sono comunicati ai muscoli e così servono per provocare i moti corporei, spesso in risposta ad una sensazione (ad esempio, sento il fuoco nel piede e rapidamente lo ritiro).

Ecco due disegni del cervello, dove si nota al centro la piccola ghiandola mobile in base all’agitazione provocata dal corpo (nervi) o dall’anima.



*L'ipotesi sulla causa del movimento muscolare quindi è la seguente:* “Di mano in mano che questi spiriti [animali] entrano così nelle concavità del cervello, passano di lì nei pori della sua sostanza e da questi pori nei nervi, ove, secondo che essi entrino o anche che tendano soltanto ad entrare, più o meno negli uni che negli altri, hanno la forza di cambiare la figura dei muscoli in cui quei nervi sono inseriti e di far muovere con questo mezzo tutte le membra. Così come potete aver visto nelle grotte e nelle fontane che sono nei giardini dei nostri re, che la sola forza con cui si muove l’acqua uscendo dalla sua fonte è sufficiente per muovere diverse macchine e anche per far loro suonare alcuni strumenti, o pronunciare alcune parole, secondo la diversa disposizione dei tubi che la conducono” (*L’homme*, in *Opere scientifiche*, cit., pp. 72-73).

Cartesio sviluppa qui con dettaglio *l’immagine della fontana*, e l’arricchisce con altri paragoni. Le sensazioni sono illustrate con la metafora di persone che continuamente entrano in queste grotte, dove c’è la fontana governata da macchine, e che camminano su piastrelle, in modo che provocano solo con la loro presenza lo scatto di meccanismi che finiscono con l’arrivare al cervello per muovere svariatamente le sue parti e in particolare la ghiandola pineale, che è come la “centrale” di arrivo delle sensazioni e di comando dei moti muscolari.

“E infine, quando *l’anima ragionevole* sarà in questa macchina, avrà la sua sede principale nel cervello e sarà lì come il fontaniere che deve essere nei castelli ove vanno a rendersi tutti i tubi di queste macchine, quando vuole provocare o impedire o cambiare in qualche maniera i loro movimenti” (*ibid.*, p. 73). L’immagine del fontaniere è simile ad altre tradizionali nel dualismo psicofisico, come quella del suonatore di uno strumento musicale o del pilota in una nave. Eppure Cartesio in altri momenti sostiene esplicitamente l’unione sostanziale anima-corpo, poiché “non è sufficiente che sia collocata nel corpo umano, così come un pilota sulla sua nave, se non forse per

muovere le sue membra, ma [che ] c'è bisogno che sia congiunta e unita più strettamente con esso per avere, oltre a ciò, dei sentimenti e degli appetiti simili ai nostri, e comporre così un vero uomo” (*ibid.*, p. 182). Il dualismo cartesiano non va troppo semplificato. Cartesio sostiene esplicitamente l'unità sostanziale, d'accordo con la tradizione, ma non la elabora ontologicamente, e in questo senso essa rimane un mistero per il filosofo francese, come ebbe occasione di notare in una lettera alla principessa Elisabetta di Boemia. Il dualismo si presenta invece al momento di dover spiegare l'interazione, e in modo particolare il ruolo dell'anima come motore del corpo.

*Lo spirito umano è introdotto dal Creatore nella macchina del corpo, precisamente come elemento nel cervello che riceve in forma di idee l'azione degli oggetti sensibili sul corpo e risponde in forma di moto fisico sul corpo.* Il corpo umano è concepito secondo il modello di una macchina. “Suppongo che il corpo non sia altro che una statua o una macchina di terra che Dio forma espressamente per renderla il più possibile simile a noi: per modo che non solo dia ad essa il colore e la figura di tutte le nostre membra, ma vi metta anche all'interno tutti i pezzi che si richiedono per fare sì che cammini, mangi, respiri e imiti infine tutte quelle nostre funzioni che si può immaginare procedano dalla materia e non dipendano che dalla disposizione degli organi. Vediamo orologi, fontane artificiali, mulini e altre macchine simili le quali, pur non essendo fatte che da uomini, nondimeno hanno la forza di muoversi da sé in parecchie maniere diverse” (*ibid.*, p. 58).

La frase citata sta all'inizio dell'opera. Le ultime parole rivendicano *il meccanicismo nei riguardi del corpo*. La vita organica è descritta secondo principi strettamente fisici e non psichici. “Queste funzioni [vitali] seguono tutte in modo naturale, in questa macchina [il corpo umano], dalla sola disposizione dei suoi organi, né più né meno di quanto fanno i movimenti di un orologio o altro automa, in seguito a quella dei suoi contrappesi e delle sue ruote; per modo che non bisogna concepire in essa alcuna altra anima vegetativa, né sensitiva, né alcun altro principio di movimento e di vita che non sia il suo sangue e i suoi spiriti, agitati dal calore del fuoco che brucia continuamente nel suo cuore e che non è di natura altra da quella di tutti i fuochi che sono nei corpi inanimati” (*ibid.*, p. 154).

Si potrà dire che il dolore, la fame o la sete sembrano *passioni* animali. Ma Cartesio assimila ogni sensazione al pensiero. “Così, a causa del fatto che non concepiamo che il corpo pensi in alcuna maniera, abbiamo ragione di credere che tutte le specie di pensieri che sono in noi appartengono all'anima” (*Les passions de l'âme*, in *Opere scientifiche*, cit., p. 255). La divisione

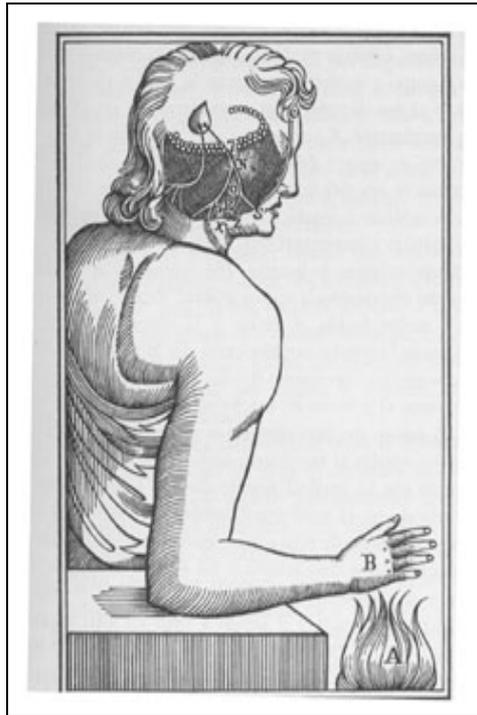
rigida stabilita tra pensiero (=coscienza) e corpo esteso obbliga Cartesio a eliminare la possibilità che la materia “senta”. La materia ha solo movimento e calore (cfr. *ibid.*).

La concezione meccanicistica del corpo umano consente di presagire facilmente la possibilità dei *robot*. E infatti Descartes afferma, alquanto oscuramente, che “senza che ci sia alcuna anima in questa macchina, essa può naturalmente essere disposta ad imitare tutti i movimenti che dei veri uomini, oppure altre macchine simili, faranno in sua presenza” (*L’homme*, in *Opere scientifiche*, *cit.*, p. 137). In modo più esplicito, nel *Discorso del metodo* egli parla della possibilità di costruire macchine che riuscirebbero a imitare le figure e i movimenti degli animali (*automi*) e perfino dell’uomo (cfr. *ibid.*, p. 179), anche se “avremmo sempre due mezzi molto certi per riconoscere che non sarebbero per ciò dei veri uomini” (*ibid.*, pp. 179-180): uno è la capacità di elaborare un linguaggio creativo in base a composizioni sempre nuove tra le parole, per rispondere a qualsiasi domanda (il linguaggio dell’automa umanoide sarebbe costituito da risposte determinate agli stimoli ricevuti); in secondo termine, la ragione è uno strumento universale, che può servire per tutto, mentre gli automi potrebbero fare molto bene, anche meglio dell’uomo, alcune cose particolari, ma mancherebbero in altre (cfr. *ibid.*, p. 180).

Ma vediamo adesso brevemente come Cartesio concepisce la funzione mediatrice tra lo spirito e il corpo della ghiandola pineale nell’ambito delle *sensazioni*. “Ora vi dirò che, quando Dio unirà un’anima ragionevole a questa macchina, così come pretendo di dirvi in seguito, le darà la sua sede principale nel cervello e la farà di natura tale che, secondo le diverse maniere con cui le entrate dei pori che sono nella superficie interna di questo cervello saranno aperte per il tramite dei nervi, avrà sentimenti diversi” (*ibid.*, p. 88). Concretamente, immagina Cartesio, se i nervi sono tirati con forza e si rompono, “il movimento che causeranno nel cervello darà occasione all’anima (...) di avere il sentimento del *dolore*” (*ibid.*, p. 88).

Il dolore, quindi le sensazioni, sono dell’anima, ma sono “occasionati” dai movimenti del corpo che vengono “comunicati” all’anima tramite i movimenti della ghiandola pineale. In modo simile egli spiega l’attuarsi delle altre sensazioni.

Un disegno in proposito è il seguente:



La ghiandola pineale è la sede dell'immaginazione e del senso comune per Cartesio (cfr. *ibid.*, p. 126). E lì che si produce la misteriosa trasformazione di un movimento corporeo in un'idea, in base alla quale "l'anima ragionevole considererà immediatamente, allorché, essendo unita a questa macchina, immaginerà o sentirà qualche oggetto. E notate che dico immaginerà o sentirà, dato che voglio comprendere generalmente sotto il nome di *idea*, tutte le impressioni che possono ricevere gli spiriti uscendo dalla ghiandola H [pineale, chiamata spesso H da Cartesio], le quali si attribuiscono tutte al senso comune, allorché dipendono dalla presenza degli oggetti" (*ibid.*, pp. 126-127). Si noti il concetto moderno di idea, valido sia per le sensazioni che per i pensieri. L'essenziale è che l'idea è una *rappresentazione cosciente*, cioè è l'anima che sente se stessa tramite immagini o pensieri (da cui si deduce che gli animali non possono avere sensazioni). Quando alla ghiandola H arriva un'immagine dall'occhio, tale immagine fisica si trasforma nella coscienza dell'immagine, in fondo diventa pensiero. La pineale è un organo mobilissimo per Cartesio. Ogni suo piegarsi verso un lato o verso l'altro, causato dal basso o dall'alto (dallo spirito o dal corpo), provoca il suo contatto con tubi diversi e così si produce tutto il dinamismo del corpo e dell'anima in un modo integrato.

In base a questa semplice ipotesi Cartesio tenta di spiegare, in un modo abbastanza dettagliato e sempre meccanico, la memoria, il sonno, i sogni, tutte le sensazioni e la motricità, quindi l'intera vita psichica. La ghiandola pineale non è un'idea secondaria nella psicofisiologia cartesiana, ma ne è il suo nucleo. Eppure, Cartesio non è stato assolutamente originale, in quanto egli in realtà prende le funzioni di tale ghiandola da Galeno e dalle idee comuni nella medicina del

XVII secolo. “C’è bisogno anche di sapere che, benché l’anima sia congiunta con tutto il corpo, c’è nondimeno in esso qualche parte nella quale esercita le sue funzioni più particolarmente che in tutte le altre. Si crede comunemente che questa parte sia il cervello o possa essere il cuore; il cervello, a causa del fatto che è ad esso che si rapportano gli organi dei sensi, e il cuore a causa del fatto che è come in esso che si sentono le passioni” (*Les passions de l’âme*, in *Opere scientifiche*, cit., pp. 272-273). Infatti Aristotele poneva il cuore come sede della sensibilità, mentre Platone propendeva per il cervello. Ma l’aristotelismo medievale considerava pure il cervello come sede centrale dei sensi, seguendo l’aristotelico Galeno, secondo il quale la massa cerebrale era il principio dei nervi, di ogni sensazione e di ogni movimento volontario. Descartes amplia notevolmente la funzione della pineale, che così diventa il cardine della sua concezione sui rapporti dinamici tra l’anima e il corpo. La sua scelta è dovuta alla piccolezza e presunta mobilità della ghiandola, sensibile a qualsiasi minimo movimento dei flussi delle minime particelle (“spiriti”), e alla sua situazione centrale nel cervello, per cui non sarebbe lateralizzata. Comunque Cartesio assegna pure una funzione speciale al cuore nella produzione delle passioni corporee, che sono collegate al flusso del sangue. Ma, in definitiva, vogliamo notare *una certa continuità tra la tesi cartesiana e il principio degli aristotelici secondo cui la sede della cogitativa è il cervello*. Tralasciando il dualismo, la tesi di Cartesio non è così bizzarra per lo stato della scienza del suo tempo.

Il seguente testo, benché lungo, sintetizza la concezione cartesiana della pineale. “Concepiamo dunque qui che l’anima ha la sua sede principale nella piccola ghiandola che è in mezzo al cervello, donde irradia in tutto il resto del corpo per il tramite degli spiriti, dei nervi e anche del sangue che, partecipando alle impressioni degli spiriti, le può portare, attraverso le arterie, in tutte le membra; e ricordandoci di ciò che è stato detto sopra della macchina del nostro corpo, cioè, che i piccoli filamenti dei nostri nervi sono distribuiti in modo tale in tutte le sue parti che, in occasione dei diversi movimenti che vi sono eccitati dagli oggetti sensibili, aprono diversamente i pori del cervello. Ciò fa sì che gli spiriti animali contenuti nelle sue cavità entrino diversamente nei muscoli e per mezzo di ciò possano muovere le membra in tutte le diverse maniere in cui sono capaci di essere mosse, e anche che tutte le altre cause che possono diversamente muovere gli spiriti bastino per condurli in diversi muscoli. Aggiungiamo qui che la piccola ghiandola che è la sede principale dell’anima è sospesa in modo tale tra le cavità che contengono questi spiriti, da poter essere mossa da essi in altrettante diverse maniere quante diversità sensibili vi sono negli oggetti; che tuttavia essa può anche essere mossa diversamente dall’anima, la quale è di natura tale da ricevere tante diverse percezioni quanti diversi movimenti accadono in questa ghiandola. Come anche, reciprocamente, la macchina del corpo è composta in modo tale che, per il solo fatto che questa ghiandola è mossa diversamente dall’anima o da qualsivoglia altra causa, spinge gli spiriti

che la circondano verso i pori del cervello, che li conducono nei muscoli attraverso i nervi e per mezzo di ciò fa muover loro le membra” (*ibid.*, pp. 274-275).

II. *Le passioni*. Una simile spiegazione psicofisica si trova nella già citata opera di Cartesio riguardo alle passioni (emozioni, sentimenti) (ved. sul tema, A. Malo, *Certezza e volontà. Saggio sull’etica cartesiana*, Armando, Roma 1994, pp. 91-133). Mentre ordinariamente l’argomento delle passioni era discusso in sede retorica o morale, qui abbiamo un tentativo di spiegazione fisico-naturale, sempre seguendo l’idea del rapporto anima-corpo tramite la ghiandola pineale. Ad esempio, si legge in *Les passions de l’âme*, quando vediamo una figura strana e spaventevole, nociva al corpo, viene eccitata nell’anima la passione del timore, o dell’ardimento, a seconda del carattere dell’anima o del temperamento del corpo. Se ciò succede, gli spiriti dell’immagine che si è formata sulla ghiandola procedono verso i nervi che servono per voltare le spalle e fuggire, altri penetrano nel cuore, agitano il sangue, il che si ripercuote a sua volta sul cervello, aumentando la paura, ecc. (cfr. *Opere scientifiche*, p. 276) In questo senso le emozioni, pur essendo atti dell’anima, sono causate dal corpo e quindi sono relative al corpo.

Vediamo altri esempi. Se i liquori dello stomaco che entrano col sangue non trovano abbastanza cibo da dissolvere, spiega Cartesio, la loro forza si volge contro lo stomaco stesso, e “agitando i piccoli filamenti dei suoi nervi in un modo più forte del consueto, fanno muovere le parti del cervello da dove vengono. Ciò sarà causa del fatto che l’anima, essendo unita a questa macchina, concepirà l’idea generale della fame” (*L’homme*, in *Opere scientifiche*, cit., p. 112; si osservi come la sensazione di fame è chiamata “idea”). Troviamo una simile spiegazione per la sete, ma sorprendentemente anche riguardo a passioni più spirituali, come la gioia e la tristezza. Se il sangue che va nel cuore è più puro e acceso del normale, dispone il nervo che lo ricollega al cuore e in questo modo si causa il sentimento di gioia, e se accade il contrario è suscitata la tristezza (cfr. *ibid.*, pp. 112-113). L’abbondanza, le dimensioni, i movimenti e differenze tra gli spiriti animali, continua il filosofo, spiegano le differenze negli umori o inclinazioni naturali negli uomini, suscitando movimenti quali la bontà, la liberalità, l’amore, la fiducia, l’arditezza, la costanza, la prontezza, la diligenza, il desiderio, la tranquillità d’animo, oppure la malignità, la timidezza, l’incostanza, la tardità, l’inquietudine. Cartesio riconosce comunque una differenza tra le passioni “che si rapportano al corpo” (fame, sete) da quelle “che si rapportano alla nostra anima”, e ammette altre cause per queste ultime, collegate alla percezione o all’immaginazione (cfr. *Les passions de l’âme*, pp. 268-269).

Se vogliamo evitare con una certa benevolenza un'interpretazione materialistica della parte emotiva dell'uomo secondo Cartesio, potremmo dire che *egli si riferisce alla base temperamentale e fisica* per cui gli uomini possono essere portati a certi stati d'animo o umori, il che non è in disaccordo con la tradizione. Ciò che interessa al filosofo francese, con cui conclude il suo trattato sulle passioni, è *il controllo umano della parte emotiva*. In generale, il dominio della volontà sulle rappresentazioni consiste nel fatto che l'anima è in grado di muovere la ghiandola pineale in diversi sensi, spingendo così gli spiriti animali, ad esempio, sulle tracce lasciate nel cervello dalle precedenti percezioni, per ottenere un ricordo (cfr. *ibid.*, p. 279), o verso altre parti, se si vogliono produrre immaginazioni di cose mai viste, e così via in rapporti ad atti come voler fissare lo sguardo, voltarlo verso un'altra parte, o muovere la lingua e le labbra per parlare.

*Il dominio sulle passioni si ottiene indirettamente*, tramite la considerazione mentale dei pensieri o ragioni che possono suscitare, ma qui Cartesio riconosce, come è tradizionale, dei conflitti tra i movimenti volontari e quelli provenienti del corpo (sentire paura, eppure non volerlo). L'anima non sempre riesce a dominare le passioni, che quindi hanno una loro autonomia dinamica a causa del loro radicamento nel corpo. Ci sono alcune anime forti, che dominano bene le loro passioni; altre sono deboli e non ci riescono. Cartesio pensa a un controllo delle passioni non solo tramite le virtù e la conoscenza della verità e del bene, ma anche *attraverso la formazione di certe abitudini fisiche*, così come i cani e altri animali domestici possono essere addestrati. Significativamente si legge: "Queste cose sono utili da sapere per dare a ciascuno il coraggio di studiarsi di regolare le proprie passioni; poiché si può, infatti, mutare i movimenti del cervello negli animali privi di ragione, è evidente che lo si può ancor meglio negli uomini e che anche quelli che hanno le anime più deboli potrebbero acquisire un imperio completamente assoluto su tutte le loro passioni, se si impiegasse abbastanza industria ad addestrarli e a condurli" (*ibid.*, p. 287).

Sembra delinarsi in questa conclusione un po' sconcertante la possibilità di un *controllo fisico completo dell'affettività umana*, simile a quello che si osserva nella condotta degli animali-macchine, i quali comunque non hanno passioni ma sono soltanto meccanismi. Cartesio è influito dallo stoicismo nella sua idea del controllo delle passioni. Eppure egli sembra passare qui dal piano delle virtù o del dominio razionale a quello del controllo medico della macchina del corpo, in particolare del cervello. All'esercizio della virtù, sembra suggerire il testo, potrebbero aggiungersi esercitazioni psicofisiche di natura tecnica o "industriosa" per raggiungere il dominio delle passioni (sentimenti, emozioni) sottomesse al controllo indiretto della volontà. Queste pratiche, seguendo la linea di pensiero accennata in queste considerazioni, sarebbero ad esempio una dieta sana, il rilassamento, esercizi fisici, un ambiente esterno favorevole. Senza pensare a peculiari interventi sul

cervello, gli antichi conoscevano alcuni metodi fisici ordinari per attenuare il ribollire di alcune passioni specialmente legate al corpo: dormire bene, mangiare, passeggiare, lavorare con intensità, evitare le occasioni e cose del genere. Queste attività sono naturali e anche umane, in quanto dipendono dalla volontà e in questo senso possono far parte dell'ambito della virtù. Ed è ovvio che esse incidono in una maniera naturale sul cervello, dove si trova la sede della sensibilità e dell'affettività legata al corpo.

*In una prospettiva tomistica, le virtù umane non sono puramente spirituali, ma possono avere una base fisica* (per non parlare delle consuetudini storico-sociali). Per temperamento ed educazione ricevuta sin dall'infanzia, alcune persone possono essere tendenzialmente più ordinate, sobrie, laboriose o altro. “La maggiore o minore intensità della concupiscenza -scrive Tommaso d'Aquino- può dipendere da una duplice causa. Certe volte procede da cause fisiche. Alcuni, a motivo della loro costituzione fisica (*ex naturali complexione*) sono più inclini da altri alla concupiscenza, oppure perché alcuni hanno più opportunità di altri per i piaceri che accendono la loro concupiscenza. Allora la debolezza della concupiscenza diminuisce il merito, mentre la sua maggiore forza lo accresce. Però altre volte l'intensità della concupiscenza può essere dovuta a qualche causa spirituale lodevole, ad esempio alla veemenza della carità o alla forza della ragione, come avviene nell'uomo che pratica la temperanza. In questo caso, la debolezza della concupiscenza aumenta il merito per motivo della causa indicata, mentre la sua maggiore forza lo diminuisce” (*S. Th.*, II-II, q. 155, a. 4, ad 2). Più avanti leggiamo: “la disposizione corporale può dare origine a passioni veementi, poiché l'appetito sensitivo è una facoltà dotata di un organo corporeo” (*ibid.* q. 156, a. 1, corpus).

Nei limiti della concezione antica, Tommaso segue Aristotele addirittura quando ritiene, in questo senso, che le donne hanno una costituzione fisica “debole” che le rende più soggette alle forze passionali (cfr. *ibid.*, ad 1 e ad 2), un punto senz'altro sbagliato, derivato dal maschilismo greco di Aristotele. L'antica teoria dei temperamenti, basata sul predominio degli “umori”, rende gli uomini svariatamente sottomessi alla forza di certe tendenze ed emozioni, nell'ottica di San Tommaso. “L'impeto delle passioni può essere dovuto alla velocità [temperamentale di reazione], come vediamo nei collerici, oppure alla veemenza, come avviene nei melanconici, i quali sono esageratamente accesi dalla passione a causa della loro costituzione fisica. Oppure può succedere che qualcuno non sia perseverante nelle sue scelte perché la sua adesione ad esse è troppo debole a causa della debolezza della sua costituzione fisica” (*ibid.*, q. 156, a. 1, ad 2). Questa base naturale della tendenza al bene non è la causa della virtù o del vizio, ma soltanto l'occasione.

Nella visione tomistica, ogni uomo riceve una natura “psicofisica” dotata di certe caratteristiche positive e negative, e tocca a lui impegnarsi, con meriti disuguali a seconda del punto di partenza, nell’acquisizione e crescita della virtù. Troviamo quindi una certa concordanza, in questo senso, tra la dottrina tradizionale e le osservazioni cartesiane sulle condizioni del dominio delle passioni. Il suo punto originale, mi pare di capire, è che tale dominio non dipende esclusivamente dall’esercizio delle virtù, ma può anche essere coadiuvato da consigli medici e psicologici per quanto si riferisce alla base fisica delle passioni. Però in Cartesio la preoccupazione fisiologica in tal senso appare sbilanciata rispetto alla dimensione propriamente morale o della virtù. Questo predominio della tecnicità per migliorare le proprie prestazioni è dovuto alla concezione del corpo come macchina e non come un vero corpo umano (cfr. A. Malo, *Certezza e volontà...*, cit., pp. 132-133).

*L’opera di Cartesio riguardo alla tematica corpo-mente è un esempio del modo concreto in cui si può sviluppare il dualismo psicofisico in un senso moderno, cioè in sintonia con la nuova scienza biologica.* I trattati biologici cartesiani come quelli cui ci siamo riferiti non rappresentarono un apporto importante per la scienza moderna. Cartesio non passa alla storia come biologo. Biologi di rilievo della sua epoca (cfr. *Storia delle scienze*, vol. 1, a cura di E. Agazzi, Città Nuova, 1984, pp. 365-392, a carico di L. Belloni) per citare soltanto alcuni, sono Andrea Vesalio (1514-1564), della scuola di Padova, il quale cominciò a demolire il sistema anatomo-fisiologico di Galeno; Michele Serveto (1511-1553), catalano, scopritore della piccola circolazione del sangue; Andrea Cesalpino (1519-1603), Girolamo Fabrizi (1537-1619), professore di anatomia e fisiologia a Padova; William Harvey (1578-1657), scopritore della circolazione del sangue, inglese formatosi a Padova (questa scoperta, accettata da Cartesio nel *Discorso del metodo*, è uno degli eventi più importanti nella storia della medicina); Marcello Malpighi (1628-1694) (Università di Pisa, Bologna, Messina), fautore dell’anatomia microscopica.

*La metodologia di Cartesio nel suo trattato L’homme* consiste nel costruire con fantasia una macchina complicata simile al corpo umano, un modello biomeccanico, ovviamente d’accordo con le conoscenze anatomiche e fisiologiche del suo tempo. L’idea del uomo-macchina avrà un notevole influsso nella filosofia posteriore. Il nucleo filosofico del pensiero cartesiano al riguardo, come abbiamo visto nell’esposizione precedente, è l’intervento causale delle mente sul corpo biomeccanico. Forse una buona immagine di questa teoria è quella delle marionette manovrate dal marionettista con le sue dita “invisibili” e i fili apposti, nello scenario che talvolta è stato chiamato -ad esempio da Dennett- il “teatrino cartesiano”. In questo senso possiamo concludere:

a. Ogni filosofia che riconosca la dimensione spirituale della persona, non riducibile a corpo, *necessariamente affronta le questioni discusse da Cartesio*, in particolare la tematica della distinzione e al contempo unità tra le due dimensioni, sia a livello di costituzione ontologica che nell'ordine delle operazioni, nonché il problema dell'influsso reciproco tra spirito e corpo. La questione della "sede" delle funzioni più alte si pone sempre, prima ancora della moderna teoria delle localizzazioni. La sede più ovvia in questo senso, se tralasciamo le antiche teorie sul cuore e altre ancora (reni, organi sessuali), è il cervello. Il ruolo assegnato da Cartesio alla ghiandola pineale, in questo senso, è solo aneddótico. Oggi un dualista parlerebbe della corteccia, dei lobi frontali, o del sistema limbico, in termini non troppo lontani dalle proposte cartesiane (come farà Eccles).

b. La soluzione cartesiana del problema indicato è *drastica*. Anche se egli continua ancora con certe idee vaghe di tipo vitalistico, recepite dalla tradizione di Galeno, come per esempio il concetto degli "spiriti animali" (senza la minima base empirica), egli comunque cerca di darli un'interpretazione il più meccanica (=chiara e distinta) possibile. Gli spiriti giocano in Cartesio il ruolo che nella fisica moderna sarebbe rappresentato dalle onde o particelle associate a onde. Il suo dualismo è *drastico* dal momento che pone una distanza abissale tra corpo meccanico e pensiero puro, a causa dell'esigenza troppo aprioristica di lavorare soltanto con quelle due idee "chiare e distinte". Così egli si pone in contrasto con l'esperienza che facilmente assegna sensibilità e passioni agli animali. Il risultato paradossale di questo sforzo di chiarezza analitica è l'oscurità con cui l'azione meccanica si dovrebbe "trasformare" in idea psichica (ad esempio, il salto enorme tra l'immagine fisica che arriva alla ghiandola pineale e l'atto immanente di un "occhio" pensante che la "vede"). Questa causalità non è spiegata, ma solo dichiarata. Ogni volta che il fuoco si avvicina al mio piede, "sento" la scottatura: come è possibile che il moto meccanico di certi nervi che muovono in un certo senso la ghiandola pineale "si traduca" in un'idea? A volte Cartesio si esprime in termini di "occasione", come avvertendo la difficoltà di concepire un transito causale da moto a idea. Perché un movimento locale dovrebbe produrre una idea? Sembra affacciarsi la tesi occasionalista, secondo cui Dio ha creato una così perfetta unione tra anima e corpo, che ogni volta che il corpo si muove in un certo senso, automaticamente l'anima subisce certe sensazioni o emozioni. L'unione anima-corpo rimane estrinseca, dovuta all'arbitrio divino.

III. *La causalità mente-corpo*. Il problema suscitato da Cartesio agli occhi dei filosofi fu soprattutto causale. La strana causalità idea-moto meccanico non era convincente e dimostrava la difficoltà di lavorare soltanto con la metodologia delle idee chiare e distinte, le quali in fondo erano

le rappresentazioni matematiche. A partire da Cartesio, nel periodo successivo della filosofia moderna, nascono *le seguenti posizioni*:

1. *Occasionalismo* (Malebranche e altri): la causalità che intercorre tra la mente e il corpo è da attribuirsi all'azione di Dio. Non appena si pongono certi atti adeguati, come l'avvicinamento della mano al fuoco, Dio interviene causalmente a provoca l'effetto apparente, come la sensazione di bruciore o il movimento fisico. La tesi occasionalista denota una crisi del concetto di causalità. Non è da sorprendersi la riduzione di Hume della casualità a successione di eventi. Questa crisi dunque non incide solo sul rapporto mente-corpo, ma anche sulla causalità tra le diverse entità del mondo. All'eliminazione della nozione di forma e di finalità segue una mancanza di intelligibilità del modo in cui le sostanze possono comunicare. Poi seguirà la crisi della nozione di sostanza.

2. *Parallelismo psicofisico* (Leibniz, Spinoza): gli eventi meccanici e mentali sono indipendenti ma paralleli, a causa di una coordinazione precisa stabilita dal Creatore. In fondo questa tesi è simile a quella dell'occasionalismo, solo che invece di postulare una continua azione divina, quasi miracolosa, preferisce trasferire la causalità divina al coordinamento generale del mondo. Dio creerebbe due sostanze mirabilmente connesse, in quanto le alterazioni della mente sarebbero covarianti alle alterazioni fisiche e viceversa. Il parallelismo di Leibniz corrisponde alla sua teoria dell'armonia prestabilita, una teoria in realtà molto sofisticata, nel senso che Leibniz in fin dei conti nasconde un monismo panpsichista, nel quale il corpo meccanico ed esteso praticamente è ridotto ad un'apparenza. Ma le monade "mentali" leibniziane, quasi come punti mentali inestesi, hanno diversi gradi di coscienza (il che apre la porta alla nozione di psiche inconscia, dotata di una percezione molto debole). Il numero delle monadi è infinito, per cui Leibniz è da ritenersi un pluralista (distinzione reale tra Dio e le molteplici monadi). Spinoza invece è parallelista in un senso più rigorosamente monista: corpo e spirito non sono che due attributi o due manifestazioni di un'unica sostanza, Dio.

Anche Kant in un certo senso è parallelista, in quanto egli considera la duplice possibilità della descrizione determinista o libera della condotta umana come un'antinomia della ragione (3<sup>a</sup> antinomia) che si risolverebbe solo con il passaggio alla ragione pratica. Quindi il problema affrontato da Cartesio, Leibniz e altri produrrebbe l'antinomia perché rimane ancorato alla filosofia dogmatica. L'antinomia kantiana nasce in un senso molto specifico a partire dalla contrapposizione tra determinismo fisico e indeterminismo della libertà, il che è uno dei lati dell'abisso esistente tra mondo spirituale e mondo fisico. Se ogni evento fenomenico è sottomesso alle rigide leggi del determinismo, come postula la visione scientifica classica, diventa problematico riconoscere che le

manifestazioni fisiche umane siano libere. La soluzione kantiana è “parallelista” in un senso trascendentale. Cioè lo spirito e il corpo sono tenuti separati e il problema dell’interazione non si pone sul piano dogmatico.

Il parallelismo psicofisico fu adottato, in un senso meno impegnativo, da uno dei fondatori della psicologia scientifica, W. Wundt (1832-1920). Il parallelismo psicofisico considera le sequenze psichiche come parallele a corrispondenti sequenze fisiche, senza crearsi problemi di interazione. *Alla sequenza “vedo un gelato → desidero mangiare un gelato → mangio il gelato e provo un particolare piacere”, corrisponde una covariante sequenza tra modifiche fisiche nell’occhio, cervello e lingua.* Ma è difficile convincersi che tra l’impressione fisica dell’immagine del gelato sul cervello e la sensazione di vederlo, col conseguente desiderio di mangiarlo, non ci sia un rapporto di reale dipendenza. Il senso comune lo suggerisce, ma gli strumenti analitici per spiegare la causalità si sono divisi tra due mondi, quello della fisica pura e quello della psicologia introspettiva. La situazione epistemologica di quest’ultima resterà instabile nel XX secolo.

*Il problema della causalità nel territorio semantico della fisica, al di là delle interpretazioni filosofiche -le quali finiranno, nel positivismo, per eliminare il concetto di causa- è che la spiegazione fisica viene normalmente assunta in un modo chiuso.* È questo anche il problema di Kant. Ogni evento fisico dev’essere spiegato da un’altro evento fisico precedente, secondo leggi precise formulate in termini matematici (equazioni). “La nozione secondo cui il mondo materiale è causalmente chiuso si rapporta alla nostra concezione delle leggi naturali. Le leggi naturali governano i rapporti causali tra eventi materiali. Tali leggi sono diverse dalle leggi approvate dai corpi legislativi. Una legge naturale non ha eccezioni: non può essere violata così come si può violare una legge di traffico. Un oggetto che si comporta in un modo strano o inaspettato si conforma tuttavia perfettamente alla legge naturale. La prova che il comportamento di un oggetto viola una data legge naturale è la prova che quello che pensavamo che era legge non lo è” (J. Heil, *Philosophy of Mind. A contemporary introduction*, Routledge, Londra 1998, p. 23). Perciò, l’intervento di una mente non fisica sul mondo fisico, anche se può sembrarci qualcosa di normale se osserviamo il comportamento umano senza la pura prospettiva fisica, agli occhi di una visione scientifica fisica totalizzante sembra un miracolo continuo, un fenomeno non previsto, inspiegabile secondo le leggi.

Si potrà dire che il fisico guarda soltanto il mondo fisico, non il mondo dell’uomo, del quale si dovrebbero occupare altre scienze, come la psicologia o la sociologia. Il problema è che nel comportamento umano abbiamo un intreccio di mondo umano e fisico: “voglio” muovere la mano

con libertà, e ecco che la muovo, un movimento su cui dovrebbe essere competente la fisica. Se ammettiamo che la mano e il cervello, a un certo livello primario, sono “mossi” dalla mia volontà (causa psichica, non materiale), stiamo riconoscendo l’intervento “anomalo” di una causa non fisica sul territorio dei fisici. Questo punto pare evidente. Anche i *risultati* dei movimenti della mano, ad esempio se pensiamo al lavoro di Michelangelo per scolpire *La pietà*, non si spiegano secondo canoni fisici.

A questo punto sorge tipicamente la difficoltà della *violazione del principio di conservazione dell’energia*. Una volontà che muova il cervello (ad esempio, la ghiandola pineale cartesiana) starebbe creando nuova energia nel mondo, il che è inconcepibile per la fisica (la stessa obiezione può essere sollevata contro la possibilità dei miracoli e della creazione divina dell’universo).

Si è tentato di rispondere a quest’obiezione in diversi modi, che adesso possiamo elencare:

a. L’intervento della libera volontà non apporta nuova energia, ma soltanto *dirige* quelle già esistente nel corpo umano (Van Melsen, *Philosophy of Nature*). Possiamo aggiungere: *usa* l’energia disponibile. Certamente se voglio alzare il braccio e non ho la forza muscolare di farlo, a causa di una paralisi, non potrò muoverlo. Ma il solo fatto di volerlo o di immaginare di farlo provoca un’alterazione nel cervello. La domanda allora è: come fa la mente a “dirigere” o ad “orientare” gli eventi neurali? Il problema sussiste.

b. La questione svanisce se diamo un’interpretazione strumentalistica della scienza. Il principio di conservazione non avrebbe una validità ontologica, ma sarebbe soltanto un presupposto o un postulato per risolvere le equazioni della fisica. Questa risposta è poco convincente. Il problema si pone veramente solo se si presuppone il realismo scientifico.

c. Non bisogna prendere il principio di conservazione, possono dire alcuni, in un modo troppo rigido. La fisica quantistica e la moderna cosmologia della gravitazione quantistica cominciano a relativizzare tale principio. L’obiezione nascerebbe, in questo senso, in un quadro di interpretazione unilaterale del suddetto principio, legata particolarmente al determinismo e alla fisica classica.

e. Popper, in *L’io e il suo cervello*, vol. 1 (Armando, Roma 1981), dove sostiene una posizione interazionista rinnovata, cerca di sdrammatizzare la difficoltà. Già Cartesio proponeva, con una certa chiarezza, che la mente cambierebbe minimamente soltanto la direzione della ghiandola pineale, la quale poteva così deviare gli spiriti animali verso tutte le dimensioni. Anche il conduttore di un’automobile cambia dal di dentro la direzione del veicolo (cfr. pp. 219-220). Confesso di non capire questa risposta: l’agente che cambia la direzione richiede pure una certa

energia propria per fare tale manovra, anche se può usare una fonte di energia indipendente. In seguito Popper aggiunge che se l'azione della mente sul corpo è concepita in modo non meccanico, bensì come un fenomeno elettrico, la difficoltà diventa “trascurabile, perché la massa della corrente elettrica deviata è quasi uguale a zero” (p. 220). Tuttavia, quasi uguale a zero non è zero. Il passaggio all'ambito elettromagnetico non annulla il problema dell'energia, benché ne riduca la necessità fino a un punto “trascurabile”. È trascurabile a effetti pratici, ma per il problema teorico.

f. Altri cercarono di scavalcare la difficoltà mediante il ricorso alle leggi statistiche. Poiché il passaggio dallo stato S1 a un nuovo stato S2 è probabile, si potrebbe pensare a un intervento della mente che non violerebbe le leggi di conservazione, dato che il nuovo evento provocato dalla mente sarebbe comunque fisicamente probabile. Semmai si potrebbe ammettere una violazione non accertabile delle leggi fisiche, come se ci fossero dietro “variabili nascoste”. È vero che tale intervento cambierebbe le probabilità, ma in modo indiscernibile (cfr. J. Heil, *Philosophy of Mind*, cit., pp. 24-26).

L'argomento prende più forza se il carattere probabilistico delle leggi fisiche è interpretato nel senso di un reale indeterminismo della materia. In questa linea, molti pensano che la difficoltà contro l'intervento della mente sul mondo nascerebbe dal determinismo. Una volta ammesso l'indeterminismo, la libertà potrebbe intervenire nel mondo senza violare alcun principio fisico (anzi, in questo senso si potrebbe anche capire la possibilità dei miracoli). Molti fisici, sin dai tempi di Heisenberg, si sono sentiti attratti da questo argomento. L'indeterminismo preconizzato dalla meccanica quantistica aprirebbe la via alla possibilità di una nuova concezione del rapporto anima-corpo, anzi aprirebbe lo spazio alla libertà e agli interventi divini sul mondo, sia miracolosi che “naturali”. La tesi è interessante in quanto non si limita meramente a risolvere un'obiezione, ma cerca di capire positivamente *come* si potrebbe concepire l'azione dello spirito sul corpo.

Questa posizione, in fondo condivisa da Popper, un grande fautore dell'indeterminismo come condizione per l'esistenza della libertà, è stata sostenuta da Penrose ed Eccles. Scrive R. Penrose che “alcuni hanno espresso la speranza che, in tali circostanze [effetti quantistici sul cervello], *l'indeterminazione quantistica* possa essere il varco attraverso cui la *mente* influenza il cervello fisico (...) Forse la “libera volontà” di una “mente esterna” potrebbe essere in grado d'influenzare le scelte quantistiche che risultano da tali processi non deterministici” (*Ombre nella mente*, Rizzoli, Milano 1996). Ma egli sostiene a tale scopo la necessità di “scoprire come l'influenza della “mente” possa aver luogo a una scala molto più grande di quella delle singole particelle quantistiche. Dobbiamo cercare dove sia mai il punto di transizione tra il livello classico e quello quantistico”

(*ibid.*, pp. 426-427). Una proposta neurologica concreta in tal senso, seguendo in qualche modo le orme di Cartesio, è quella del popperiano J. Eccles. “La nuova luce sul problema mente-cervello viene dall’ipotesi che gli eventi mentali non-materiali (il Mondo 2 di Popper) siano legati agli eventi nervosi del cervello (il Mondo 1 di materia ed energia) attraverso processi conformi ai principi della fisica quantistica” (*Come l’io controlla il suo cervello*, Rizzoli, Milano 1994).

La soluzione di Eccles all’obiezione relativa alla conservazione dell’energia è la seguente: “Secondo i critici materialisti, l’ipotesi che gli eventi mentali immateriali possano agire in molti modi su strutture materiali come i neuroni incontrerebbe difficoltà insuperabili. Tale effetto presunto degli eventi mentali sarebbe incompatibile con le leggi di conservazione della fisica, in particolare con la prima legge della termodinamica. Questa obiezione sarebbe stata certamente sostenuta dai fisici del XIX secolo e dai neuroscienziati e filosofi che, ideologicamente, sono rimasti alla fisica del XIX secolo, senza riconoscere la rivoluzione operata dai fisici quantistici nel XX secolo (...) In un recente libro [*The Miracle of Existence*, Ox Bow, Woodbridge 1984], l’eminente fisico quantista Margenau ha offerto un contributo fondamentale. Si tratta di una notevole trasformazione dalla fisica del XIX secolo il sentirsi dire che ‘alcuni campi, come i campi di probabilità della meccanica quantistica, non hanno né energia né materia’ (p. 22). Riassumendo, la mente può essere considerata un campo, nel comune senso del termine. Ma si tratta di un campo non-materiale; l’analogo più simile è forse un campo di probabilità” (*Come l’io controlla il suo cervello*, cit., pp. 99-100).

Sinceramente la teoria di Eccles non mi pare molto convincente. La nozione di Margenau di “campi non materiali” correlati a campi di probabilità per poter così stabilire un “ponte” tra lo spirito e la materia appare oscura e rivela una certa confusione tra spirito e materia.

g. Un’altra forma di aggirare la difficoltà del principio conservativo è il ricorso alla nozione di *informazione*. L’intervento della mente sul cervello consisterebbe nell’aggiunta di informazione, il che non richiederebbe l’utilizzo di energia, così come, per fare un paragone, mediante un computer noi possiamo dare comandi che scatenino importanti conseguenze fisiche, col semplice utilizzo di una tastiera. Comunque il problema del transito della sfera mentale a quella fisica persiste. Esiste pure la tentazione del monismo, secondo cui i poteri mentali sarebbero una sorta di “energia fisica” particolare capace di intervenire in un mondo fisico aperto grazie alla sua indeterminazione.

*Personalmente ritengo che sia difficile ovviare le difficoltà di capire l’intervento causale dello spirito sull’universo materiale seguendo un concetto di causa situato nel quadro epistemologico della fisica moderna, a meno che non riduciamo in qualche modo lo spirito a una*

sorta di entità materiale superiore, un po' come gli stoici immaginavano che le realtà spirituali, perfino Dio, erano materiali, dando un senso più alto e flessibile al concetto di materia, oppure che, all'inversa, riduciamo la realtà materiale a una forma prefigurante lo spirito, come avviene nel pansichismo o in proposte affini come quelle di Leibniz. Soluzioni classiche estreme in questo senso, molto naturali se vogliamo evitare un dualismo assoluto che rende incomunicabili la mente e la materia, tranne meccanismi inventati, esterni e quasi miracolosi, sono le dottrine di Berkeley (riduzione schietta di tutta la realtà a percezione psichica) o di Hobbes (riduzione di tutto a materia, tesi ripresa nelle attuali teorie dell'identità).

La forma in cui la causalità è vista nella fisica moderna è utile per capire le realtà materiali o le componenti materiali di qualsiasi tipo di entità più alta o complessa, ma è *inadeguata per la comprensione delle realtà non materiali*. In questo senso il problema dell'interazione causale mente-corpo, così come si profila da Cartesio in avanti, *a mio avviso è uno pseudo-problema, poiché si cerca di capire la causalità della mente sul corpo con uno schema adeguato solo per la spiegazione della causalità materiale* (la questione del principio di conservazione dell'energia appartiene interamente a questo piano di spiegazione). La fisica può spiegare i movimenti del corpo umano secondo le leggi fisiche, che non sono mai violate neppure negli atti fisici volontari. Ma non può pretendere di dar ragione così di tutta l'attività umana. Di conseguenza, essa deve rimanere aperta ad altre prospettive epistemologiche.

Il piano più alto di spiegazione comunque non dovrà essere semplicemente "aggiunto" alla spiegazione fisica, appunto perché la prospettiva epistemologica cambia radicalmente. Neanche si può tentare di elevarsi a un punto di vista più abbracciante come quando si passa, ad esempio, dalla meccanica classica alla meccanica relativistica o quantistica, nel tentativo di unificazione della fisica. In tale unificazione rimane sempre una certa omogeneità di metodo fisico-matematico. Non è sufficiente, anche se può essere utile, il ricorso all'indeterminazione, se non si passa al livello veramente superiore di spiegazione. Il fisico dovrebbe avvertire questo punto anche senza uno speciale approfondimento nella sfera "interiore" della libertà, solo guardando i risultati della prassi umana. Benché i movimenti del corpo umano, come dicevamo, siano fisicamente spiegabili, i loro risultati, come sono tutte le opere di arte e della civiltà, tutta la storia umana, sono completamente al di fuori della portata delle scienze naturali.

*Risposta specifica al problema della presunta violazione del principio di conservazione dell'energia nel caso dell'intervento di una causalità immateriale sui corpi.*

In un modo più generale, il problema è nato nell'ambito della fisica tradizionale in rapporto a qualsiasi intervento spirituale causale sui corpi. Concretamente: creazione divina; miracoli; causalità della libertà umana sul corpo umano.

Riguardo alla *creazione divina iniziale*, il problema non esiste in alcun modo, dal momento che essa è causa creativa ontologica in un senso completamente diverso dalla causalità fisica. Dio crea l'universo con tutte le sue leggi, perfino quella relativa alla conservazione dell'energia. Se una violazione di questo principio comporta una "creazione di energia", oppure no, è un problema interno alla scienza della fisica, non filosofico. In ogni caso, un'eventuale "creazione di energia" non va necessariamente vincolata alla causalità di Dio, poiché il problema è prettamente fisico, e nemmeno bisogna concepire la causalità creativa divina come una sorta di "fonte originaria di energia". Tutto ciò comporterebbe una confusione tra causalità di Dio e causalità fisica.

Rispetto ai *miracoli*, non si vede neppure una vera difficoltà in relazione al menzionato principio della fisica. I miracoli sono effetti straordinari compiuti dalla potenza divina nella natura, al di sopra delle potenze naturali e di ogni legge fisica. Quindi, un miracolo non può "violare" il principio conservativo. Semmai, il miracolo opera al di sopra di esso, così come opera al di sopra di ogni legge fisica. Ma non per questo motivo diciamo che il miracolo è una *violazione* delle leggi fisiche. Del resto, una causa non fisica può agire al di sopra delle leggi fisiche. Le leggi fisiche non sono "inviolabili" come se fossero dei principi metafisici, il cui contrario comporta una contraddizione. Dio è libero da condizionamenti fisici per agire liberamente nel cosmo fisico.

Qualcosa di analogo si può dire in rapporto ad eventuali *interventi divini creativi* nel corso dello sviluppo del cosmo, postulati da alcuni filosofi o teologi (per esempio, i "teologi del processo") per dar ragione di certi salti ontologici che altrimenti sarebbero riconducibili a eventi emergenti del processo evolutivo della natura. Nell'ipotesi di un tale intervento, non si violerebbe propriamente alcun principio fisico, ma soltanto si agirebbe al sopra della causalità naturale.

Qualcuno forse potrà domandare: ma se ci sono miracoli o tali interventi di Dio, bisognerebbe comunque dire che "nel mondo viene creata miracolosamente nuova energia". Se Dio risuscita un morto, si crea nuova energia? Lasciamo parte della risposta al fisico: sarebbe per lui verificabile tale incremento anomalo di energia, senza ricorrere all'ipotesi di una causa intramondana di tale risurrezione? Possiamo pure ammettere (senza concedere) che nel mondo si fosse introdotta in modo "non legale" nuova energia. E allora? Credo che per il fisico ciò sia irrilevante. Se fosse necessario affermare che si crea nuova energia, ciò sarà comunque una parte dell'evento miracoloso. Ma come si potrebbe verificare tale apporto di energia secondo canoni sperimentali?

Oggi il principio conservativo, come è stato ricordato sopra, è alquanto relativizzato (non negato). Pare che l'energia "non sempre" si dovrebbe "conservare", o non in modo assoluto. D'altra parte, come prima ho pure detto, alcuni autori, cercando alcune "condizioni fisiche che rendano possibili gli interventi immateriali sul mondo fisico, al di sopra delle leggi fisiche", hanno fatto ricorso all'indeterminismo, alle leggi statistiche o al concetto d'informazione. Mi pare interessante, ma sinceramente non ne vedo la necessità, anzi intravedo il rischio di "abbassare" gli interventi causali divini ad un livello troppo fisico. La causalità di Dio è più alta. Lui è il Creatore. Non dobbiamo concepire la sua causalità secondo la modalità delle cause naturali.

Per risolvere questo problema, abbiamo anche notato come alcuni hanno segnalato l'importanza di non prendere la causalità fisica delle scienze in un modo *chiuso*, come se non ci fossero altri tipi di cause. L'osservazione è giusta. Ma non nel senso di poter aggiungere una causalità superiore "concorrente" ma in fondo analoga alla causalità fisica, bensì nel senso che *la causalità contemplata dalle scienze naturali è pur sempre una causalità parziale*. Perciò, la causalità fisico-chimica non arriva mai a spiegare pienamente l'ambito biologico, e tanto meno l'ambito sensitivo della vita animale. Alla fine, tale causalità è *del tutto inadeguata* per spiegare ciò che è specificamente umano (scienze, arti, cultura, atti umani). Man mano ci eleviamo nell'ordine dei gradi dell'essere, la causalità fisica delle scienze naturali diventa sempre più parziale. L'economia non si può spiegare, ad esempio, con concetti elettrici o termici. Ricordiamoci di Socrate in carcere (secondo Platone), consapevole che lui si trovava lì non a causa dei movimenti dei muscoli delle sue gambe, ecc., ma a conseguenza del voto di alcuni cittadini di un tribunale di Atene.

Andiamo adesso al problema della *causalità animal/corpo*. Esiste un vero problema causale in questo campo, da affrontare filosoficamente. Mi rifiuto, invece, di riconoscere anche qui un reale problema riguardo alla "conservazione dell'energia". È uno pseudo-problema, nato da una confusione concettuale in rapporto alla comprensione della causalità.

La scelta volontaria di muovere una parte del corpo, secondo San Tommaso, comporta l'esercizio di una vera causalità non fisica dell'anima sul corpo (l'Aquinate parlava dell'*usus* o applicazione alle forze naturali del corpo, o dell'*impero* o comando volontario). Anzi, tale causalità, per San Tommaso, è la più alta nell'universo, poiché gli *agenti liberi* sono i "veri agenti" efficienti del cosmo, cioè sono coloro che realizzano in un modo più alto la nozione analogica di causalità. Ma la causalità volontaria, o dell'anima intellettuale sul corpo, *non è fisica*: non è meccanica, chimica, elettrica, ecc., pur essendo una causalità del tutto naturale (nel senso di non miracolosa o

straordinaria). Non comporta una rara eccezione dei principi fisici, in quanto è una forma superiore di causalità che domina meglio le forze fisiche, nel rispetto della loro autonomia e leggi proprie. La costruzione di una cattedrale non viola alcuna legge fisica, ma non si spiega (se non parzialmente) secondo alcuna legge della fisica. Si spiega *adeguatamente* solo secondo i principi della causalità intenzionale e volontaria dell'agente razionale o umano.

Quando la volontà o un atto di pensiero muovono *qualche parte del corpo*, secondo canali precisi, quindi in conformità anche con certe leggi dei livelli inferiori, si può dire che tali atti causali spirituali *creano nuova energia*? Non mi pare, ma dico lo stesso che ho detto sopra riguardo ai miracoli (solo che in questo caso la causalità in gioco è naturale, non miracolosa). Se tali atti dovessero creare nuova energia, ciò sarebbe irrilevante e inverificabile. Che cosa si può invece osservare dal punto di vista empirico? Semplicemente una *nuova distribuzione di energia*, per cui mi pare molto ragionevole dire che la causalità spirituale sul corpo *utilizza e non crea* un'energia fisica preesistente (in mancanza di tale energia, la causalità spirituale sarà inefficace). Sono anche d'accordo nel dire che la causalità mentale aggiunge *informazione* agli organi che possono riceverla, e che la mente può anch'essa *astrarre* l'informazione contenuta negli organi sensoriali (cervello).

Naturalmente il fisico, *come fisico*, non può sapere che la causa del movimento fisico che egli sta eventualmente osservando (ad esempio, una determinata attivazione del cervello) è nella sua radice spirituale. Egli ne osserva la parte fisica. Forse egli potrà supporre che tale modifica fisica dovrà avere certi *antecedenti fisici* (ad esempio, determinati stimoli). Questo è sempre vero, *ma solo parzialmente*. Soltanto impiegando un dogma determinista *a priori* egli potrà sostenere che in quegli antecedenti fisici bisognerebbe trovare la *causa completa adeguata* delle azioni umane in quanto umane (ad esempio, la creazione di un'opera letteraria).

### **3. Le indagini sperimentali moderne e il loro ruolo nel problema mente-corpo**

Dopo Cartesio e le susseguenti filosofie empiriste e razionaliste, non è facile tracciare un quadro lineare in relazione al problema mente-corpo. Come è avvenuto in tante altre tematiche, la questione può essere seguita nel campo strettamente filosofico, ma adesso viene pure affrontata anche sul terreno scientifico, includendo spesso qualche rapporto con la filosofia, almeno a livello ideologico. Nell'epoca di Cartesio la filosofia e la medicina erano scienze chiaramente diverse, ma non esisteva ancora la psicologia sperimentale.

Sul versante filosofico, dopo il cartesianesimo l'approccio al soggetto umano non fa più parte di un programma di filosofia naturale che culminerebbe nell'uomo, secondo il disegno aristotelico. La metodologia cartesiana dell'oggettivismo rappresentazionista provoca la scomparsa della dimensione metafisica nell'empirismo. L'io in *Hume* appare solo come il soggetto delle rappresentazioni in flusso, un soggetto insostenibile perché non cade sotto il principio d'intelligibilità del rappresentazionismo. Lo spirito pensante cartesiano così si dissolve in una serie di rappresentazioni più o meno associate secondo leggi psicologiche puramente empiriche. L'io è un fascio di sensazioni, la cui unità non è neanche assolutamente garantita attraverso la memoria. L'impostazione humanea è psicologista, ma pone in crisi il concetto anche psicologico di identità personale. Le nozioni di persona, soggetto, anima, io, non si giustificano filosoficamente, anche se possono essere utilizzate, almeno alcune, nella vita pratica. In ciò che resta della mente cartesiana rimane pur sempre uno sfondo di visione gnoseologista esasperata, con la riduzione della conoscenza a coscienza indubitabile. Ma adesso tale riduzione porta all'atomismo istantaneo della psiche (soltanto merita certezza ogni stato puntuale e presente della coscienza) (cfr. S. Scuyper, *Hacia una comprensión no atomista de la identidad personal*, en "Anuario filosófico" (Pamplona), XXVI/2, 1993, pp. 223-248).

L'io in *Kant* verrà salvato dalla dissoluzione empirista solo a titolo di funzione trascendentale di ultima istanza, il che lo separa del tutto da un suo possibile ruolo psicologico. L'io kantiano è il *cogito* cartesiano ridotto a funzione sintetizzante di carattere gnoseologico. Ma stranamente, l'io rinasce sotto un'altra forma in Kant, non come io=coscienza pensante, bensì come *persona* nel senso di soggetto di moralità e di Diritto (ma anche in Hume, più psicologicamente, l'io dissolto dall'analisi filosofica ricompariva come l'oggetto delle passioni quali l'orgoglio, la vanità o l'autostima).

Mentre l'io dei filosofi prosegue la sua strada nella storia della filosofia moderna verso la sua esaltazione idealistica (io puramente mentale, cosciente, non corporeo, attivo e creatore nell'idealismo, infinito nel romanticismo, fondante delle rappresentazioni intenzionali nella fenomenologia di Husserl), nel XIX secolo prendono forza le ricerche *neurologiche* e si sviluppa a poco a poco la *psicologia sperimentale* (cfr. R. Luccio, *La psicologia: un profilo storico*, Laterza, Roma-Bari 2000; *Storia delle scienze*, vol. 2, a cura di E. Agazzi, Città Nuova, Roma 1984). Spesso i due campi s'intrecciano nel campo della fisiologia del sistema nervoso.

Sin dall'epoca del Rinascimento e nei secoli XVII e XVIII si svilupparono le ricerche anatomiche sul cervello e sul sistema nervoso, conosciute anche da Cartesio. Nel XIX secolo si apre

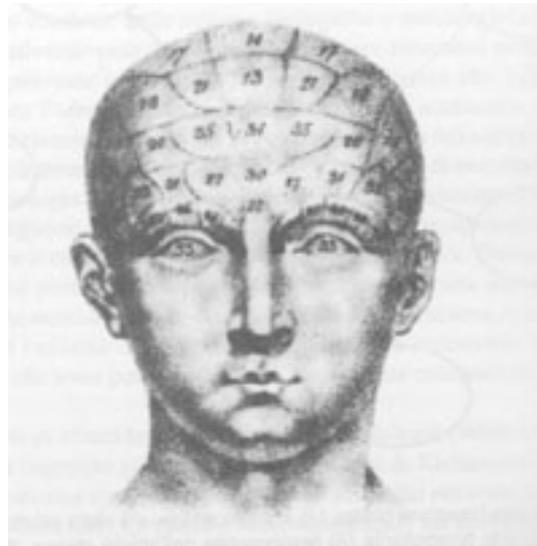
la strada delle indagini neurofisiologiche. L'ambiente scientifico di questo periodo, specialmente nelle aree della biologia, medicina, psicologia e psichiatria, riceve l'impronta degli orientamenti filosofici dell'illuminismo e più tardi del positivismo e naturalismo. È in tale cornice che bisogna vedere certe polemiche in parte metodologiche, ma che anche versano sulla concezione di fondo della vita e dell'uomo (ad esempio, in biologia, tra vitalisti e meccanicisti, e in neurofisiologia tra coloro che sostengono l'autonomia del pensiero e quelli che tendono a ridurlo all'attività cerebrale). All'intreccio tra scienza e filosofia, in questi campi, si aggiunge l'intervento di posizioni ideologiche, che in un certo senso sono "cattiva filosofia", nonché arma di combattimento per la difesa delle proprie idee scientifiche o filosofiche (pure oggi vi sono atteggiamenti ideologici nel modo di impostare la storia della scienza e d'interpretare gli eventi).

Ma vediamo alcune indicazioni concrete sulle indagini e scoperte neurologiche della *prima metà XIX secolo*. Prima dell'Ottocento, il sistema nervoso era visto come costituito da un tessuto indifferenziato, incapace di funzionalità specifiche. Charles Bell (medico scozzese) nel 1811 e François Magendie (medico francese) nel 1822 scoprirono l'indipendenza delle vie *sensoriali* e *motorie* dei nervi periferici ("legge di Bell-Magendie"). Per la prima volta si svelava un'attività non indifferenziata del sistema nervoso, aprendo la via alla ricerca di nuove funzioni specifiche. Altre ricerche, sin dalla fine del Settecento o gli inizi dell'Ottocento (Whyatt, Hall), indicarono nel midollo (con indipendenza dal cervello) l'origine dei movimenti riflessi non volontari. Verso la metà del XIX secolo Claude Bernard (1813-1878) scoprì l'esistenza di un duplice sistemi di nervi (simpatico e parasimpatico), con funzioni di stimolazione e di inibizione.

Contemporaneamente, Johannes Peter Müller (1801-1858) propose la *legge delle energie specifiche sensoriali*, secondo cui l'organo di senso reagisce in modo specifico di fronte a stimoli diversi (meccanico, luminoso, elettrico, chimico, termico), mentre lo stesso stimolo provoca in ogni senso una risposta propria. Questa tesi fu ulteriormente approfondita da von Helmholtz. In altre parole, la natura degli impulsi trasmessi dai nervi ai centri nervosi è propria e relativamente indipendente dalla natura dell'agente. Ciò induceva a pensare a un certo "a priori" dell'organo del senso nei confronti della stimolazione esterna. Müller sosteneva idee vitalistiche, opponendosi alla riduzione alla materia dei fenomeni materiali.

L'ipotesi di una corrispondenza tra funzioni cerebrali e attività mentali fu sostenuta dall'austriaco *Franz Joseph Gall* (1758-1828). Egli credeva a una corrispondenza piuttosto rigida tra le facoltà mentali (di cui faceva una numerosa lista: istinto di propagazione, istinto ferino, istinto di difesa, amore per la prole, orgoglio, amicizia, senso della meccanica, spirito metafisico, talento

poetico, sentimenti religiosi ecc.) e le localizzazioni cerebrali. Le facoltà psicologiche sarebbero responsabili dei tratti comportamentali. Ecco un disegno rappresentativo della concezione galliana:

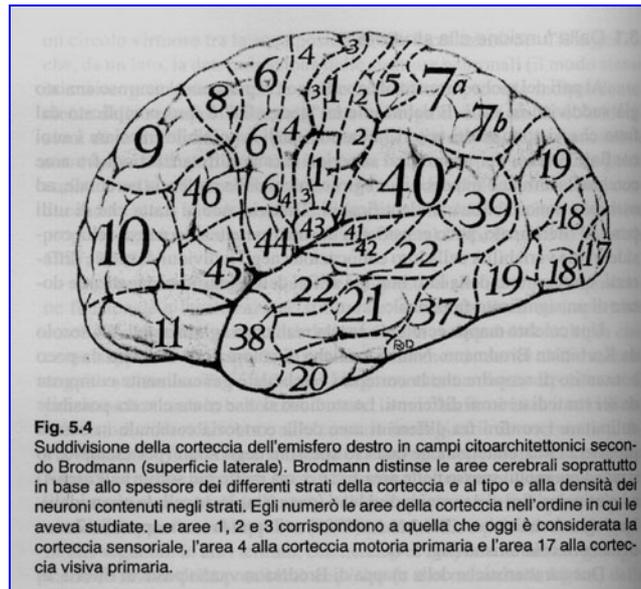


Gall ipotizzò una distribuzione correlativa di aree del cervello, pensando che l'esercizio di alcune delle facoltà dovrebbe comportare una crescita dell'area cerebrale corrispondente, il che premerebbe contro la scatola cranica, deformandola. Di conseguenza, la presenza di particolari asimmetrie nel cranio delle persone rivelerebbe le loro predisposizioni psichiche. La sbagliata teoria di Gall fu chiamata *frenologia* ed esercitò un certo fascino per parecchio tempo nell'opinione pubblica. Uno dei suoi errori era vedere la funzionalità delle aree cerebrali secondo una semplice architettura di aggregazione, senza tener conto dell'organizzazione dell'insieme. Un'altra svista fu la scomposizione cerebrale in funzioni troppo specifiche, invece di pensare ad altre più generali, che potessero stare al servizio di tutte le facoltà (memoria, linguaggio, ecc.).

La frenologia galliana si sviluppò in un'ambientazione materialista (Gall fu cacciato via dalla docenza in Austria e portò le sue teorie a Parigi). Anche se le sue tesi non ebbero il favore degli scienziati, esse sono in sintonia con il naturalismo dell'Ottocento. Contro Gall si alzò Marie-Jean-Pierre Flourens (1794-1867), il quale considerava gli emisferi cerebrali come una struttura funzionalmente unitaria, sede della mente, pure unitaria con le sue funzioni superiori di intelligenza, volontà, percezione e memoria. Le idee di Flourens erano legate al dualismo cartesiano. Ancora non era giunto il momento di scoprire veramente delle funzionalità specifiche nel cervello, come invece si veniva facendo nelle strutture nervose ad esso inferiori (cervelletto, midollo allungato, midollo spinale). Prima del 1870 le aree nervose inferiori, relative al controllo delle sensazioni, movimenti e funzioni vitali, erano tenute separate dalle aree cerebrali, poco conosciute.

Ormai nella *seconda metà del XIX secolo*, E. du Bois-Reymond (1818-1896) dimostrò -negli anni 1848-49 che l'eccitazione nervosa era un'onda elettrica, rivelata dal galvanometro, con cui si poteva finalmente superare ogni resto dell'antica teoria degli spiriti animali. A sua volta, nel 1850 Hermann von Helmholtz (1821-1894), fisico termodinamico ma anche fisiologo, dimostrò sperimentalmente che l'impulso nervoso viaggiava a una velocità finita e misurabile (compresa tra 42 e 25 metri al secondo, cioè molto più lenta rispetto a quella della luce e del suono), contro le tesi vitalistiche contrarie. Lo studioso tedesco contribuì pure agli studi psicofisici con le sue ricerche sulle basi fisiologiche delle sensazioni. Egli scoprì che la percezione del colore nasceva dall'esistenza nella retina di tre tipi di coni (cellule sensibili alla lunghezza d'onda della luce), e compì studi sulla percezione di altezze tonali nell'udito.

Nel 1870 due scienziati tedeschi, Gustav Theodor Fritsch (1838-1927) e Julius Eduard Hitzig (1838-1907) rivelano per la prima volta l'eccitabilità elettrica della corteccia cerebrale. Provocando tali eccitazioni, riuscirono a produrre delle contrazioni muscolari, arrivando così a localizzare le prime aree motorie corticali. Continuò tali ricerche David Ferrier (inglese, 1843-1928), il quale rilevò con la tecnica della stimolazione elettrica numerose aree motorie e le cinque aree sensitive. La legge di Bell-Magendie si estendeva quindi alla corteccia. La sua opera principale, *The functions of the brain* (1876) significa il passaggio a una nuova epoca negli studi del sistema nervoso. Altresì importante fu la scoperta dell'unità morfologica e fisiologica delle cellule nervose (neuroni) agli inizi del XX secolo (Santiago Ramón y Cajal, 1852-1934, collegato alle ricerche di Camillo Golgi, 1844-1926). Posteriormente iniziò la ricerca sui collegamenti sinaptici, nonché quella relativa alla descrizione anatomica del cervello tenendo conto della sua struttura citologica. Agli inizi del XX secolo Korbinian Brodmann elaborò una "mappa" cerebrale, indicando 47 aree della corteccia umana.



Le indagini sulle localizzazioni cerebrali di funzioni psichiche ricevettero un notevole successo con la scoperta compiuta da *Pierre-Paul Broca* (1824-1880) nel 1861 del centro cerebrale del linguaggio articolato (“area di Broca”). Il chirurgo francese ottenne tale risultato grazie allo studio di pazienti soggetti ad afasia motoria (deficit nella produzione del linguaggio), a causa di lesioni in quell’area dell’emisfero cerebrale sinistro. Ugualmente *Karl Wernicke*, neurologo e psichiatra tedesco (1848-1905) localizzò nel 1876 il centro (“area di Wernicke”) la cui lesione è responsabile dei deficit di comprensione linguistica (afasia sensoriale). Quindi le fantasiose teorie di Gall non riuscirono a fermare la neurofisiologia della corteccia, che vedrà nel XX secolo numerose scoperte. Broca e Wernicke inaugurarono lo studio delle “localizzazioni”, anche se ovviamente lo studio del cervello non si riduce a questo punto.

Marginalmente vorrei menzionare pure in questo periodo la nascita dell’*antropologia*, concepita come una *scienza naturale*, con obiettivi quali lo studio delle caratteristiche morfologiche dei tipi umani e della diverse etnie o razze, in rapporto a tempi, distribuzione geografica e fattori come il clima, il nutrimento e i costumi. Si assegnò molta importanza in questo periodo alla descrizione e misurazione del cranio (craniometria). Broca, che era pure un antropologo, studioso della morfologia umana e delle razze, fondò nel 1850 la Società di Antropologia di Parigi, costituita da medici e naturalisti (queste società proliferarono in Europa rapidamente, estendendosi a Londra, Madrid, Berlino, Mosca, Firenze).

Gli studi di antropologia razziale e culturale, improntati da metodi naturalisti, comportano una certa prospettiva del problema mente-corpo. Talvolta furono collegati con gli studi sull’evoluzione nel senso darwiniano, studi contemporanei a quanto veniamo esponendo. La mentalità umana, in

questa visione, si ritiene plasmata da fattori fisici genetici e ambientali. In tale contesto si possono capire meglio, ad esempio, le tesi di C. Lombroso (1835-1909), psichiatra e antropologo e positivista giuridico, sul rapporto tra costituzione fisica e tendenze criminali delle persone.

Le ricerche antropologiche sulle razze furono di moda in quel periodo e purtroppo costituirono un supporto per la nascita delle ideologie razziste, con le loro note derivazioni politiche. È illustrativo, in questo senso, consultare le voci *Antropologia* e *Razza* nell'Enciclopedia Italiana Treccani, pubblicata nel periodo fascista. Queste voci, firmate da Gioacchino Sera, professore di antropologia presso l'Università di Napoli, sono molto ben elaborate dal punto di vista scientifico, ma non possono occultare una certa sudditanza, sia pure equilibrata, verso una visione dove la razza acquista un'importanza notevole, proprio per superare quello che viene chiamato il dogma della rivoluzione francese di una "generica e astratta eguaglianza degli uomini" (voce *Antropologia*, vol. 3, 1929). Colpisce la parte dedicata alle differenze psichiche delle razze (voce *Razza*, vol. 28, 1925-26). L'autore fa notare il grande sviluppo in Germania della letteratura scientifica sulle razze, dove nota con un certo disagio il grande rilievo che si dà alla razza nordica. Menziona l'antropologia sociale di Fischer, in Germania, che studia le relazioni dei caratteri antropologici con i gruppi sociali, e osserva che tale movimento, improntato dall'idea della necessità di una "igiene razziale" scientificamente fondata, è collegato al pangermanismo nazionale, anche se le sue idee sono dissimulate dall'apparenza di oggettività scientifica.

Ho fatto questa parentesi al tema specifico per far vedere fino a che punto gli studi scientifici, se impostati in un certo modo, possono diventare stimolo per posizioni ideologiche, non solo per imposizione dell'autorità, ma anche in un modo spontaneo e quasi "innocente". Questo punto oggi lo vediamo chiaro, nell'esempio citato, solo perché abbiamo un consenso sul fatto che il nazismo è indiscutibilmente perverso. Ma quando manca il consenso intorno a certi valori etici, non ci si accorge così facilmente fino a che punto è operante l'intreccio tra scienza e ideologia, nel contesto dell'interazione tra cultura, valori e pensiero scientifico.

Per finire il quadro di alcuni sviluppi neurologici rilevanti per la psicologia, aggiungerò tre punti (cfr. Michel Simon, ed., *La peau de l'âme*, ed. du Cerf, Parigi 1994; Massimo Marraffa, *Scienza cognitiva. Un'introduzione filosofica*, Cleup, Padova 2002, pp. 165-180):

a. La neuropsicologia della seconda metà del XIX secolo preparò il rifiuto dell'equazione *psichismo=coscienza*. I movimenti riflessi -inconsci- dipendono dai centri nervosi inferiori, ma poteva avvenire lo stesso nei centri superiori. Già alla fine del XIX secolo si comincia a parlare di inconscio. La stessa difficoltà nell'eseguire e interpretare i metodi introspettivi con interrogazioni e

prove rigorose, per renderli il più oggettivi possibili, dimostrava che non tutto era cosciente nei meccanismi psichici umani.

b. Al modello elettrico del cervello, dopo gli anni '20 del XX secolo, seguirà un modello complementare di tipo *chimico*. Otto Loewi scoprirà in questo periodo il ruolo delle sostanze chimiche sul cervello. Nasce così la neurochimica, da cui dipendono la psicofarmacologia e la chemioterapia (il fatto era noto in un certo senso dagli antichi, poiché si conoscevano gli effetti di certe sostanze -droghe- sugli stati di coscienza). Posteriormente verranno scoperti i neurotrasmettitori, che rinnoveranno la psichiatria. Alcune disfunzioni mentali (insonnia, ansietà, aggressività, anomalie della sessualità) possono essere causate da una cattiva regolazione delle endosostanze e quindi possono essere trattate farmacologicamente.

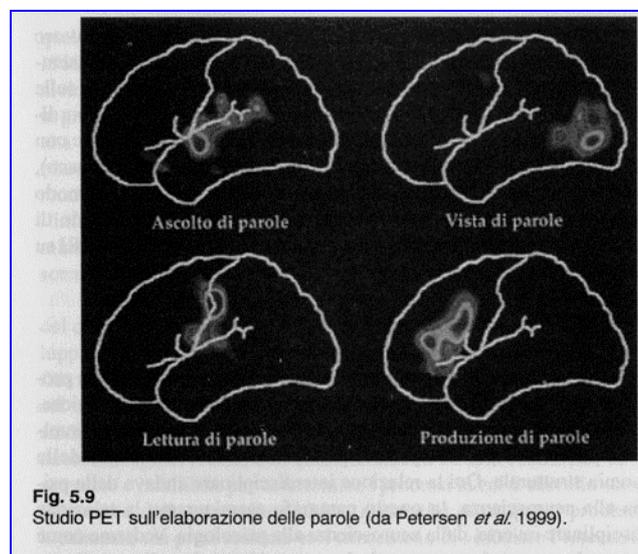
Nascono così due correnti nel trattamento di alcune malattie nervose che hanno a che vedere con il problema mente-corpo: l'indirizzo *farmacologico* e quello *psicologico* (ricordiamoci qui del problema cartesiano sul dominio delle passioni). A causa del rapporto vicendevole tra la parte fisica e psichica, le disfunzioni fisiche non solo perturbano la vita psichica, ma anche viceversa. Le depressioni, i cambiamenti di umore, producono sintomi digestivi, endocrini, alterazioni del ritmo cardiaco e della pressione arteriale, problemi di sonno, ecc. Non sempre il trattamento farmacologico va alla radice. Per certe malattie, si può pensare alla necessità di accompagnare la cura medica con quella psicologica. Ma neanche la psicologia è tutto. Essa non è competente per la questione del senso della vita e per problematiche profonde dell'uomo, di tipo etico, antropologico e religioso. La psicologia non può entrare nel campo oggettivo di queste sfere, in quanto per natura si deve limitare al lato "soggettivo-formale" della vita umana (funzioni psichiche). Ma ciò non toglie la sua validità. Il comportamento etico e religioso può essere anche vissuto in un modo psicologicamente patologico (ad esempio, esasperazioni, complessi di colpa, misticismi patologici).

c. La ricerca sulle localizzazioni cerebrali (aree funzionali) e sulle vie nervose afferenti, efferenti e associative (visiva, uditiva, ecc.) continuò con grande ampiezza nel XX secolo. Sono state usate a questo scopo diverse tecniche osservative, applicate ad animali o uomini, come lo studio degli effetti di lesioni o ablazioni o delle correlazioni tra danni cerebrali e deficit psicologici, le stimolazioni elettriche, fino ad altre più recenti e meno invasive, come la tomografia a emissioni di positroni (PET) e la risonanza magnetica funzionale (fMRI) (si ottengono così "neuroimmagini" del cervello, da interpretare funzionalmente). In queste indagini si segue una duplice strada, che consiste nell'ipotizzare funzioni o sottofunzioni e cercarne le localizzazioni (*top-down*) o viceversa (*bottom-up*), il che dà luogo spesso ad una convergenza tra la psicologia (che preferisce la prima

strada, specialmente la psicologia cognitiva, col ricorso a esperimenti o test cronometrici nell'esecuzione di compiti cognitivi) e la neurologia (che invece predilige il secondo metodo).

D'altra parte, la scoperta di localizzazioni non porta semplicemente ad evidenziare funzioni immediate (come pensava Gall), il che è troppo semplicistico. Contano pure le connessioni e i rapporti interattivi tra le aree funzionalmente rilevanti, aree in cui molte funzioni sono smembrate o distribuite. Ci sono funzioni fisiologiche, regolative, ormonali, sensoriali, motorie, ecc. Per la nostra tematica interessano quelle psicologiche, ad esempio di tipo cognitivo o emotivo.

A titolo di esempio, ecco una figura relativa ad uno studio condotto col metodo PET (1989), che rileva aree corticali attivate per l'ascolto di parole, la visione di parole, la loro lettura passiva e la produzione attiva di verbi correlati:



Nella seconda metà del XX secolo vi sono state indagini approfondite sulle asimmetrie funzionali degli emisferi cerebrali. Negli anni '60 Roger W. Sperry studiò la specializzazione funzionale emisferica mediante ricerche sul "cervello diviso" (*split brain*) di alcuni pazienti. Vengono pure testati soggetti normali tramite la presentazione tachistoscopica di stimoli visivi in uno dei due emicampi (analogamente per l'udito). La lateralizzazione linguistica dell'emisfero sinistro (o comunque di uno dei due emisferi) non è assoluta. Per un po' di tempo è andata di moda l'idea che l'emisfero sinistro sarebbe la sede di processamenti più analitici degli stimoli (linguistici, aritmetici), mentre quello destro sarebbe invece sede di processamenti più sintetici od "olistici" (collegati alle capacità artistiche e creative). Esiste certamente una "dominanza" emisferica per alcune funzioni, ma non va esagerata né semplificata. Inoltre il cervello è molto adattabile e agisce con un coinvolgimento totale nei suoi diversi compiti cognitivi, con una incredibile complessità che ancora non conosciamo bene.

d. Dopo la seconda guerra mondiale si è verificata una nuova spinta negli studi del sistema nervoso, contemporanea alla nascita dell'orientamento cognitivo nel campo psicologico. Una caratteristica di queste due linee scientifiche è il loro intreccio quasi connaturale. Si viene a delineare così, nella seconda metà del XX secolo, l'ambito delle *scienze cognitive*, nel quale diverse discipline, come la psicologia, la neurofisiologia, la psichiatria, la linguistica e l'informatica si collegano a vicenda nello studio degli eventi cognitivi dell'uomo, apportando ciascuna la propria prospettiva. Su questa piattaforma si alzeranno le nuove interpretazioni filosofiche sul rapporto mente-corpo.

La *neuropsicologia cognitiva* (da non confondere con la psicofisiologia) non sta solo attenta alle basi cerebrali dei processi mentali, secondo l'approccio "localizzazionista", basato specialmente sullo studio di cerebrolesi, ma s'ispira anche ai modelli cognitivi teorici (struttura e componenti, interrelazioni, integrazione con altri processi mentali) e mira alla conoscenza dei processi mentali normali. Si utilizzano diversi metodi, come lo studio sperimentale -tramite test- di gruppi di soggetti, con l'impiego di statistiche, anche casi singoli (malati), nonché tecniche psicofisiologiche e di ottenimento di neuroimmagini (cfr. Maria Pia Viggiano, *Introduzione alla psicologia cognitiva*, Laterza, Roma-Bari 1995, pp. 147-173).

Donald Hebb, nella sua opera *Organization of Behavior* (1949), propose l'esistenza di gruppi di neuroni sensori e motori che operano unitariamente formando "assembramenti cellulari" o circuiti neuronali che mantengono l'eccitazione in un modo circolare e distribuito lungo il cervello (s'ispirò alla teoria precedente di Lorente de Nó sui "circuiti reverberanti"). Tali assembramenti o "reti neurali" sarebbero i correlati neurali dei concetti. Assistiamo così alla nascita della teoria "connessionistica", che avrebbe avuto notevoli sviluppi posteriormente.

Questi fenomeni dimostrano che il sistema nervoso non è soltanto canale di trasmissione, ma anche di elaborazione dell'informazione ricevuta. Ci sono meccanismi di attivazione o di inibizione delle cellule nervose, da cui seguono diversi modi di trattare l'informazione trasmessa all'interno delle vie nervose (selezione, integrazione). Concretamente, si scoprì che la sostanza reticolare del tronco cerebrale distribuisce diffusamente gli impulsi ricevuti su ampie zone della corteccia, conservando in essa un tono di attività di fondo e così mantenendola in stato di vigilanza (studi di G. Moruzzi e Magoun del 1945). Dalla sua attività dipende quindi lo stato di veglia o di torpore, il che è legato al fenomeno dell'attenzione e al rendimento percettivo discriminatorio.

Questi sono soltanto alcuni esempi di campi di ricerca neurologica contemporanea. In realtà le scoperte e le indagini sono sterminate in numerosi altri campi (linguaggio, motricità, sensazioni e

percezioni, memoria, emotività, psicopatologia), e ancora abbiamo la sensazione di non conoscere troppo bene il funzionamento del cervello.

Vediamo adesso alcune linee fondamentali sullo sviluppo storico della *psicologia scientifica* (sperimentale). Le sue prime origini sono collegate agli sviluppi della fisiologia del sistema nervoso. Possono essere considerati fondatori della psicologia sperimentale Ernst Heinrich *Weber* (1795-1878), Gustav Theodor *Fechner* (1801-1887), e Wilhelm *Wundt* (1832-1920). In Weber e Fechner l'approccio psicologico non è più filosofico, ma naturalistico e sperimentale, concentrandosi sullo studio dei fenomeni percettivi, col tentativo di misurarne gli elementi quantitativi. Fechner scoprì, mediante esperimenti di laboratorio, che la grandezza soggettiva ("mentale") di una sensazione è proporzionale al logaritmo della grandezza fisica del suo stimolo ("fisico") (legge di Weber-Fechner, posteriormente riveduta e migliorata da altri autori). Siamo alla base della psicofisica (criticata invece da Bergson come tentativo inadeguato di matematizzare i fenomeni della coscienza), elaborata col metodo della misurazione dei tempi come criterio di rilevamento degli atti psichici e delle loro differenze. Tale scoperta e altri studi furono interpretati da Fechner in una linea parallelista di stile spinoziano. La differenza tra corpo ed anima sarebbe una differenza di modi di vedere la stessa realtà, così come lo stesso oggetto può essere visto da prospettive diverse.

Forse l'anno simbolico della fondazione della psicologia sperimentale è il 1879, quando Wundt istituì a Lipsia un laboratorio di psicologia accanto a un'università. Wundt approfondì la psicofisiologia dei sensi (vista e udito in particolare) e continuò la linea fechneriana basata sul metodo di studio dei tempi di reazione, che si rivelò più complessa del previsto. Egli era convinto dell'importanza del metodo introspettivo, cercando però di adottare delle cautele allo scopo di arrivare ad un'adeguata oggettività tramite i metodi sperimentali. Filosoficamente condivideva il parallelismo psicofisico di Fechner. Un suo allievo, Edward Titchener (1867-1927) portò con successo la scuola wundtiana negli Stati Uniti (chiamata "strutturalismo").

Verso la fine dell'Ottocento e poi i primi decenni del Novecento, l'indirizzo psicofisiologico wundtiano venne contestato in Europa dalla psicoanalisi e da un gruppo di psicologi ispirati a Brentano, cui seguirà con grande successo la psicologia della *Gestalt*. Negli Stati Uniti sarà importante l'influsso esercitato da James e dalla psicologia funzionalistica, anche se presto dominerà il comportamentismo (per questi punti, cfr. R. Luccio, *La psicologia: un profilo storico*, cit., pp. 89-130). Dovremo soffermarci su quest'ultimo, a causa del suo ruolo teoretico nella storia della filosofia della mente.

A. *Scuole psicoanalitiche*. Sigmund Freud (1856-1939) è il creatore della psicologia dell'inconscio. Egli cominciò le sue ricerche nei campi psicofisiologico e neuropatologico, passando dagli studi dei metodi ipnotici e dell'isterismo (Charcot) all'impostazione psicoanalitica. La scoperta della sfera psichica inconscia supponeva una nuova concezione strutturale della psiche, in cui molte forze ordinariamente unite potevano essere disgregate. L'io cosciente non risultava pienamente signore di se stesso, poiché vi erano in gioco altre forze oscure negli strati inconsci.

La scuola psicoanalitica ebbe numerosi sviluppi (Adler, Jung). Una sua derivazione poco nota è la logoterapia di Viktor Frankl, più aperta alle dimensioni più alte e complete della personalità umana. In una prospettiva postmoderna, Jacques Lacan reimpostò la teoria freudiana utilizzando elementi dello strutturalismo, per arrivare alla dissoluzione totale dell'io e della coscienza, molto al di là di Freud (da notare la vicinanza tra Lacan e Derrida).

Col passare degli anni, alla scuola freudiana venne rimproverata la mancanza di un serio metodo scientifico e gli eccessi di una visione "superinterpretativa" di tutte le sfere della vita, secondo schemi antropologici prefissati (il terapeuta è un ermeneuta di tutti gli ambiti coscienti cui il paziente ha un accesso diretto). All'inizio di questo corso abbiamo visto il medesimo difetto nelle attuali critiche alla "psicologia popolare" da parte della neurofisiologia riduttivista (coniugi Churchland).

B. *Brentaniani e pregestaltisti*. Questo nuovo gruppo di psicologi manifestò un'insofferenza nei confronti dell'impostazione puramente psicofisica, troppo associazionistica e ristretta allo studio di fenomeni elementari della psiche. Le note critiche di Husserl alla psicofisica del suo tempo sono un'eco di questa nuova impostazione, pienamente psicologica. Franz Brentano (1838-1917) pubblica nel 1874 la sua *Psicologia da un punto di vista empirico*. Com'è noto, egli caratterizzò l'atto psichico come intrinsecamente intenzionale, cioè indirizzato a un oggetto interiore o rappresentazione (la sua linea fu continuata dal suo allievo Meinong). In questo modo si ponevano le premesse della fenomenologia, la quale comporta una reazione oggettivista di fronte allo psicologismo (Popper direbbe: la riduzione del mondo 3 al mondo 2).

La nuova linea antiwundtiana è rappresentata in primo luogo dalla scuola di Graz, fondata da Meinong. Un suo allievo, von Ehrenfels, può essere considerato il precursore della *Gestaltpsychologie*. Egli coniò l'espressione "qualità gestaltica", ad esempio la melodia colta in un brano musicale, o la triangolarità, che si mantiene ugualmente percepita in mezzo a tutte le variabili figure triangolari che vengono viste. Tra gli allievi di Wundt, Otto Külpe (1862-1915), fondatore della scuola di Würzburg, prese pure un orientamento più globalista. Egli fece indagini

sull'esistenza di "pensieri senza immagini" (attesa, finalità, sicurezza, intenzionalità) con esperimenti guidati dal metodo introspettivo. Queste e altre scuole analoghe (ad esempio Gemelli in Italia) dominano nella prima metà del Novecento.

C. *Funzionalismo americano*. Sin dalla fine dell'Ottocento fino ai primi decenni del Novecento nasce negli Stati Uniti, come reazione contro la psicologia wundtiana di Titchener, un nuovo orientamento che venne denominato "funzionalismo". Sotto l'influsso delle teorie evoluzionistiche, iniziano gli studi di psicologia animale e di psicologia dell'età evolutiva. William James (1842-1910), sebbene non sia un funzionalista in senso stretto, sviluppò un nuovo modo di fare psicologia (*Principi di psicologia*, 1890). La coscienza non va vista, secondo James, come qualcosa di statico, colta da atti singoli di introspezione, bensì come un processo continuo e globale, dove conta molto l'esperienza e la memoria (una tesi simile si trova in Bergson). I funzionalisti, collegati alla filosofia pragmatista, svilupparono l'uso di test e cominciarono a studiare la psicologia applicata, ad esempio nell'area del lavoro, l'arte, la storia, e studiarono pure l'adattamento dell'individuo all'ambiente.

D. *Gestaltpsychologie*. La psicologia della *Gestalt* (forma, *pattern*) si concentrò specialmente sui fenomeni percettivi. Nacque per il pubblico nel 1912, quando Max Wertheimer (ebreo di Praga, 1880-1943) pubblicò i risultati delle sue ricerche nell'istituto di psicologia di Francoforte, condotte insieme a Wolfgang Köhler (1887-1967) e Kurt Koffka (1886-1941). Un suo antecedente è Carl Stumpf, maestro di Wertheimer (e anche di Husserl, cui guidò nella preparazione del suo saggio nel 1891 sulla filosofia dell'aritmetica). La scuola gestaltistica trovò un grande ascolto nel mondo intellettuale tedesco. La sua impostazione non psicofisica ma solo "fenomenalistica" la rese affine alle "scienze dello spirito" predominanti nel culturalismo tedesco dei primi decenni del Novecento. Con l'avvento del nazismo, molti gestaltisti emigrarono negli Stati Uniti, dove comunque non trovarono un grande ascolto a causa della loro metodologia poco sperimentale e poco funzionalistica (tranne Kurt Lewin, 1890-1947, promotore della psicologia sociale). Ma oggi le scienze cognitive sono nuovamente interessate alle teorie della *Gestalt*.

Il punto centrale della scuola della Gestalt è la scoperta del carattere originario della percezione di strutture organizzate rispetto agli elementi componenti (anti-elementarismo). Gli stessi elementi talvolta possono essere percepiti come strutture diverse. Wertheimer ad esempio studiò il movimento apparente "stroboscopico", che sta alla base della percezione del movimento cinematografico. Lo spettatore unifica nella sua percezione i fotogrammi con un senso di movimento continuo, invece di vedere immagini statiche, il che non ha una corrispondenza con le

stimolazioni retiniche. Certi *patterns* sono *pregnanti*, cioè si tende spontaneamente a percepirli nonostante le irregolarità (legge della pregnanza).

Ecco un esempio molto noto di percezione gestaltica di una figura ambigua. La stessa stimolazione retinica può essere vista come due figure (donna anziana, donna giovane). Se siamo abituati, noi stessi possiamo decidere di scambiare i *pattern* percepiti:



I gestaltisti si sono preoccupati pure del rapporto tra percezione e cervello, sostenendo una teoria denominata *isomorfismo* (Köhler, Wertheimer). Köhler sosteneva l'esistenza di un'identità strutturale tra eventi psicologici ed eventi del sistema nervoso centrale. Alle forme psicologiche dovrebbero corrispondere "forme fisiologiche", cioè totalità strutturali fisiologiche a livello corticale che risponderebbero come tali alle eccitazioni periferiche. Wertheimer parlava di un "cortocircuito cerebrale" in tal senso. Alla dominanza della forma sugli elementi sul piano percettivo corrisponde, secondo Köhler, una correlativa dominanza della forma nel mondo fisico e neurologico. Quindi, anziché pensare a messaggi periferici che si trasmettono linearmente fino ai centri delle localizzazioni cerebrali, bisognerebbe puntare a risposte globali del cervello. In questo senso i gestaltisti si sono rifatti alla nozione fisica di *campo*. Il sistema percettivo va visto quindi come un sistema fisico che tende all'equilibrio secondo il "principio di minimo". Il campo è un'unità di forze interagenti, dove ogni oggetto introdotto modifica l'equilibrio delle forze. Il comportamento percettivo si gioca a livello di totalità, con tendenza all'autoregolazione.

L'isomorfismo gestaltico comporta una posizione intorno al problema mente-corpo che tende al monismo parallelista. Le forme percettive, per Köhler, sarebbero la versione "fenomenale" di forme fisiologiche soggiacenti. Ma gli autori non sono stati più espliciti al riguardo.

Cornelio Fabro si è occupato con una notevole accuratezza della teoria della Gestalt nella sua opera di gioventù *Fenomenologia della Percezione*, pubblicata per la prima volta nel 1941, con una prefazione di Agostino Gemelli (Vita e pensiero, Milano; ved. una seconda edizione riveduta nella Morcelliana, Brescia 1961). L'opera è continuata in *Percezione e Pensiero*, pubblicata quasi simultaneamente in entrambe le edizioni (cfr. i nostri commenti a queste opere in *Ermeneutica dell'apprendimento percettivo*, "Euntes Docete", 50, 1997, pp. 195-212). Sia Fabro che Gemelli, nel campo del tomismo, hanno dimostrato un interesse scientifico per gli sviluppi della psicologia, intervenendo nel dibattito in merito alle proprie competenze (Gemelli guidò per molti anni una scuola di psicologia sperimentale a Milano). Nella prima opera indicata, Fabro loda nella *Gestalttheorie* la rilevanza data alle strutture, convergente con il primato aristotelico della forma, ma non risparmia alcuni apprezzamenti critici.

Sul punto che più ci riguarda (rapporto pensiero-cervello), Fabro manifesta un'insoddisfazione nei confronti della tesi dell'isomorfismo, per cui alla fine propende per l'esistenza di "un'incommensurabilità fra i fenomeni fisiologici ed i contenuti fenomenali" (p. 459 dell'edizione della Morcelliana). In realtà, gli argomenti contrari di Fabro si concentrano sulla nostra ignoranza su quanto avviene nel nostro cervello -in collegamento con i fenomeni psichici- in rapporto allo stato delle ricerche neurologiche nell'epoca in cui egli scriveva (l'isomorfismo dei gestaltisti era un'ipotesi di lavoro). Eppure egli riconosce che una visione aristotelica porta a sostenere una connessione intrinseca tra organo e facoltà, riflesso dell'unione sostanziale fra anima e corpo (cfr. *Percezione e pensiero*, Morcelliana, Brescia 1962, pp. 140-142).

Per quanto riguarda gli atti sensoriali (non quelli spirituali, intellettivi e volontari), propendiamo non per una semplice connessione, ma per una vera *integrazione ilemorfica*. Comunque alcuni gestaltisti manifestarono una concezione materialista della conoscenza (in rapporto al "principio dell'isomorfismo"). Ma non sviluppiamo adesso questo punto.

L'altro elemento critico, rilevato da parecchi autori, come Agostino Gemelli, Janet, Jean Piaget (1896-1980), Michotte, è la mancanza nel gestaltismo della distinzione tra "forma" e "significato", oltre all'emarginazione dei processi genetici degli atti percettivi. Il gestaltismo è rimasto su un piano troppo statico nel suo studio della percezione. Nel suo secondo volume sulla psicologia (*Percezione e pensiero*), Fabro compie un superamento della teoria della *Gestalt* appellandosi alla facoltà tomistica della cogitativa per inquadrare gli elementi "funzionali" della percezione, nei confronti dei significati vissuti dei percetti per la vita del conoscente.

La psicologia della percezione gestaltica venne superata anche da alcuni psicologi americani

degli anni '50, appartenenti al movimento chiamato *New Look*, come Bruner, Postman, Goodman, McGinnies, i quali cominciarono a tener conto degli elementi soggettivi che intervengono nell'atto percettivo -bisogni, motivazioni-, ed esaminarono il dinamismo dei processi percettivi. Il riconoscimento di forme può avvenire attraverso ipotesi percettive che inducono ad ulteriori esplorazioni contestuali, seguite da conferme o smentite. La percezione quindi ha una sua storia e l'esperienza passata orienta verso i nuovi riconoscimenti. Alcune percezioni "subliminari" non superano la soglia della coscienza eppure vengono recepite. La percezione insomma gioca su diverse dimensioni e non può essere ridotta al puro aspetto strutturale.

5. *Psicologie della personalità*. Per amore di completezza, vogliamo menzionare qui l'esistenza di un'altra corrente psicologica, molto attiva negli USA negli anni '30 e '40 (Allport, Kretschmer, Sheldon, Cattell, Odbert). Le psicologie della personalità non analizzano i processi psicologici in modo separato, ma si interessano della descrizione di *tipi* umani e di *tratti* globali della personalità, spesso bipolari (introversione-estroversione, dominanza-sottomissione, conservatorismo-radicalismo, ecc.). I "tipi" continuano in parte la prospettiva tradizionale che, individuando nelle persone un insieme di caratteristiche, parlano di *carattere*, più morale e volitivo, e di *temperamento*, più basato sul rapporto tra aspetti psicologici e fisici (così come la tradizione gallelica si riferiva ai sanguigni, collerici, flemmatici e melanconici). La tipologia di Kretschmer contempla in tal senso le categorie dei ciclotimici, ixtimici e schizotimici. La personologia non segue un modello *nomotetico* (ricerca delle leggi), ma *idiografico* (l'oggetto di studio è l'individuo). Il problema del rapporto tra caratteristiche somatiche e personologiche è collegato alla questione mente-corpo.

#### 4. Il comportamentismo

A. *Comportamentismo psicologico*. Vogliamo soffermarci sul comportamentismo a causa della sua importanza nella filosofia della mente. Possiamo distinguere tra il comportamentismo come movimento psicologico e il comportamentismo filosofico. Cominciamo col primo, che dà piede al secondo. Il precedente del comportamentismo nella psicologia sono gli studi russi sui *riflessi condizionati*, negli ultimi anni del XIX secolo. Questi studi appartengono di per sé alla fisiologia, ma comportano conseguenze psicologiche. L'arco riflesso, con la sua struttura di stimolo-risposta, diventa una nuova unità di analisi per la psicologia.

Ivan Petrovich Pavlov (1849-1936), medico fisiologo, scoprì nel 1905 che la secrezione salivale dei cani non dipendeva soltanto dall'introduzione dell'alimento nella bocca, ma anche dalla presentazione visiva e olfattiva del cibo, anzi dalla comparsa ripetuta di altri stimoli contestuali

associati allo stimolo principale (rumori o altri segni). La risposta riflessa quindi non era solo collegata a stimoli incondizionati (cibo), ma ad altri stimoli associati, prima neutri ma poi efficaci, quindi condizionati. I riflessi potevano essere dunque *condizionati*. Si potevano così studiare la genesi dei riflessi condizionati, i loro rinforzi, differenziazione, associazione ed estinzione. L'idea di Pavlov era di ricostruire su questa base un sistema psicologico completo che comprendesse anche i processi superiori e la personalità.

La psicologia riflessologica offriva la base per una psicologia "oggettiva", affine alle scienze naturali e quindi diversa dalla tradizionale psicologia introspezionistica. Una linea analoga nacque contemporaneamente negli Stati Uniti con Edward Lee Thorndike (1847-1949), il quale elaborò una teoria dell'apprendimento animale in base a processi di tentativi (*trials*) ed errori. Siamo agli albori del behaviorismo, scuola psicologica orientata all'osservazione di schemi di condotta come quelli appena descritti e non all'esplorazione dei fatti interni della coscienza. John B. Watson (1878-1958) è il fondatore di questa corrente, specialmente dopo la pubblicazione nel 1913 dell'articolo *Psychology as the Behaviorist Views It*.

Il comportamentismo si presenta come un programma di ricerca mirante allo studio sperimentale del comportamento osservabile dei soggetti, tralasciando lo studio della coscienza o della mente. *Lo schema comportamentale è costituito da rapporti tra stimoli ambientali -eccitazioni sull'organismo- e risposte dell'organismo*, intese come cambiamenti nella muscolatura e nelle secrezioni ghiandolari (non si puntava alla fisiologia nervosa, probabilmente perché era ancora troppo poco conosciuta). Vengono abbandonati i metodi dell'introspezione e del colloquio personale con i soggetti studiati (il comportamentismo nacque in contesto di psicologia degli animali). Il nuovo metodo *causalistico* è simile a quello delle scienze naturali e quindi poteva soddisfare le esigenze del positivismo e del *fisicalismo*, col quale però non era inizialmente in contatto. L'obiettivo del nuovo programma era cercare delle leggi tra le variabili indipendenti degli stimoli S e le variabili dipendenti delle risposte R, secondo la formula  $R=f(S)$ , cioè *la risposta è funzione dello stimolo*. Il comportamentismo insisteva sulla nozione di *apprendimento* e così eludeva la questione dei meccanismi istintivi.

La corrente comportamentista dominò massicciamente lo scenario psicologico americano durante la prima metà del XX secolo. Ben presto però -sin dagli anni '20- i postulati comportamentisti dovettero modificarsi, man mano che lo schema S-R (stimolo-risposta) diveniva più complesso del previsto. Nacquero così ulteriori posizioni denominate nell'insieme *neocomportamentismo*. Altri autori (Edward Chace Tolman, 1886-1959) fecero intervenire nuove

variabili (“intervenienti”) nello schema descrittivo della condotta (organismo, eredità, età, sesso, abitudini). Tolman parlò di un *purposive behaviorism* (“comportamentismo diretto a un scopo”). Poi si è portati a postulare costrutti ipotetici (accomodamento, rinforzo), collegati al metodo ipotetico deduttivo. Burrhus Frederik Skinner (1904-1990) propose un neocomportamentismo più ortodosso, con una visione più ampia rispetto a Watson, dando una particolare importanza al rinforzo dei condizionamenti (rinforzo positivo e negativo). Dopo gli anni ‘50, nuovi studi prepararono il declino del comportamentismo psicologico. D’una parte, la conferma in base a osservazioni che i comportamenti animali (*etologia*) sono particolari alle specie, il che poneva un limite alle possibilità di apprendimento. Non tutti gli animali possono apprendere le stesse cose. D’altra parte, vi si aggiunse il ruolo del linguaggio come mediatore tra stimoli e risposte.

Si può leggere in Popper, *L’io e il suo cervello*, vol. 1 (cit.), pp. 167-172, una critica alla teoria dei riflessi condizionati, nella prospettiva di una visione più ampia della condotta animale. La spiegazione della condotta dell’animale secondo l’arco riflesso stimolo-risposte, rileva Popper, è troppo semplicistica. L’animale affronta l’ambiente in una maniera più attiva, contando con un “programma di azione” inconscio, che egli mette alla prova nella sua esplorazione dell’ambiente. Questo “programma di azione” è uno di quei “costrutti teorici” rifiutati dai comportamentisti rigorosi, i quali volevano eliminare ogni forma di “teoria” introdotta in mezzo alle regolarità osservate.

In conclusione, il comportamentismo psicologico portò l’attenzione allo studio della condotta esterna dell’animale, vista come reazione fisiologica agli stimoli ambientali, una reazione non innata, ma imparata nell’esperienza. Lo schema era valido per una comprensione “non interioristica” della vita animale, ma era troppo ridotto. Era necessario contare anche con una base innata genetica, e con l’intervento “mediatore” di forme di rappresentazione, di aspettative o di impulsi intenzionali, “scatola nera” in cui non si voleva entrare.

B. *Comportamentismo filosofico*. La teoria comportamentista in psicologia è convergente con l’analoga dottrina nel campo filosofico. Quest’ultima prende le mosse dal tentativo del Circolo di Vienna di arrivare all’eliminazione nelle scienze di ogni concetto o proposizione non verificabile, da dove nasce il progetto “fisicalistico” di *riduzione di tutto il sapere alla fisica*. Anche la psicologia dovrebbe essere ridotta al modello epistemologico fisico (perdendo quindi autonomia). Un altro elemento del comportamentismo filosofico, presente nei neopositivisti e in alcuni autori della filosofia analitica con diverse sfumature, è la critica del linguaggio “mentalista”, un linguaggio puramente interno o “rappresentazionale” (brentaniano) che sarebbe privo di senso. Gli

psicologi comportamentisti propendevano per una metodologia exteriorista a causa delle carenze dell'introspezione, il che era un motivo parzialmente valido. Invece i filosofi aderenti a questa corrente introducevano motivi di più ampia portata, riconducibili al nominalismo, al pragmatismo e al materialismo.

Nell'analisi delle posizioni filosofiche e scientifiche bisogna distinguere tra i *problemi reali* e le *soluzioni inadeguate*. Cartesio e i comportamentisti segnalavano problemi veri, da risolvere con equilibrio. Purtroppo con frequenza gli autori assumono posizioni estreme, credendo di aver trovato in qualche principio la chiave per ridimensionare tutte le tematiche. Cartesio, ad esempio, era ambiguo, ma i suoi principi, una volta radicalizzati, portano ad aporie, contro le quali poi altri reagiranno con non minore estremismo, onde sorgeranno nuove aporie e via dicendo. I recuperi degli elementi persi in questi radicalismi antagonisti sono talvolta interessanti, ma fragili, e siamo così trascinati da una sorta di dialettica tra estremi opposti. Questi movimenti della "ragione problematizzata" si produrranno a non finire nell'avventura della filosofia della mente.

Nelle sue *Ricerche filosofiche*, Wittgenstein (1889-1951), al di là della sua posizione teoretica (in realtà inesistente), troviamo certi elementi che facilmente inducono al comportamentismo. Uno di essi è la *critica del linguaggio puramente privato*, attribuito a Cartesio. Il termine "dolore", se significasse soltanto ciò che io provo e che nessun altro può provare, cioè la mia esperienza privata, non disponibile al pubblico, sembra che non potrebbe servire per comunicarmi con gli altri. Il significato delle parole, in quanto è chiaramente pubblico, dovrebbe essere correlato piuttosto a certe esperienze pubbliche, ad esempio al modo di reagire nei confronti delle manifestazioni di dolore (pianto, contorsioni del corpo). Una simile critica è avanzata nei confronti delle rappresentazioni concettuali. La tendenza di Wittgenstein è di collegare il significato delle parole non a semplici "contenuti" mentali fissi e astratti, ma a forme di condotta pubblica o forse all'uso pragmatico che noi facciamo con le parole in diversi contesti vitali e seguendo certe convenzioni ("giochi linguistici").

I rilevamenti euristici e problematici di Wittgenstein sul senso del linguaggio "mentalistico" non costituiscono un vero behaviorismo dogmatico (ma sono impostati in un'atmosfera ad esso vicina). Direi che essi si muovono nella direzione di un approfondimento dei significati, in lontananza dalla matrice cartesiana, che vedeva in ogni espressione psicologica un puro e semplice riferimento ad uno stato interno della coscienza, indubitabile e inaccessibile ad altri, e in ogni espressione concettuale credeva di trovare soltanto l'espressione di una rappresentazione fissa e invulnerabile. Si può però dare una risposta non comportamentista a tale problematica, senza

ricadere nel cartesianismo.

Tendo a pensare che Wittgenstein avesse in mente un modello di comunicazione intersoggettiva dove l'interno e l'esterno delle persone che s'incontrano non sono separate. "Il mio atteggiamento nei confronti di un uomo è un atteggiamento nei confronti della sua anima" (*Ricerche filosofiche*, appendice IV: verificare testo italiano). Per questo motivo egli scrive: "Ma non è assurdo dire che un *corpo* prova dolore? E perché si avverte in ciò un'assurdità? In che senso non è la mia mano a sentire dolore, ma sono io che sento dolore nella mano? Che genere di questione controversa è mai questa: È il *corpo* che sente dolore? Come può essere risolta? Come è possibile avvertire che *non* è il corpo? Ebbene, forse così: Quando uno sente dolore alla mano, non è la mano che lo dice, a meno che non lo scriva; e non si rivolgono parole di conforto alla mano, ma al sofferente; a lui si guarda negli occhi" (*Ricerche filosofiche*, n. 286). Come dire: il dolore non si esaurisce nella pura fisicità, ma ha a che vedere con un soggetto.

Queste parole sono lontane dal cartesianismo e dal comportamentismo. Esse intendono riconoscere nella persona dolente l'espressione naturale del suo dolore, non qualcosa di nascosto. Wittgenstein fa notare la difficoltà di provare indifferenza "epistemica" di fronte alla persona sofferente: "prova un po' a mettere in dubbio -in un caso reale- l'angoscia, il dolore di un'altra persona" (*Ricerche filosofiche*, n. 303). Ancora: "quando vedo qualcuno torcersi dal dolore per cause evidenti, non penso: quello che prova mi è nascosto" (*Ricerche filosofiche*, parte II, XI). In modo simile: "Come avviene che io sia pieno di compassione *per quest'uomo*? Come si vede qual è l'oggetto della compassione? (la compassione è, si può dire, una forma di convinzione che un altro prova dolore)" (*Ricerche filosofiche*, n. 287).

In definitiva, mi pare che Wittgenstein abbia in mente una concezione della conoscenza intersoggettiva più vicina a San Tommaso (per il quale l'altra persona si "vede" con gli occhi, a titolo di sensibile *per accidens*) anziché al comportamentismo. "Quando si vede il comportamento di un essere vivente, si vede la sua anima" (*Ricerche filosofiche*, n. 357, verificare testo italiano). Per questo motivo egli concludeva l'opera citata manifestando perplessità riguardo ai metodi della psicologia: "la confusione e la sterilità della psicologia non si possono spiegare dicendo che è una 'scienza giovane'; il suo stato non si può paragonare, ad esempio, con quello della fisica ai suoi primordi (...) In altre parole, in psicologia sussistono metodi sperimentali e confusione concettuale" (*Ricerche filosofiche*, parte II, XIV).

Alcuni filosofi interpretarono le difficoltà wittgensteniane riscontrate nelle sue analisi linguistiche in un senso materialistico "esteriorista" e "pragmatista". A partire dalle premesse

comportamentistiche, per molti autori diverrà un problema *reale* -che io non condivido- il fatto di “sapere se il mio dolore di denti è uguale al tuo”. Noi non potremmo comunicarci a vicenda i nostri stati interni. Con tale impostazione, neanche si potrà sapere se “tu intendi per amore ciò che intendo io”. Il nominalismo radicale porta alla mancanza di comunicazione. D’allora in poi, in certi settori della filosofia della mente si porrà come un problema da non facile soluzione l’interpretazione dei termini interioristi, legati alla coscienza e, come vedremo più avanti, si cercherà di eludere il linguaggio “di coscienza”. È questa un’eredità del comportamentismo di cui la filosofia contemporanea della mente risente ancora.

Curiosamente, alla radice di quest’eredità si trova il solipsismo cartesiano (nella misura in cui un cartesianesimo rigoroso dovrebbe essere solipsista). Se io sono sicuro soltanto dei miei stati presenti di coscienza e interpreto in questo senso il linguaggio interiorista (ma quale linguaggio non è in alcuna misura interiorista? Noi diciamo ciò che pensiamo), allora il senso del linguaggio interiorista diventa un problema insuperabile, e quindi si giustifica la proposta comportamentista, che comunque è ugualmente destinata al fallimento, in quanto essa si compie inevitabilmente in un linguaggio che esprime le nostre idee. Potremmo dire al comportamentista: come siamo sicuri che il tuo concetto di comportamentista è uguale al mio? E se egli, attraverso il linguaggio, mi segnala delle azioni esterne, si potrà sempre domandare: come siamo sicuri che tali parole rimandano esattamente a tali azioni esterne, interpretate in un certo modo? Un comportamentismo radicale è autocontraddittorio. Tra l’altro, noi capiamo questo movimento filosofico attraverso libri e articoli che comprendiamo, il che non può essere spiegato dal comportamentismo.

Un comportamentista “analitico” nel senso della filosofia linguistica (di Oxford) è *Gilbert Ryle* (1900-1976), resosi famoso con il suo libro *The Concept of Mind* (1949, tradotto in italiano col titolo *Lo spirito come comportamento*). Ryle, come Wittgenstein, si scaglia contro il dualismo cartesiano che riduce l’uomo a un corpo abitato da una “mente” fantomatica. La filosofia di Ryle cerca di smontare le difficoltà linguistiche, se andiamo all’uso ordinario della lingua, derivate dal dogma cartesiano dello “spettro nella macchina”. Il linguaggio mentalista invece si riferisce al comportamento pubblicamente osservabile degli uomini. Se in un pranzo chiedo il sale a un mio vicino e me lo passa, posso dire che lo ha fatto “per educazione”. Ora con quest’espressione non mi riferisco a qualcosa di nascosto nel suo intimo, ma semplicemente alla sua disposizione a comportarsi in un modo educato in determinate circostanze. Il linguaggio mentalista quindi è *disposizionale*. Esso indica *disposizioni comportamentali*. “*La mia mente* significa la mia capacità o tendenza a far certi tipi di cose, non un apparecchio personale senza il quale non potrei farle” (G. Ryle, *Lo spirito come comportamento*, Laterza, Roma-Bari 1982, p. 143). Se dico “ho capito lo

scherzo”, ad esempio, non mi riferisco ad uno strano atto interiore, ma indico la mia capacità di fare certe cose che dimostreranno che l’ho capito, come spiegare lo scherzo, rispondere a certe domande, e cose simili.

La teoria di Ryle è diversa del comportamentismo psicologico. Non considera i meccanismi fisiologici regolari ai quali si dovrebbe ridurre la nostra condotta, né segue la struttura stimolo-risposta, ma vuole piuttosto convincerci che il linguaggio ordinario non è favorevole al dualismo tradizionale mente-corpo. “Io” è semplicemente un pronome che indica la persona che parla e agisce, ricordare un evento è essere capace di raccontarlo di nuovo, e tutti gli altri termini in apparenza mentalisti si potrebbero risolvere in azioni esterne della condotta umana o almeno nelle disposizioni ad agire in un certo senso.

*Comportamentismo come “fisicalismo causalista”.* Altri comportamentisti filosofi, come Carl Hempel, seguendo una simile strada proposero di *definire* e di *tradurre* i vocaboli psicologici in termini di condotta esterna osservabile. Questo programma riduzionistico appartiene al *fisicalismo* e nasce dall’orientamento del *Circolo di Vienna*. Un dolore di denti dovrebbe essere definito mediante la descrizione di eventi osservabili, ad esempio il lamento, certi movimenti facciali, lesioni fisiche ai denti e la corrispondente base neurologica. Il comportamentismo fisicalistico (Carnap, Hempel) si basa sull’analisi del linguaggio, ma intende a ridurre il linguaggio psicologico, tramite definizioni che consentano di effettuare traduzioni, al linguaggio scientifico. Invece il comportamentismo “analitico” di Ryle, più vicino al metodo di Wittgenstein, preferisce la riduzione a comportamento umano esterno (ma cerca anche una sorta di equivalenza logica tra il presunto linguaggio “mentalista” e le disposizioni comportamentali).

*La riduzione dell’evento cognitivo alla causalità fisica* è uno dei tratti del comportamentismo filosofico neopositivista o fisicalista, anzi è la chiave della riduzione della psicologia ad una scienza naturale. Questo punto si manifesta nella versione nominalistico-causale della conoscenza. Lo stimolo sensibile, invece di essere considerato come un segno intenzionale che fa “conoscere”, viene interpretato dal comportamentismo fisicalista come un segnale (causale) che scatena una condotta. Questa concezione è più drastica quando è applicata all’uomo. I segnali sensibili inviati agli animali, in fin dei conti, hanno lo scopo di indurre una certa condotta pratica (l’animale vede il cibo per andare a prenderlo). Nella conoscenza intellettuale umana, invece, i segni non sono “segnali per fare qualcosa”, bensì rimandano a ciò di cui sono segni con indipendenza dal fatto che da ciò segua una prassi esterna. L’animale vive immerso nel mondo fisico, al quale vede soltanto come oggetto dei suoi bisogni pratici. L’uomo invece è capace di separare il significato dagli aspetti

pratici (cfr. su questo punto il mio lavoro *Realismo y referencia*, in *Atti "XXV Reuniones filosóficas, Universidad de Navarra, El hombre: inmanencia y trascendencia"*, Servicio de Publicaciones, Pamplona 1991, pp.259-275).

Un'interpretazione comportamentista dei segni si trova in *W. Van O. Quine* (cfr. *From a Logical Point of View*, Harvard Univ. Press, Cambridge (Mass.), pp. 47-64), il quale sviluppa con tale premessa una forma di ontologia linguistico-trascendentale di tipo olistico e naturalistico. La visione "ontologica" del mondo, impressa in un linguaggio inteso secondo il comportamentismo nominalista, è poco intelligibile -e non poteva non esserlo per un nominalista coerente-, per cui si conclude l'impossibilità di una traduzione radicale tra i diversi linguaggi. Approdiamo all'incommensurabilità di fondo tra i linguaggi delle culture. Tale visione ontologica è fondata sul nostro inserimento nella natura così come è descritto dal comportamentismo. Tuttavia la filosofia di Quine non è un puro comportamentismo, ma piuttosto un "kantismo linguista ed empirico", senz'altro radicalmente naturalista e materialista.

*Rilevamenti critici.* La "dissoluzione" degli atti o situazioni psichiche o psicofisiche in azioni o disposizioni ad agire o semplicemente ad atti esterni comporta un'analisi deficiente delle azioni umane. La realtà è molto più complessa.

a) *Ci sono diversi tipi di azioni umane:* 1. Reazioni fisiologiche (salivazione, respirazione, tremore, battiti del cuore). 2. Movimenti del corpo (alzare un braccio, correre). 3. Azioni *umane* che comportano un movimento del corpo: chiamare per telefono, leggere un libro, visitare un amico. 4. Azioni che non comportano un movimento del corpo: ragionare, indovinare, giudicare, decidere (cfr. J. Kim, *Philosophy of Mind*, Westview Press, Boulder (Colorado) 1996, p. 28).

Queste ultime azioni sono gli "atti psichici" esclusi dal comportamentismo, a meno che non si risolvano in forme pubbliche di condotta umana. I numeri 1 y 2 sono i tipi di atti preferiti dal comportamentismo psicologico. Il n. 3 potrebbe includere per un "mentalista" qualche tipo di atto interiore, mentre il comportamentismo cerca di interpretarle escludendo gli elementi interiori, ad esempio le rappresentazioni concettuali, le intenzioni o le emozioni.

*Molte delle nostre azioni sono esterne e pubbliche -come il linguaggio-, ma non si riducono ad atti fisici così come vengono descritti dalla fisica, dalla fisiologia o dalla neurologia. Salutare o ringraziare sono atti pubblici, visibili ad altri che sappiano interpretarli, proprio perché includono che il soggetto cui si saluta o cui si ringrazia riconosca nell'agente un'azione esterna del corpo che esprime un'intenzione, un atto di volontà o un pensiero. I comportamentisti del linguaggio ordinario*

si dispensano dall'analizzare la condotta umana in termini antropologici. Il cosiddetto "atto umano pubblico" è un'azione esterna animata da un significato più alto. Il movimento di una mano può esprimere un giudizio affermativo o negativo, una minaccia, un saluto, un sì volontario.

b) *Molti atti interiori sono irriducibili ad atti esterni.* Nessuna argomentazione comportamentista convincerà sul serio alla gente che il dolore si debba "tradurre" a condotta esterna. Atti come "mi ricordo che ieri ho comprato il giornale", "non credo che questo partito possa vincere le elezioni", "i numeri sono infiniti" non sono traducibili ad azioni esterne. Ciò non vuol dire che questi atti interni non abbiano a che vedere con nessun atto esterno o con ulteriori atti interni. Ciò è impossibile, poiché qualsiasi atto interiore si può almeno esprimere esternamente e sempre comporta rapporti con altri atti interiori (cognitivi e volitivi). Il pensiero "non credo che questo candidato possa vincere le elezioni", oltre ad esprimersi in modo linguistico, come abbiamo fatto qui, presuppone la comprensione di un insieme di significati correlati e impliciti (ad esempio, i motivi per cui non si crede tal cosa).

Questo stesso atto ("non credo che...") comporta inoltre la disposizione a compiere altri atti interni, come "sono preparato per mostrare i motivi in favore di tale opinione", "sono disposto e discutere con chi nega la verità di tale frase". Sarebbe sciocco afferrarsi esclusivamente al lato esterno di questi atti, i quali non avrebbero alcun senso se non venisse colta la loro parte interna. L'aspetto di verità che possiamo riconoscere nel comportamentismo è il fatto che i nostri atti interiori, nel loro dinamismo di solito non sono isolati, ma sono intrecciati con molte altre azioni umane. La nostra rappresentazione di tram comporta che noi quando vedremo un tram potremo riconoscerlo e utilizzarlo come mezzo di trasporto, ma proprio grazie al fatto di avere tale rappresentazione.

K. Popper ha formulato una serie di argomentazioni sull'eterogeneità irriducibile tra il linguaggio fisico e quello psichico. (Cfr. *Congetture e confutazioni*, il Mulino, Bologna 1972, pp. 499-509: *Il linguaggio e il problema del rapporto corpo-mente*). Egli indica due modi comportamentistici di affrontare questo problema:

i. Si ammette con benevolenza l'esistenza di due linguaggi diversi, uno "mentalista" e l'altro "fisicalista". Ma l'atto interno viene ridotto ad un modo di parlare, traducibile nel linguaggio vero e proprio, che sarebbe il linguaggio scientifico. Per Popper questa "traduzione" è illegittima: quando dico "sento dentro di me l'ira", affermo qualcosa di diverso da quando dico "muovo le mie mani per dare un colpo al mio avversario". D'altra parte, parliamo ordinariamente un unico linguaggio che contiene elementi fisici e psicologici (non ci sono "due linguaggi"), come quando diciamo

“immaginare delle pecore mi aiuta ad addormentarmi”.

ii. Si ritiene più giusto parlare solo in termini di comportamento esterno, dichiarando anomalo il linguaggio mentalista, che sarebbe legato a Cartesio. Anche questa posizione è falsa per Popper: se ho *mal di denti*, è assurdo ridurre questo dolore ad un certo modo esterno di comportarsi, a causa del pregiudizio verificazionista. Se un capo stazione è *convinto* che il treno sta per partire, questa sua convinzione è assai diversa dal suo comportamento esterno.

c) *Alcuni stati personali si chiarificano con l'indicazione della potenzialità (“disposizione”) per compiere certi atti.* Questo è il punto positivo che possiamo imparare da Ryle, se tralasciamo il riduzionismo. Ad esempio, essere pigri, generosi, furbi, ecc. sono qualità definibili in termini di disposizioni o potenzialità a fare o a non fare certi atti. Sarebbe assurdo pretendere che la generosità sia una sorta di qualità statica interiore, indipendente dalla condotta. I classici si riferivano a questi tratti personali mediante il concetto di *abito*, inteso come predisposizione attiva verso un tipo di azioni. Una persona è intelligente, in questo senso, se è capace di agire intelligentemente (fare calcoli, cogliere relazioni, prevedere, programmare, risolvere problemi).

D'altra parte, la *disposizione* è una realtà potenziale non osservabile, non troppo consone con il verificazionismo, il quale se viene preso rigorosamente dovrebbe limitarsi a costatare ciò che è osservabile, senza postulare entità inosservabili. Portato all'estremo, il verificazionismo non regge, poiché deve ricorrere almeno a leggi statistiche (non osservabili) in virtù delle quali si possano fare delle previsioni. Ma quelle leggi statistiche, se sono pure delle regolarità senza alcun'altra indicazione causale, non sono molto informative, ma si riducono a semplici descrizioni che possono cambiare in qualsiasi momento (empirismo estremo humeano).

d) *Certi atti interiori (amare, credere, sperare) sono autentici o coerenti se il soggetto è in grado di compiere determinati atti esterni in virtù di essi.* In questo senso, l'Apostolo Giacomo dichiarava di non credere ad una fede senza opere, o più precisamente diceva che una fede senza opere era morta (“mostrami la tua fede senza le opere, ed io con le mie opere ti mostrerò la mia fede”: *Giacomo 2, 18*). Se qualcuno dice di amare una persona e la maltratta, con i fatti dimostra di non amarla davvero. L'amore o la giustizia chiedono atti, ma neppure s'identificano *tout court* con gli atti esterni, poiché le manifestazioni di amore o di giustizia possono essere molteplici, e comunque l'opera esterna è priva di valore se non è una reale espressione di un atto interno (un robot non può voler bene le persone, anche se può simularlo). Il vero credere impone, per coerenza, l'agire secondo il giudizio di fede (se *credo* che quest'autobus mi porta alla stazione, lo *prendo*).

*Ma non è possibile stabilire una totale equiparazione tra fede, desideri, intenzioni, e gli atti esterni corrispondenti o coerenti con tali atti interiori.* “Credo che pioverà” provoca che io prenda l’ombrello quando esco da casa. Ma potrei pure prendere un altro mezzo per ripararmi dalla pioggia, o non prenderlo perché voglio correre il rischio di bagnarmi, o perché è di un colore che mi mette in ridicolo, o perché devo prestare l’ombrello a un mio fratello, oppure posso uscire con l’ombrello perché voglio ripararlo, e così via all’infinito. Le possibilità sterminate di combinazione e di variazione tra le nostre disposizioni interiori e la nostra condotta esterna, in tutta la sua indefinita varietà, impediscono di stabilire una semplice equazione tra atti interni e condotta esterna. Esiste soltanto una corrispondenza variabile, non un’identificazione, tra atti interiori ed atti esterni. Il comportamentismo poteva apparire forse plausibile con alcuni pochi esempi molto semplici, ma chiaramente non regge di fronte alla complessità dell’agire umano (e anche animale).

Un’azione razionale invece, come spiega Kim, spesso segue lo schema “desidero p, e credo che facendo A, otterrò p, per cui tento di fare A” (desiderio, credenza, ragionamento, azione). Se comprendo questi elementi, posso capire l’azione di una persona, e anche prevederla in qualche modo, se so a che cosa crede e che cosa desidera. Ma non ci sono leggi vere e proprie che consentano di fare predizioni assolute della condotta umana, perché le persone potranno sempre avere uno stato mentale diverso che cambi le loro decisioni. Ad esempio, “Sally desidera aria fresca e crede che può ottenerlo aprendo la finestra”, ma forse non la apre perché ritiene che se lo fa ci sarà troppo rumore, oppure forse addirittura la apre, nonostante tutto, perché crede che sua madre malata ha bisogno di aria fresca...e così all’infinito (cfr. J. Kim, *Philosophy of Mind*, cit., pp. 34-35). Con un altro esempio: saluto qualcuno per la strada. Questo semplice atto include parecchi atti mentali, come riconoscere la persona, volerla bene, sperare che l’altro riconosca il mio gesto e che lo interpreti come espressione dell’intenzione di salutare. Ora, non esiste un tipo di movimento fisico unico che significhi il salutare. Si può farlo in molti modi, diversi nelle culture, affinché il saluto sia capito come tale, non come un insulto o come mille altre cose (cfr. *ibid.*, pp. 35-36). Quindi le difficoltà del comportamentismo sono insormontabili.

Verso o la fine degli anni ‘50 il comportamentismo cominciò a declinare negli Stati Uniti. Un colpo decisivo per comportamentismo psicologico fu una recensione di Noam Chomsky, pubblicata nel 1959 nella rivista *Language*, dell’opera di Skinner *Verbal Behavior*. In quest’opera Skinner cercava di spiegare il linguaggio in termini comportamentali, ma Chomsky evidenziò come esso non poteva comprendersi senza supporre un bagaglio previo di strutture cognitive che governavano il comportamento verbale esterno.

Il declino del comportamentismo coincide con la caduta dell'empirismo logico. Le scienze ricorrevano a ipotesi inosservabili per la spiegazione dei fenomeni. Non era logico restringere l'indagine della condotta umana alla descrizione regolare di sequenze di atti fenomenici secondo leggi di probabilità (descrizione impossibile), tralasciando il fatto che tra l'input esterno sul corpo umano e l'output della condotta c'era una "scatola nera" che in realtà era piena di un'intensa attività (desideri, intenzioni, convinzioni, propositi, emozioni).

## 5. Teoria dell'identità (mente=cervello)

Il comportamentismo aveva tralasciato la base neurologica degli stati psicologici. La crisi di questa corrente indicava la necessità di indagare sul *black box* che stava tra gli stimoli ambientali e le risposte comportamentali. Per un'impostazione materialista, questa scatola nera erano i fenomeni studiati dalla neurologia. Verso la fine degli anni '50 si fece avanti la cosiddetta teoria "dell'identità", che sostiene l'equiparazione tra fenomeni psichici e neurologici o, in altre parole, tra la mente e il cervello. *Ma le argomentazioni in favore di tale teoria spesso si riferiscono alle sensazioni* (l'esempio prediletto è il dolore), non alle attività superiori della mente, come il pensiero, il giudizio, i ragionamenti o le intenzioni. Solo in questo settore tali argomentazioni possono sembrare più plausibili, specialmente se il dualismo cartesiano è rifiutato. Di solito l'identità viene però estesa a qualsiasi fenomeno mentale o spirituale, *il che non è dimostrato, ma postulato*.

Entriamo in un campo dove le polemiche sono state ampie e non sempre chiare. Conviene non perdere il senso critico nei loro confronti. *La tendenza attuale nella filosofia della mente*, capeggiata da autori americani (anche inglesi e australiani) è *materialista* (sia pure in un modo sofisticato). Questo fatto è un'eredità di certi orientamenti filosofici della cultura di lingua inglese: *empirismo, nominalismo, positivismo logico, scientismo, evolucionismo filosofico, comportamentismo*. Le idee antropologiche della filosofia europea (esistenzialismo, fenomenologia, personalismo, ermeneutica) non sono state accolte in profondità nella cultura anglosassone, che è molto polarizzata verso le scienze naturali, tranne poche eccezioni. Le discussioni che intercorrono in queste aree, sin dall'epoca del comportamentismo fino ad oggi, sembrano "polemiche di famiglia tra materialisti".

Indicherò in seguito alcune caratteristiche generali di queste correnti della filosofia della mente (comportamentismo, già visto, teoria dell'identità, eliminativismo e alcune forme di funzionalismo):

1. Rifiuto netto e costantemente ripetuto del dualismo cartesiano o "mentalismo", considerato

come bersaglio preferito delle critiche. Gli errori del cartesianesimo sembrano ovvi a questi autori. Spesso, però, le tesi cartesiane sono presentate in un modo semplificato e anche un po' ridicolo ("teatrino cartesiano", "entità spirituale nascosta e privata, fantomatica", *homunculus*).

2. Ignoranza storica delle posizioni spiritualistiche della tradizione e anche della modernità. Lo spiritualismo viene spesso liquidato come cartesiano. Le filosofie di Aristotele e di Tommaso d'Aquino sono ignorate (Aristotele talvolta è considerato come un funzionalista o un materialista). *Confusione, quindi, tra spiritualismo e dualismo cartesiano*. Di conseguenza, indirettamente o implicitamente le critiche contro Cartesio ricadono sull'aristotelismo e sulla tradizione cristiana.

3. *Stile scientifico* nella presentazione e difesa delle tesi sulla filosofia della mente. Il metodo discorsivo di questi autori segue un certo modello delle scienze naturali, in quanto intende elaborare una "filosofia scientifica". Ne seguono una serie di note: discussioni tematiche specifiche; tono oggettivo, talvolta facilmente concludente, senza le fatiche "inesauribili", ad esempio, di un filosofo come Wittgenstein; assenza di prospettive storiche ampie, tranne quelle che si limitano alle discussioni di scuola. La restrizione tematica, l'oggettività univoca e l'incuranza storica sono tipiche della letteratura scientifica. Queste note si trasmettono allo stile in cui solitamente si fa filosofia nel campo del materialismo e dello scientismo. Non intendo additare queste caratteristiche "poco filosofiche" ad ogni autore (sono più tipiche dei filosofi mediocri, non nei grandi autori).

4. Dallo stile scientifico nell'elaborazione letteraria di questa filosofia seguono altre note che chiamerei *scolari*, il che comporta alcuni aspetti positivi e altri negativi: una certa uniformità nei problemi o tesi, citazioni preferenziali all'interno del "club" di pensatori accettati, il gusto per divisioni "scolastiche" delle diverse posizioni, schieramenti molto netti in torno alle posizioni, polemiche entro certi parametri, chiarezza e definizione nelle premesse e, in questo senso, non di rado petizioni di principio dissimulate, che intendono ribadire le proprie idee ed escluderne altre (ad esempio, le "prove" in favore dell'identificazione tra atti psichici e fisici di solito contengono petizioni di principio).

La teoria dell'identità venne diffusa verso la fine degli anni '50 negli USA come una conseguenza del fisicalismo di Carnap. Uno dei suoi rappresentanti, sin dagli anni '30, era stato *Herbert Feigl*, neopositivista logico (vedere articoli "*Il 'mentale' e il 'fisico'*", 1958; *Mente-corpo, non uno pseudo-problema*, 1960). Per Feigl, i termini psicologici e quelli fisici corrispondenti hanno lo stesso riferimento (gli stati cerebrali), mentre differiscono nel significato. Il termine psicologico significa piuttosto l'apparenza, così come l'acqua, avendo una struttura reale di H<sub>2</sub>O, ci si presenta con certe caratteristiche fenomenologiche o antropomorfiche.

In modo simile, l'australiano *Ullin T. Place* riservò il comportamentismo disposizionale per le funzioni cognitive e volitive -chiamate da alcuni analitici, sulla scorta di Quine, *atteggiamenti proposizionali*, come “credo che p”, “desidero p”, “voglio p”, in quanto hanno di fronte delle proposizioni a titolo di oggetti-, ma non per gli stati o atti di coscienza (sensazioni, immagini, esperienza), *nei confronti dei quali Place propose la loro equiparazione ai processi cerebrali*. Tale identità viene dichiarata *contingente*, non analitica o in base a una definizione, in quanto è il risultato di una scoperta empirica (questa è una differenza metodologica rispetto al comportamentismo logico, per il quale gli stati mentali si dovevano definire *a priori* in termini di condotta). Anche se in un primo momento può sembrare che i processi psichici non abbiano niente a che vedere con gli eventi cerebrali, così come nessuno pensa *a priori* che i fulmini siano scariche elettriche, posteriormente si potrebbe scoprire tale identità di fatto. Nel 1958 *Jack J. C. Smart*, un altro australiano, fece conoscere la teoria dell'identità negli Stati Uniti in un ciclo di conferenze.

Le versioni di Place e Smart della teoria dell'identità sono iniziali. A causa delle critiche mosse dal funzionalismo contro tale teoria (cfr. più avanti), negli anni successivi si è fatta avanti una teoria dell'identità più complicata, chiamata *causale* (o *del ruolo causale*). *David K. Lewis*, americano vicino ai filosofi australiani, in un articolo del 1966 (*Un argomento a favore della teoria dell'identità*) accolse l'idea di Carnap secondo cui i concetti psicologici potrebbero essere visti come “concetti teorici” (analoghi a quelli usati nella fisica). Visto il ruolo causale dei concetti teorici, le nozioni psicologiche potrebbero essere definite tramite i rapporti di causa-effetto stabiliti con gli stimoli oppure tra di esse (si veda la distanza dal comportamentismo). Lo schema causale sarebbe: *stimolo -> stato mentale -> un altro stato mentale -> comportamento esterno* (vedo un libro -> desidero comprarlo -> ritengo che per comprarlo debbo entrare nella libreria -> entro nella libreria e lo compro). Riprendiamo così la tematica della causalità tra atti mentali e atti fisici, o tra gruppi di atti mentale). La teoria dell'identità di Lewis accetta questa complicazione, ma ritorna sul punto specifico della prospettiva materialista: la causalità degli atti mentali va ricercata negli stati del sistema nervoso.

Un altro australiano, *David M. Armstrong*, propose una sua versione della teoria causale nel volume *A Materialist Theory of Mind* (1968) (gli autori menzionati appartengono alla cosiddetta “scuola australiana materialista”). Come gli altri esponenti, anziché cercare argomenti in favore di un'identificazione degli atti mentali con i fenomeni cerebrali, un compito lasciato alla neurologia, Armstrong vuole argomentarne la possibilità, scongiurando le obiezioni di principio contro tale ipotesi. La sua proposta, da lui chiamata “teoria dello stato centrale” (in riferimento al sistema nervoso), solo nei dettagli si distingue da quella di Lewis. L'analisi concettuale delle nozioni

psicologiche considera gli intrecci causali corrispondenti, ma il sistema nervoso centrale è il candidato a svolgere il ruolo “causale” degli stati psicologici.

Le *teorie dell'identità causale* di Lewis-Armstrong sono più sofisticate di quelle iniziali di Place-Smart, al punto che certi autori le hanno catalogate come funzionalistiche, o come un gruppo separato, da chiamarsi “funzionalismo causale” o altro. In realtà quegli autori sostengono chiaramente la tesi dell'identità. Sono materialisti espliciti, a differenza dei funzionalisti. La loro “sofisticazione” nasce dalle difficoltà sollevate dal funzionalismo e dalle proposte di quest'ultima corrente, coinvolta nelle tematiche causali. Tale complicazione non aumenta il valore dimostrativo del materialismo, anche se lo aiuta a sopravvivere.

Una teoria materialista vicina a quelle precedenti è il vecchio *epifenomenismo* (*Th. H. Huxley*, fine del XIX secolo e inizi del XX secolo). Per Huxley, gli stati fisici “fanno apparire” gli eventi psicologici, ma non all'inversa, un po' come il corpo produce la sua ombra. Ovviamente sarebbe inadeguato pensare che l'ombra può causare qualcosa del corpo. Essa è soltanto un prodotto derivato, privo di efficacia causale, covariante con i mutamenti del corpo. Se “desidero muovere un dito”, la vera causa starà nel cervello, non nel mio “desiderio”. La teoria causale dell'identità, invece, “sembra” assegnare un ruolo causale ai fenomeni psichici, ma in realtà lo fa solo per poi ridurre tale causalità agli eventi fisici.

Presenterò adesso una visione panoramica delle posizioni filosofiche sul rapporto mente-corpo e sulle relazioni tra gli atti psicologici e gli atti fisici:

\* *Dualismo interazionista* (Cartesio, Popper, Eccles): si riconosce la distinzione reale tra atti mentali e fisici e l'esistenza di una vera interazione tra di essi (una puntura provoca un dolore al braccio, inteso come sensazione psichica che causa il movimento del braccio). Si distingue spesso tra dualismo *ontologico* o *sostanzialista* (viene riconosciuta l'esistenza di due sostanze) e dualismo delle sole *proprietà* (non si parla di due entità, ma si ammette l'eterogeneità di due tipi di atti interattivi).

\* *Monismo psichista* o *panpsichismo* (Berkeley): non c'è una reale distinzione tra atto materiale e atto psichico. La materia contiene una forma di psichismo (mente, razionalità) o si riduce a fenomeno psichico (Berkeley). Lo stoicismo, l'idealismo, il razionalismo di Leibniz, possono considerarsi in qualche modo come filosofie panpsichiste.

\* *Parallelismo psicofisico* (Spinoza, Wundt): c'è una qualche distinzione tra gli atti mentali e fisici, ma non si accetta che possano interagire. Quindi le vie causali psichica e fisica sono chiuse.

Di per sé il parallelismo è una dottrina ambigua, che può essere interpretata diversamente. Alcuni parallelisti risolvono il problema causale direttamente in Dio (*occasionalismo*: Malebranche), o nel coordinamento estrinseco (operato da Dio) tra il mondo psichico e il mondo fisico (*armonia prestabilita*: Leibniz), o semplicemente si fermano metodologicamente al parallelismo, come un modo di distinguere tra la fisica e la psicologia (Wundt), forse pensando che esiste un rapporto sconosciuto tra quei due ambiti. Alcuni parallelisti tendono, invece, al monismo: atti fisici e psichici sarebbero due manifestazioni di un'identica realtà (Spinoza li vede come due attributi della Natura o Sostanza, che è Dio). Altri autori dimostrano una propensione per il panpsichismo (Leibniz) oppure per il materialismo.

\* *Monismo materialista*: la distinzione reale tra atti psichici e fisici è respinta. Gli atti psichici si risolvono o si riducono a eventi materiali. Finora abbiamo considerato tre varietà di materialismo:

- *comportamentismo*: gli atti psichici si risolvono in comportamenti fisiologici secondo lo schema “stimolo sensibile-risposta motoria o ghiandolare” (Pavlov, Skinner: *comportamentismo psicologico*); oppure gli atti psichici si “dissolvono” in quanto si definiscono in termini di comportamento pubblico o di disposizioni comportamentali (Ryle: *comportamentismo filosofico o analitico*), o in termini degli atti fisici scoperti dalla scienza (fisicalismo di Carnap o di Hempel, vicini alla teoria dell'identità);

- *epifenomenismo* (Huxley): gli atti psichici sono sottoprodotti della causalità fisica e non possono esercitare alcuna influenza causale sulla realtà;

- *teoria dell'identità*: gli atti psichici sono manifestazioni apparenti di quello che in realtà è soltanto attività cerebrale (mente=cervello). Possono avere un “ruolo causale”, ma riducibile all'attività nervosa (teorie del “ruolo causale”).

Per completare le posizioni materialistiche, bisognerebbe menzionare il *funzionalismo* e l'*emergentismo*, posizioni meno radicali e forse aperte ad un'uscita dal materialismo, e infine l'*eliminativismo*, molto più radicale. Ci riferiremo più avanti a queste posizioni, secondo il progresso concettuale della nostra esposizione.

Il quadro fin qui presentato è solo orientativo. Tante posizioni sottili degli autori non sono riconducibili a queste tipologie, che restano incomplete ed aperte. Sarebbe difficile inquadrare in quelle categorie le filosofie, ad esempio, di Kant, Russell o Wittgenstein, per non parlare di Aristotele e San Tommaso (che non sono propriamente né dualisti né monisti). Ho voluto delineare questo quadro solo perché le divisioni tipologiche come queste vanno di moda nella filosofia della

mente. Malgrado le semplificazioni, esse proiettano un po' di luce sulle posizioni dei singoli filosofi. Di fronte a un autore di filosofia della mente, è bene domandarsi: distingue tra atti fisici e psichici? che tipo di correlazione ammette tra queste due categorie di atti? riconosce l'autonomia della psicologia rispetto alla fisica o alla neurologia? assegna un valore alle apparenze psicologiche ("psicologia del senso comune")?

*Consideriamo in seguito le difficoltà che possono essere sollevate contro la teoria dell'identità.* Restringo le mie osservazioni *al campo delle sensazioni*, dove un aristotelico è disposto a concedere il massimo alla menzionata teoria. Seguendo l'esempio preferito della teoria dell'identità, la sensazione di dolore sarebbe *identica* all'eccitazione delle fibre C del sistema nervoso (questa correlazione neurologicamente non è esatta, ma non intacca le argomentazioni). È vera tale identità? In che senso? Si può estendere ad ogni fenomeno psichico?

Siamo consapevoli della *correlazione rigorosa di molti dei nostri atti psichici con alcuni eventi fisici del nostro organismo* (toccare un oggetto ad alta temperatura produce la sensazione di caldo nella mano). L'interpretazione dualista cartesiana non è naturale in casi simili. È strano pensare che quando un oggetto riscaldato tocca la pelle, "l'anima umana sentirebbe il caldo" (il che porta alla curiosa idea dell'insensibilità degli animali). È più naturale ritenere che la nostra pelle, sotto certe alterazioni, *sente* (freddo, dolore, ecc.). Non è sempre vero che la "psicologia del senso comune" segua il dualismo cartesiano. I termini del problema cambiano se passiamo alle sensazioni intenzionali. Vedere o udire non comporta propriamente sentire il proprio corpo, ma consiste fenomenicamente nell'essere informati, in un modo vivo e diretto di certe proprietà degli oggetti fisici. Quindi possiamo avere "l'impressione" che gli atti di vedere e di udire siano immateriali, almeno nel senso che non si rivelano come un atto del nostro corpo (l'attuazione del nostro corpo passa inosservata nelle sensazioni intenzionali). Eppure anche per il senso comune esiste un'evidente correlazione tra l'attività fisica dell'organo di senso e il suo operare intenzionale. Nelle *sensazioni interne* (immaginazione, memoria) ormai non abbiamo nessuna notizia "fenomenica" dell'attività fisica del proprio corpo. L'oggetto intenzionale qui appare "solo", anche se il suo contenuto si riferisce a caratteristiche materiali dei corpi.

*La neurofisiologia dimostra l'esistenza di tale correlazione a tutti i livelli percettivi e dei sensi interni.* Rimane però fermo il modello della dualità occhio-vedere (organo fisiologico e atto psichico). "Vedo": vedo cose esterne, vedo pure il mio occhio (in uno specchio). Ciò che vediamo sono figure e colori, quindi non vediamo il nostro atto di vedere, eppure lo avvertiamo (indirettamente, poiché il vedere è intenzionale, cioè dobbiamo vedere cose esterne al vedere stesso.

Il caso del dolore è molto diverso). Ci accorgiamo, quasi ragionando, che l'atto di vedere è legato al nostro organo visivo, quindi diciamo naturalmente "il mio occhio vede", "vedo con i miei occhi". Eppure nel vedere, preso formalmente come operazione cognitiva, non "notiamo" l'occhio (notiamo, solo che per vedere dobbiamo muovere gli occhi, mantenerli aperti). Quindi *ci accorgiamo*, ma non "*sentiamo*" che vediamo con gli occhi. Proprio per questo, l'atto di vedere non ci informa sulla struttura interna dell'occhio e sulla sua attività fisiologica al momento di vedere, e tace assolutamente su tutto il retroscena fisiologico della sensazione (vie nervose visive, attività cerebrale). Man mano che l'attività sensitiva sale di livello (immaginazione, memoria), la parte organica si nasconde sempre di più, al punto che il cervello, essendo la sede della nostra sensibilità e della nostra emotività, si occulta *assolutamente* alla conoscenza ordinaria. L'uomo ha impiegato secoli per scoprire il ruolo del sistema nervoso nella conoscenza.

Dal punto di vista aristotelico, possiamo trovare un senso teleologico a questa dialettica di fenomenologia percettiva e nascondimento organico. Si nasconde alla nostra coscienza ciò che si deve nascondere, per dar luogo all'attività che più interessa. Sentire l'attività del cervello nell'atto di vedere sarebbe disturbante. Sentiamo, invece, ciò che è adeguato sentire, ad esempio, certi stati del corpo gradevoli o sgradevoli. Tali sensazioni comportano un'informazione utile o ci spingono ad azioni convenienti per la vita.

Arriviamo alle seguenti conclusioni sulla *natura dell'atto percettivo e sulla sua duplice dimensione, psichica e fisiologica*. Le indicazioni fenomeniche primarie, pur nella loro frammentarietà, sono convergenti con le scoperte neurologiche. Da ciò si conclude l'esistenza di *una correlazione intrinseca, essenziale e necessaria, tra l'atto neurofisiologico e l'atto psichico* (sensitivo). Talvolta si è detto che questa correlazione sarebbe *contingente*, un punto usato anche come argomento contro la teoria dell'identità (si potrebbe pensare ad un supporto organico della vista diverso dagli occhi, quindi sembrerebbe che vedere ≠ atto neurofisiologico). A mio parere questo argomento è debole, poiché separa troppo le due dimensioni dell'atto percettivo, come se la loro unità fosse estrinseca o arbitraria, e così ci riavvicina al dualismo platonico o cartesiano.

*L'apparenza di contingenza "arbitraria" tra le due dimensioni dell'atto sensitivo è dovuta alla nostra ignoranza*. Inizialmente non sappiamo nulla sul rapporto tra l'atto di vedere e l'attivazione dei lobi occipitali della corteccia. Una volta che scopriamo tale rapporto, tuttavia, non possiamo non riconoscerne una *necessità naturale* (non di tipo logico), così come la scoperta del rapporto tra l'acqua comunemente nota e la struttura H<sub>2</sub>O è l'accertamento *a posteriori* di una necessità naturale. Saul Kripke ha insistito su questo punto in *Naming and Necessità*. Ovviamente

ci potrebbero essere altre modalità visive, non necessariamente legate alla struttura dei nostri occhi, ma entro limiti che probabilmente non conosciamo. Non qualsiasi struttura fisica ammette la visione. Essa è legata a una forma fisica, che entro certi limiti è “molteplimente realizzabile”. Anche i coltelli possono essere fatti in molti modi, purché servano al loro scopo, da cui nascono certi limiti materiali.

*La teoria dell'identità è inadeguata perché, nell'affermare la pura e semplice omogeneità dell'atto neurologico con l'atto psichico, finisce per annullare la realtà di quest'ultimo. La “spiegazione” della base neurologica dell'atto psichico finisce per essere qui una riduzione o un'eliminazione di un versante reale dell'atto. Si conclude così con la riduzione dell'atto psichico alla sua causalità materiale. Sarebbe come dire che un bicchiere d'acqua è soltanto il materiale di cui il bicchiere è fatto, dimenticando la sua forma. Ma neppure è vera la teoria dualista, secondo cui l'atto sensitivo si duplica in due atti separabili e contingentemente collegati. La teoria aristotelico-tomista sostiene, invece, la presenza di un unico atto a due dimensioni, paragonabile a quanto succede nell'unità ilemorfica, dove una materia e una forma sono ricongiunte inseparabilmente come potenza ed atto. Ma nell'atto sensitivo siamo in presenza non di una semplice struttura morfologica o funzionale, bensì di un vero atto, atto cognitivo e anche corporeo, irriducibile ad altre dimensioni del corpo.*

*I sostenitori della teoria dell'identità trovano difficile di accettare questo punto perché, nel loro scientismo, essi sono sensibili soltanto alle spiegazioni fisiche o meccaniche. La spiegazione neurofisiologica degli atti sensitivi è, sì, una vera spiegazione, in quanto indica certe cause relativamente necessarie, ma i termini in cui essa si esprime si collocano in una dimensione non cognitiva (meccanismi fisiologici, attivazioni elettriche, attività chimica). La descrizione fisica ed organica non esaurisce tuttavia tutta la realtà dell'atto sensitivo, e non per questo bisogna postulare l'esistenza di un mondo spirituale. La dimensione psichica è reale e di essa si occupa la scienza psicologica.*

*Vi sono, quindi, due dimensioni intrinsecamente unite e realmente diverse, non come due cose, ma come è diversa la forma di una cosa materiale e i suoi costitutivi materiali, anche se in questo caso la forma è un'operazione fisica (quindi un'operazione fisica più alta rispetto alle operazioni non cognitive). La conseguenza epistemologica di questa unità nella dualità è la distinzione e allo stesso tempo il rapporto tra neuropsicologia e psicologia fenomenica.*

Ho menzionato il modello dell'acqua “fenomenica” e dell'acqua come H<sub>2</sub>O come analogo alla dualità sensazione/base fisiologica. *Quest'analogia non è del tutto sufficiente. L'acqua*

“fenomenica” è la presentazione di una realtà naturale al nostro modo umano di conoscerla. Invece l’atto sensitivo in quanto tale (non nel suo eventuale contenuto intenzionale) non è un semplice “modo di presentarsi” di una realtà, ma è esso stesso una realtà. Questo è ovvio nel caso degli atti sensitivi non intenzionali, come il dolore o il piacere. Tali atti attestano la presenza del proprio corpo nelle sue attività e nei suoi stati. Quindi la spiegazione neurologica delle sensazioni non rimanda ad una realtà “più profonda” o “più vera”, ma si riferisce soltanto alla dimensione materiale, causalmente importante, ma “inferiore” (la dimensione materiale è sempre inferiore alla dimensione formale).

Non occorre dunque lasciare la conoscenza sensitiva comune in una situazione “epistemicamente” inferiore, come invece succede con la conoscenza volgare delle cose materiali, una volta che la scienza ne ha fornito le spiegazioni più profonde. La conoscenza neurofisiologica è “più profonda” riguardo agli atti sensitivi solo nell’ambito fisico, dal quale non può uscire per motivi di metodo (pena lasciare di essere una vera spiegazione neurofisiologica). Ma di per sé, la sola spiegazione neurologica non può dare nemmeno un passo verso la conoscenza del fenomeno psichico. Semplicemente lo presuppone e ne chiarifica la base materiale. Quindi non bisogna dire che i fenomeni psichici, senza il riduzionismo della teoria dell’identità, rimarrebbero un campo “non governato dalle leggi fisiche”. *Le leggi fisiche danno piena ragione dell’atto psichico nel suo versante materiale.* Su questo versante non trovano “concorrenza” con la psicologia. Solo che la scienza naturale, ribadiamo ancora, non spiega tutto della vita sensitiva, e tanto meno la vita umana. Pretendere il contrario sarebbe come obiettare che gli oggetti matematici sono inaccettabili perché non sono governati dalle leggi della fisica.

Fin qui ci siamo riferiti soltanto agli atti sensitivi. Purtroppo gli autori usano parlano indistintamente delle operazioni *sensitive* e *intelligenti* e nemmeno utilizzano questa distinzione. Questa mancata distinzione è un relitto dell’eredità cartesiana. Così si genera oscurità nel dibattito filosofico, poiché, come ho anticipato, risulta più favorevole alla tesi dell’identità piazzarsi sul piano delle sensazioni, argomentando semplicemente contro il dualismo cartesiano (un bersaglio troppo facile). Le cose sono più problematiche quando si arriva al piano delle conoscenze intellettive (scienze, idee morali, concetti universali), dove la tesi delle correlazioni cerebrali è tutt’altro che chiara.

In questo momento non esaminerò il rapporto tra l’attività spirituale umana e il sistema nervoso. I nostri chiarimenti sull’attività sensitiva, comunque, sono fondamentali. Il riconoscimento di una dimensione fisica superiore (psicofisica), cui la neuroscienza non arriva, è un passo molto

importante, non solo per evitare il totalitarismo scientifico -base del materialismo-, ma anche per poter capire meglio la dimensione spirituale dell'uomo. Anche qui ci allontaniamo, ancora una volta, dal cartesianismo. Infatti, risulta forzato concepire lo spirito come l'abitante di una macchina fisiologica. Invece, se il corpo animale comincia a presentare delle dimensioni più alte, non riducibili a fisica e a chimica, la via verso la comprensione della "spiritualità materiale" della persona umana è già ben predisposta.

Per finire, indicherò *due obiezioni storicamente importanti contro la teoria dell'identità* (ce ne sono state altre e le vedremo dopo). Tutte e due sottolineano l'irriducibilità delle sensazioni soggettive, abitualmente denominate *qualia* dai filosofi della mente. Nessuna conoscenza oggettiva è in grado di sostituirle, per cui non regge l'analogia della riduzione del lampo a scarica elettrica come illustrazione della pretesa riduzione delle sensazioni alla loro base neurologica.

a. Una delle obiezioni è stata sollevata da Thomas Nagel nel suo saggio *What Is It Like to Be a Bat*, del 1974 (*Che effetto fa essere un pipistrello?*). I pipistrelli sono in grado di localizzare oggetti nell'oscurità grazie all'individuazione degli ultrasuoni (sensazioni uditive). Ma noi non abbiamo la minima idea su che cosa "prova" l'animale in questo tipo di percezione. Il mondo delle sue sensazioni resta chiuso per noi. Lo scarto tra la soggettività e la spiegazione neurologica è abissale. Ma non c'è da stupirsi, dal momento che la scienza cartesiana sin dall'inizio si proponeva di seguire un metodo oggettivo, indipendente dal modo in cui esso appare ai nostri sensi.

b. L'altra obiezione, proposta da Franck Jackson nel saggio del 1982 *Epiphenomenal Qualia (I qualia epifenomenici)*, è simile alla precedente. La neurofisiologa Mary, possiamo ipotizzare, sa tutto sul funzionamento del cervello e del sistema visivo. Mary ha studiato profondamente la percezione dei colori, ma resta sempre chiusa in una stanza dove tutto si vede solo in bianco e nero. Quando la neurofisiologa esce dalla stanza e vede per la prima volta i colori, impara qualcosa di nuovo, che prima le era inaccessibile.

Questi argomenti sono ineccepibili. I sostenitori della teoria dell'identità hanno cercato di aggirarli, ma non hanno trovato una risposta soddisfacente. Semplicemente li ignorano.

## 6. Cognitivismo e scienze cognitive

Alla fine della sezione sul comportamentismo avevamo menzionato una recensione critica di Chomsky (1959) al *Verbal Behavior* di Skinner come il colpo di grazia a quella corrente della

psicologia. Sono gli anni in cui si stava producendo nella psicologia degli USA la grande rivoluzione cognitiva, da cui uscirebbe il funzionalismo come nuova tendenza della filosofia della mente. La teoria dell'identità fu invece un capitolo piuttosto breve, dovuto all'impatto del comportamentismo analitico di Ryle e alla forza del fisicalismo carnapiano in molti ambienti della filosofia in Nordamerica.

Non è facile sintetizzare in poche pagine la storia, tuttora in corso, della cosiddetta "psicologia cognitiva". In termini generali, essa nasce come una forma di reazione contro la psicologia comportamentista, rapportandosi in un modo molto naturale all'indagine, sia sperimentale che teorica, dei contenuti cognitivi quali l'attenzione, la memoria, l'elaborazione dell'informazione, il linguaggio e le sue rappresentazioni, con una certa autonomia ma anche con un collegamento con le ricerche neurofisiologiche (anzi, abbiamo pure accennato addietro all'influsso del cognitivismo sulla neurofisiologia, con la nascita di una neuropsicologia cognitiva che si colloca alla stregua dei nuovi modelli cognitivi). Il termine *cognitivo* è dovuto al fatto che questo programma di ricerca si concentra specialmente sui processi della conoscenza (percezione, memoria, rappresentazioni, ragionamenti), anche se posteriormente s'interesserà anche dei fenomeni emotivi.

In quanto programma di ricerca, anziché vera e propria scuola, il "cognitivismo" è molto fluido. Non ci soffermiamo sui precedenti, che in realtà possono essere molto ampi, come la psicologia aristotelica, la teoria del presente e del passato psicologico di William James, la *Gestalttheorie*, o il comportamentismo meno ortodosso di Tolman, il quale ad esempio ipotizzò che i ratti in un labirinto costruirebbero una sorta di "mappa cognitivo" per cercare di uscirne. Anche un autore alquanto isolato dal punto di vista delle correnti ufficiali, come Piaget, potrebbe considerarsi come un valido precedente della linea di ricerca cognitiva, nonché gli psicologi americani del *New Look*, come Bruner (Piaget, Michotte, il *New Look*, da un altro punto di vista sarebbero "postgestaltisti", o "funzionalisti gestaltici").

In questa corrente e anche nelle sue vicinanze direi che possiamo distinguere quattro indirizzi:

a) la *scienza computazionale* porterà a concepire la mente sul modello del calcolatore, da cui usciranno le ricerche sull'*intelligenza artificiale*;

b) il *primo cognitivismo* (psicologia cognitiva) cui è vicino la psicolinguistica di Chomsky (anni '50 e '60);

c) il *secondo cognitivismo* o periodo della "scienza cognitiva", più ampia della psicologia

(fine degli anni '60 e anni '70), dove spicca la figura di Fodor, importante per il funzionalismo filosofico;

d) la linea di ricerca delle *reti neurali* o *connessionismo* (anni '80, anche se riprende accenni degli anni '40).

Più difficile è indicare le linee portanti della psicologia degli ultimi dieci anni, dove si segnala l'importanza crescente, ancora una volta, del campo neurologico (*neuroscienza cognitiva*), e gli sviluppi della *robotica*, con nuovi approcci degli studi sull'intelligenza artificiale, che prima era subentrata in un periodo di relativa stagnazione. Sembrano utili per la psicologia gli sviluppi fisici dei sistemi dinamici non lineari. Si parla di una visione cognitiva *incorporata*, che tiene più conto del corpo e dell'ambiente, superando così il funzionalismo, eccessivamente "astratto" a causa del modello computazionale preso in un modo troppo univoco. In questo senso, Marraffa (*Scienza cognitiva, cit.*) parla di una "scienza cognitiva postclassica", con una maggiore vicinanza alla base biologica e alla situazione ambientale in cui si svolge l'attività cognitiva del soggetto. In qualsiasi caso, nella complessa situazione attuale le tematiche cognitive non si possono studiare senza una molteplicità di prospettive, che continuamente vanno integrate a vicenda.

A. *Dalla teoria dell'informazione alla scienza computazionale.* Il cognitivismo è stato preceduto dalla diffusione, sin dagli anni '30 e '40, della *teoria dell'informazione*, la *cibernetica* e la *scienza computazionale* (C. Shannon, Wiener, Turing, von Neumann). La teoria dell'informazione venne preparata, in un certo senso, dalle tecniche innovative delle telecomunicazioni nel tardo Ottocento e nella prima metà del Novecento. Queste tecniche posero il problema della trasmissione di messaggi espressi in una lingua (sistema di simboli) attraverso impulsi elettrici o segnali in cui venivano codificati (col compito, quindi, di effettuare traduzioni vicendevoli tra il messaggio verbale e la codificazione elettronica). "La nozione di informazione viene così a designare ciò che si mantiene stabile sotto le successive trasformazioni e traslazioni spazio-temporali e, quindi, essa può essere definita operativamente come ciò che denota tali operazioni di selezione, trasmissione e ricezione di una struttura linguistica" (Paolo Aldo Rossi, in *Storia delle scienze*, vol. 2, cit., p. 426).

Al concetto di *trasmissione* dell'informazione si aggiunge quello della sua *elaborazione* ovvero del suo *processamento*. Questo fatto comporta la trasformazione dell'informazione in vista di un determinato fine (ad esempio, per tradurre un testo o calcolare un risultato). L'idea fu applicata alla creazione dei moderni calcolatori (elettronici e digitali, cioè non semplicemente meccanici ed analogici, come prima erano stati concepiti). Nel 1937 H. Aiken costruì ad Harvard, in

collaborazione con la IBM, il primo grande calcolatore elettronico numerico, cui seguirono ben presto tanti altri. Nasceva così la scienza della computazione o del “calcolatore”, che consentiva l’automazione di processi finora riservati alla mente umana (come la semplice operazione di eseguire un calcolo matematico o sintattico).

Il calcolatore era munito di un programma o serie di istruzioni, scritte nel “linguaggio-macchina”, e di una memoria contenente l’informazione. Quindi il calcolatore era in grado di elaborare l’informazione ricevuta (*input*), compiendo determinate operazioni (*processing*), per fornire alla fine una risposta (*output*). Tali operazioni materiali della macchina (stabilire sequenze tra bits), grazie al simbolismo guidato dall’uomo, potevano conformarsi a processi inferenziali, a scelte “intelligenti”. L’output poteva essere anche una risposta motoria (robotica). Di conseguenza, la tecnica computazionale era in grado di simulare i circuiti cerebrali che intercorrevano a partire dai dati ricevuti dai sensi fino ai processi decisionali e alle relative risposte motorie (e non solo simulare, ma anche migliorare certe prestazioni “intelligenti”). Si era così creato un nuovo fronte nel problema del rapporto mente-cervello, a causa dell’analogia del computer, quasi spontanea, con una sorta di “cervello elettronico” (così fu chiamato il calcolatore per alcuni anni).

Nel periodo iniziale della scienza computazionale fu altresì importante la *cibernetica*, scienza proposta da Norbert Wiener (nel 1948 egli pubblicò *Cibernetica ovvero controllo e informazione nell’animale e nella macchina*), che posteriormente si sarebbe fusa con l’informatica. Il punto notevole fu la scoperta di un meccanismo di autocontrollo (controllo automatico) dell’informazione richiesta in un processo teleologico o “intenzionale” (ad esempio, l’autocorrezione degli spari di proiettili nell’inseguimento di un obiettivo mobile). Il meccanismo automatico si basava sul *feedback* o *retroazione negativa*, in virtù del quale i segnali in arrivo sono costantemente confrontati con quelli previsti, per poter così effettuare successive correzioni ai meccanismi di uscita (questo meccanismo di autoregolazione era già stato scoperto da J. Watt nel 1788, per la regolazione meccanica di forze centrifughe). Il *feedback negativo* provoca una diminuzione degli errori (il *feedback positivo*, invece, ha un effetto di amplificazione).

L’autoregolazione cibernetica (*servomeccanismo*), implementata mediante un calcolatore, era un meccanismo di controllo del comportamento di un oggetto mirante a un fine, in base a una forma di processamento dell’informazione. Veniva così suggerita una sorta di “apprendimento” adattativo in un ambiente variabile, con un accumulo d’informazione passata o “memoria”. L’analogia con la condotta “spontanea” degli animali e del cervello umano era quasi scontata, specialmente in collegamento, verso gli anni ‘20 e ‘30 alle tesi del fisiologo Walter Cannon (collega ad Harvard di

Wiener), che aveva proposto nell'ambito fisiologico il concetto di *omeostasi* come meccanismo, a *feedback* negativo, in grado di mantenere l'equilibrio di certi parametri biologici nell'organismo. Wiener, difatti, era in contatto con neurofisiologi come A. Rosenblueth. Il coordinamento motorio dell'uomo, per esempio il movimento delle mani per afferrare un oggetto visto o l'equilibrio del corpo per non cadere mentre va in un autobus o in bicicletta, è chiaramente autoregolativo, e corre a carico del processamento informativo operatosi particolarmente nel cervelletto, in comunicazione con l'informazione sensitiva e con le risposte muscolari.

La cibernetica scatenò la costruzione di automi che imitavano i comportamenti vitali. R. Ashby creò un *omeostato* o congegno elettronico capace di mantenere la stabilità in mezzo a situazioni ambientali variabili. Il neurofisiologo G. Walter progettò "animali sintetici", come tartarughe meccaniche, automi a riflessi condizionati e macchine che imparavano per tentativi ed errori. Dal punto di vista teoretico, si affacciò così l'idea di un'identità strutturale tra macchina, organismo e cervello umano (in quanto la condotta umana è controllata dal sistema nervoso). Ashby, in un articolo del 1948, *Progetto per un cervello*, avanzò l'ipotesi che il cervello sarebbe una macchina omeostatica. Si poteva così pensare che si sarebbe sempre in grado di costruire automi che imitassero (e migliorassero) gli organismi viventi.

Negli stessi anni, nacquero le ricerche sulle *reti neurali*, grazie al lavoro di Warren McCulloch e W. Pitts (cfr. il loro articolo *A Logical Calculus of the Ideas Immanent in Nervous Activity*, del 1943) e del già citato D. Hebb. Nel 1943, McCulloch e Pitts tentarono di simulare il comportamento associativo dei neuroni (applicandovi l'algebra di Boole e l'insiemistica), tenendo conto anche dei loro suoi meccanismi di *on/off* (attivazione ed inibizione). Divenne così possibile la creazione delle reti neurali, che simulano le vie nervose e acquistano e trasmettono informazione. L'idea si discosta dalla cibernetica tradizionale. Ciò che conta primariamente non è il controllo "centralizzato" delle azioni, bensì i circuiti "neurali" presi globalmente. È questa una nuova forma di elaborazione dell'informazione, non basata sul simbolismo che caratterizzava invece la nascente scienza computazionale. La prima ricerca "connessionistica" ottenne alcuni risultato nel riconoscimento di forme percettive e portò, in questa linea, alla creazione di macchine in grado di imparare e di migliorare l'individuazione di forme (*pattern recognition*), come il *Perceptron* di F. Rosenblatt e il *Pandemonium* di Selfridge.

I programmi di ricerca cibernetico e connessionistico (reti neurali) potevano collegarsi con la psicologia e la neurofisiologia, e in questo senso potevano dar luce sulla base fisiologica di concetti psicologici quali l'apprendimento, l'adattamento vitale e l'esperienza basata su associazioni tra

elementi psichici. Il programma connessionistico subì una battuta d'arresto negli anni '50 a causa degli sviluppi delle prime teorie sull'intelligenza artificiale, che si orientarono in una linea simbolica, più lontana dalla neuropsicologia. Le ricerche su reti neurali riprenderanno fiato, però, a partire degli anni '80, ad esempio con gli studi di J. McClelland e D. Rumelhart (cfr. *PDP Microstruttura dei processi cognitivi*, il Mulino, Bologna 1991) (PDP significa *Parallel Distributed Processing*, termine con cui è designata la ricerca sul processamento a rete dell'informazione).

*B. I computer "von Neumann". L'intelligenza artificiale.* La teoria dell'informazione degli anni '40 seguì invece in modo massiccio la strada vincente della creazione dei *calcolatori elettronici*, i cui pionieri nel campo teorico sono stati i matematici Alan M. Turing (inglese, 1912-1954) e J. Ludwig von Neumann (ungherese, 1903-1957). Questi due scienziati possono considerarsi i veri creatori dell'architettura dei computer così come oggi gli conosciamo. I computer secondo "l'architettura di von Neumann" (i nostri computer abituali) sono basati sulla sequenzialità o *serialità*, in quanto operano con una sola sequenza alla volta (le reti neurali, invece, funzionano secondo il principio del *parallelismo*, dove il processamento dell'informazione si compie in maniera distribuita in tutta la rete), e sono anche fondati sul controllo delle operazioni computazionali a carico di un *programma* dotato di istruzioni scritte in un determinato *linguaggio* (col ricorso, quindi, al simbolismo, di cui invece non si vede traccia nel processamento biologico e nervoso dell'informazione, che i modelli a rete imitano).

Il computer tradizionale è, in definitiva, *uno strumento di manipolazione formale di simboli*. Questa manipolazione corre a carico del *software* del calcolatore, come se fosse quasi la sua "anima", mentre l'*hardware* ne costituisce la base materiale, relativamente indipendente dell'implementazione del programma (in futuro, l'*hardware* potrebbe situarsi a livello di particelle quantistiche anziché di processi elettronici).

Il passaggio dal computer all'*intelligenza artificiale* (AI, *Artificial Intelligence*) non è concettualmente discontinuo dalla computazione in generale. Le macchine calcolatrici, infatti, elaborano l'informazione in un modo tale che sin dal principio sembrano sostituirsi a procedure tradizionalmente collegate all'intelligenza, come la semplice realizzazione di operazioni aritmetiche o di deduzione. I termini *computer* o *calcolatrice*, infatti, denotano il *calcolo*, vale a dire *l'operazione di passaggio da una sequenza di simboli a un'altra sequenza di simboli secondo certe regole o istruzioni* (come quando facciamo una somma aritmetica in modo automatico, perché abbiamo imparato ad applicare "ciecamente" certe regole). Si comincia a parlare di "intelligenza artificiale" quando le operazioni computazionali abbracciano sempre più funzioni come la

risoluzione di problemi, il gioco a scacchi, la presa di decisioni in base a dati, la traduzione, la dimostrazione di teoremi, l'apprendimento, l'elaborazione di opere artistiche, ecc., funzioni che normalmente vengono attribuite agli essere umani in quanto intelligenti (riteniamo molto intelligente una persona capace di vincere una partita di scacchi a un avversario difficile). Il nome di IA -ideato da McCarthy- si diffuse sin dal 1956, in occasione di un *workshop* tenutosi a Dartmouth College (New Hampshire), dove si riunirono per la prima volta gli esperti internazionali in questa materia (John McCarthy, Allen Newell, Herbert Simon, Roger Schank e altri).

*C. La macchina di Turing. Limiti del pensiero formale: teoremi di limitazione.* Dal punto di vista teoretico, il nucleo del problema posto alla filosofia della mente con l'invenzione degli elaboratori elettronici e dell'IA sta nella nozione di *macchina di Turing*, proposta dal menzionato matematico Turing negli anni '30. Questa macchina ideale, molto semplice e astratta e mai realizzata veramente, di cui ogni computer è un'esemplificazione, include il concetto di *algoritmo*, inteso come *procedimento che consente di risolvere un problema "automaticamente", seguendo un insieme finito di istruzioni -chiare e univoche- e compiendo un numero finito di passi. Effettuare tale passaggio è computare*. Il termine "automatico" si contrappone qui alla nozione gnoseologica di "intuitivo" (o "eidetico"). Così come noi crediamo di "vedere con la mente" la verità evidente secondo cui  $2+2=4$  (intuizione), diverso è il caso in cui effettuiamo un'addizione ciecamente (senza intuizioni), semplicemente seguendo regole, così come sappiamo che  $8 \times 8 = 64$  perché applichiamo le tavole pitagoriche della moltiplicazione che abbiamo imparato a memoria. Quest'ultimo caso è appunto l'uso di un procedimento algoritmico.

La macchina di Turing (MT) definisce, di per sé (nel suo disegno) le funzioni algoritmiche o di calcolo computazionale. Semplificando a scopo didattico, possiamo dire che la macchina di Turing è come una scatola che riceve certe entrate (simboli) e produce certe uscite, passando così "causalmente" da uno stato ad un altro stato "interno" (che il funzionalismo identificherà come "stato mentale"). Con la lettura dell'input A, la macchina che si trovava nello stato S1 passa allo stato S2.

A questo punto subentra il concetto generalizzato di *macchina universale di Turing*, cioè una MT, sempre ideale, capace di eseguire *tutti* i processi di calcolo algoritmico. Dal momento che l'output della MT può includere l'ordine di compiere una determinata azione materiale, si può pure pensare a un costruttore universale, una sorta di robot computerizzato capace di realizzare "qualsiasi cosa" che sia stata previamente computerizzata. È questa la nozione di "macchina di von Neumann",

o “costruttore universale”, capace di costruire idealmente qualsiasi cosa materiale e di replicare se stessa programmando la sua auto-replicazione.

Alonzo Church (matematico, lavorò a Princeton) e Turing proposero nel 1936 la tesi (“di Church-Turing”) dell’*equivalenza tra la macchina di Turing e qualsiasi sistema algoritmicamente computabile* (in realtà, essa è già inclusa nella nozione di MT). Questa tesi indica il “raggio di azione” delle funzioni computabili e quindi della meccanizzazione della razionalità. Ogni sistema formale finito potrà sempre essere “meccanizzato” in un calcolatore digitale, automatizzando tutte le sue possibili derivazioni.

Tale equivalenza include il *limite della meccanizzabilità della ragione calcolatrice*, dal momento che valgono per la menzionata tesi i limiti del formalismo logico-matematico segnalati nel 1931 dai teoremi di limitazione di Gödel e di altri matematici posteriori, quali Church e Turing. Si conclude che non tutte le funzioni sono computabili. I sistemi formali sono sempre aperti o *incompleti*, cioè all’interno di un dato sistema logico-formale sorgono sempre infiniti enunciati indecidibili (=non derivabili algoritmicamente). A quest’ultima idea punta il *teorema dell’alt o dell’arresto*. I passi finiti della MT suppongono un’ultima istruzione che dichiara la fine dell’operazione computazionale e quindi l’arresto del funzionamento della macchina nell’esecuzione di tale compito. Invece, se la funzione risulta incomputabile, non si produce l’alt e la macchina gira a vuoto, senza produrre alcuna uscita.

*In definitiva, non tutto è automaticamente deducibile, ma solo ciò che può derivare da certe regole date, e le regole possono variare all’infinito. La ragione “calcolatrice” non esaurisce il campo della razionalità, e tanto meno quello della creatività intellettuale (e della stessa realtà, la cui razionalità in un certo senso è “infinita”).* Questo punto non significa che anche la mente umana non sia soggetta ai limiti della computabilità quando lavora in un modo computazionale. Questo è talvolta necessario, dal momento che la nostra intuitività non arriva a tutto (per eseguire una complicata moltiplicazione o divisione aritmetica, ordinariamente non possiamo cercare di “intuire” il risultato. Dobbiamo affidarci alle regole e al simbolismo numerico prescelto). Ma l’intelligenza umana sta al di sopra di ogni regola formalizzata, e proprio per questo l’uomo è capace di inventare indefiniti sistemi formali e di collocarsi all’infuori di tutti (al contrario, è segno di scarsa intelligenza essere troppo automatici nel modo di pensare).

*La MT fece supporre ad alcuni pensatori l’esistenza di un’equivalenza ideale tra il pensiero deduttivo dell’uomo e forse la sua intelligenza “tout court” (confusa con la ragione calcolatrice formale) e le operazioni delle macchine “deduttive” (che includono le realizzazioni degli automi*

cibernetici e le future riuscite tecniche dell'IA). Ancora una volta si affacciava nella storia delle idee la tentazione dell'uomo macchina. Il problema non era soltanto antropologico o sociale, cioè il fatto che l'uomo sembrava "superato" dai suoi stessi prodotti, che un giorno emulerebbero le sue prestazioni intelligenti. Il problema era più profondo. Il pensiero umano sembrava ridursi a ciò che può fare una macchina computazionale. Al riduzionismo della teoria dell'identità (**mente=cervello**) si aggiungeva un nuovo riduzionismo (**mente=computer**), il che poteva far pensare, per completare il quadro, che la mente non sarebbe altro che il cervello in quanto macchina computazionale (**cervello=computer**). Le interpretazioni dell'IA che seguono questa tesi si dicono teorie dell'IA *forte*.

*D. Il test di Turing. Discussione.* Nel 1950 Turing pubblicò in *Mind* l'articolo *Macchine calcolatrici e intelligenza* (*Computing Machinery and Intelligence*, "Mind", 59 (1950), pp. 433-460; ristampato in A. R. Anderson (ed.), *Minds and Machines*, Englewood Cliffs, N. Jersey 1964, pp. 18 ss), con il suggerimento dell'equivalenza tra l'intelligenza umana e le prestazioni del computer, indipendentemente dai risultati concreti. La questione non era tanto "se la macchina poteva pensare", ma piuttosto se la macchina potrebbe fare tutto ciò che fa l'uomo (e molto di più). Egli propose il celebre *test di Turing*: se possiamo fare domande e conversare con un uomo e con una macchina nascosti (e la macchina può ingannare, dicendo di essere un uomo e reagendo emotivamente come una persona umana, anche con eventuali errori), e se dovesse arrivare un momento in cui, in base alle risposte, non ci fosse modo di distinguere tra uomo e macchina, forse si dovrebbe concludere che il pensiero umano non è ontologicamente diverso di quello che fa la macchina (il test è ideale, poiché in pratica l'interrogatorio può essere poco intelligente, e le risposte sia dell'uomo che del computer possono essere molto diverse in qualità).

Il test di Turing è impostato sul versante della *simulazione* e quindi non costituisce, a mio avviso, un autentico problema teoretico. Anche un demone cartesiano potrebbe ingannarci e farci credere, con un ipotetico sogno perfetto, che tutto fosse reale, e non avremmo modo di "argomentare" contro di lui in favore del realismo. La simulazione e l'inganno sono sempre possibili. Il test di Turing sarebbe ancora più forte se, anziché pensare ad una conversazione con un computer, pensassimo ad un robot umanoide che fosse in grado di imitare perfettamente le reazioni umane, del tipo di un "demone" di Cartesio. Invece, nel caso di una conversazione eseguita nelle condizioni del test di Turing, forse la questione è più semplice: non possiamo pretendere di riconoscere una persona ascoltando semplicemente le sue risposte e domande rivolte a noi, senza vedere e toccare il suo corpo.

Potrebbe essere utile qui la risposta di Cartesio agli ipotetici robot indiscernibili dall'uomo, così come egli riuscì a immaginarli, secondo cui l'uomo darebbe risposte sempre *creative* e in ambiti non specifici, ma completamente *universali* (come abbiamo visto nella nostra sezione dedicata a Cartesio)? Queste due caratteristiche sono originarie dell'uomo, è vero, ma non servono per discernere la presenza di un uomo o di una macchina in una conversazione, che è sempre finita. Solo a lunga scadenza, ad esempio, in conversazioni che durassero anni, o magari guardando le realizzazioni dell'umanità, confrontate con quelle di molte generazioni di computer "intelligenti", si potrebbe finalmente intuire quando siamo davanti a una vera persona e non di fronte a una macchina (questa risposta vale per un'IA ideale, che veramente non esiste). In questo senso, forse è parzialmente valida l'argomentazione secondo cui il test di Turing è male impostato, poiché si colloca in una cornice "comportamentistica", in quanto considera soltanto i risultati e non gli atti interiori. Però, si ricordi che il test in questione è proprio una sfida all'interiorismo. Quindi, per situare la nostra risposta nel terreno della sfida proposta, non serve appellarsi all'esperienza interiore. Inoltre è vero che noi riconosciamo le persone "senza telepatia", osservando il loro comportamento personale e razionale. Perciò, ritengo che la risposta data nel paragrafo di sopra sia sufficiente.

Il problema si potrebbe anche dividere in due aspetti:

a) Aspetti *informativi* e di pensiero in materie di *calcolo razionale* (ad esempio, gioco di scacchi, teoremi matematici): in questo campo, un computer dotato di una sofisticata IA potrebbe alla fine superare sempre l'uomo, poiché esso può avere molta più informazione e una maggiore potenza di calcolo di quella che sta alla nostra portata. Quindi, anche se un computer rispondesse molto bene a ogni tipo di domanda relativa a quelle due sfere cognitive, ciò non significherebbe niente di speciale. Anzi, per "sembrare" umano, egli dovrebbe perfino "sapere un po' meno" e forse sbagliare ogni tanto.

b) Aspetti *umani*, come credenze metafisiche, religiose, morali, o atteggiamenti esistenziali. In questo campo un computer dotato di IA non è competente (il suo vero "raggio di azione" è solo il "pensiero computazionale"). Ma è sempre pensabile, almeno in teoria, che l'IA sia preparata per *imitare* queste dimensioni dell'uomo. Allora non ci sarà mai una determinata risposta "speciale" che sia "*propria* dell'uomo" e che nessun computer potrebbe imitare. Tutto si può imitare.

Possiamo segnalare qui due questioni:

1) Come riconosciamo *una persona umana*? Una risposta completa a questa domanda sarebbe

un intero capitolo della gnoseologia. In poche parole: riconosciamo ogni tipo di ente (viventi, animali, esseri umani) in base a una quantità complessiva di elementi convergenti ed unitari, colti nell'esperienza intelligente. L'inganno nel riconoscimento di forme è sempre possibile, ma ciò non significa nulla.

2) Come riconoscere quando siamo di fronte ad *un'intelligenza umana*? Risposta: solo quando vediamo che tale intelligenza è esercitata da una persona umana (anche se è una modesta intelligenza). Separato dalla persona, ogni processo razionale è imitabile da una macchina. Ma ricordiamo che il "pensiero delle macchine" lo abbiamo fatto noi, esseri umani, per cui, nella sua radice, è "il nostro pensiero".

*E. I limiti sostanziali dell'intelligenza artificiale.* L'equivalenza uomo-macchina comporta la riduzione del pensiero umano al calcolo computazionale, il quale di per sé è automatico e finito (l'intelligenza è così ridotta a una funzione della ragione calcolatrice). Senz'altro, i risultati dell'IA sono stati sbalorditivi (eseguire calcoli matematici di portata astronomica, vincere negli scacchi ai migliori campioni, ecc.). Questo fatto alimentò in modo pratico la tentazione del riduzionismo della mente al computer. Ma è naturale che nel terreno computazionale le "macchine intelligenti" trascendano la capacità intuitiva dell'uomo "a piedi", così come l'automobile, creazione tecnica dell'uomo, supera in velocità la nostra capacità fisica di correre. Gli strumenti superano *secundum quid* il creatore umano, ma in definitiva qui siamo davanti all'uomo che supera se stesso. Gli strumenti sono soltanto strumenti, creati ed usati dall'uomo (se questi non può controllarli, è perché li progetta o li usa male).

L'IA risulta utile in numerosi campi concreti dove è necessaria una certa computazione, organizzazione e razionalizzazione del pensiero: problemi organizzativi e quantitativi, soluzioni di problemi tramite inferenze, prove di teoremi matematici, traduzione, programmazione automatica, apprendimento mediante strategie di inferenza analogica, consigli "esperti" in campi dove occorre la trattazione di una grande quantità di dati (diagnosi medica, tecnica bancaria, giurisprudenza, analisi chimica, ecc., con la possibilità di interagire con l'utente, come avveniva nei sistemi esperti), robotica (con riconoscimento di forme sensibili), o giochi. Tramite algoritmi euristici, l'IA può anche manifestare una relativa adattabilità e creatività, con risultati che l'uomo non può prevedere.

Ma osserviamo i seguenti punti:

1) Le regole, i programmi, gli algoritmi, *sono stati creati dall'uomo* e introdotti da lui nei congegni elettronici. Tali regole creano una sorta di "mondo predefinito e ristretto", che tocca alla

persona umana interpretare e usare convenientemente. Anche la nostra mente può talvolta subire queste restrizioni (ad esempio, pensando quasi “come un computer”), ma in fin dei conti l’uomo possiede sempre la libertà di organizzare in modi infiniti le questioni, restando libero dagli schemi oggettivi da lui creati. L’uomo sta fuori dai sistemi cognitivi che egli può configurare di volta in volta nella cultura e nella tecnica.

In questo senso, secondo Popper, non basta segnalare che un computer è in grado di eseguire *qualsiasi compito specificato che possa fare l’uomo*. Proprio questo non bisogna fare, cercare cioè di *specificare* una prova che possa essere superata soltanto dall’uomo e non dalla macchina, come se ci fosse un “particolare” tipo di attività che definisca l’intelligenza umana (cfr. *The Open Universe. An Argument for Indeterminism*, a cura di W. W. Bartley III, Hutchinson, Londra 1982, nn. 22-23). Invece è proprio dell’uomo la creatività intellettuale, imprevedibile e non specificabile. D’altronde, rileva Popper, se le macchine superano di gran lunga l’uomo nel potere di calcolare, soltanto l’uomo le costruisce e distingue tra conoscenze rilevanti e triviali.

2) La macchina pensante *non conosce propriamente fini, valori, principi*, le cose essenziali, i contesti o le priorità. Tutto ciò, semmai, le viene dato con la programmazione, che sempre può essere cambiata o migliorata in funzione di nuove finalità. Poi, tocca all’utente adeguare lo strumento computazionale ai contesti infiniti della vita umana e metterlo al servizio delle priorità volute dalla persona. Può anche non farlo, ma ciò sarebbe un deficit della nostra intelligenza naturale.

Osserva in questo senso Winograd che le persone umane, quando usano un computer, “usano le strutture del loro linguaggio naturale per interagire con un sistema che non comprende il linguaggio ma è abile per manipolare alcune delle sue strutture. In secondo termine, le risposte [del computer] riflettono una particolare rappresentazione che è stata creata da un gruppo di persone, e incorpora una cecità di cui perfino i costruttori non sono del tutto consapevoli” (Terry Winograd e Fernando Flores, *Understanding Computers and Cognition*, Ablex Pub. Corporation, Norwood (N. Jersey) 1987, p. 124). Winograd non è un filosofo teorico, bensì uno dei massimi realizzatori di programmi di IA presso il MIT e l’Università di Stanford. In quest’opera egli si rifà, insieme a Flores, a elementi di Heidegger, Wittgenstein e di filosofia ermeneutica.

3) L’IA dell’architettura di von Neumann è una “macchina simbolica” e quindi *tocca all’uomo interpretarne il simbolismo*, incorporandolo alla sua comprensione essenziale delle cose, sia teoretica che pratica. Ancora Winograd: “i computer non esistono, nel senso di cose che posseggono caratteristiche e funzioni oggettive, all’infuori del linguaggio” (*ibid.*, p. 78).

*F. Aspetti storici dell'IA.* Dal punto di vista storico, gli sviluppi dell'IA non sono stati così utopici come i primi momenti potevano far pensare (cfr. Jean-Gabriel Ganascia, *L'âme-machine. Les enjeux de l'intelligence artificielle*, Seuil, Parigi 1990). Nei primi anni (gli anni '50), anni caldi e molto promettenti, si pensò a meccanismi computazionali generali (ad esempio, nel programma conosciuto col nome di *General Problem Solver*), che sarebbero stati capaci di risolvere ogni tipo di problema logico, di dimostrare quasi tutto nel campo della matematica, di tradurre automaticamente e di riempire il mondo di robot.

Ma i risultati reali, poveri in certi campi, dimostrarono i limiti di questi programmi. È vero, ad esempio, che nel campo degli scacchi l'uomo si dimostra un rivale sempre più debole nei confronti di macchine potentissime programmate per questo gioco. Ma il gioco degli scacchi è situato in un "universo di oggetti" chiuso e molto univoco, lontano dalla vita reale, che è infinitamente più complessa. In matematica i computer possono eseguire dimostrazioni sorprendenti, ma in contesti formali preparati dall'uomo, e finora non si è visto che la computazione abbia portato un progresso creativo nella storia della matematica. Come osserva Penrose, ciò che conta in matematica non sono gli algoritmi, ma la scelta degli algoritmi giusti. "Si può prendere l'avvio da alcuni assiomi, dai quali si devono derivare varie proposizioni matematiche. Quest'ultimo procedimento potrebbe essere in effetti algoritmico, ma per decidere se gli assiomi siano o no appropriati occorre il giudizio di un matematico cosciente" (*La mente nuova dell'imperatore*, Rizzoli, Milano 1992, p. 522, e più ampiamente pp. 517-537). L'algoritmizzazione può essere resa automatica e quindi può appartenere alle macchine, ma al di sopra del pensiero deduttivo di tipo "meccanico" c'è il pensiero comprensivo e intelligente in un senso profondo. I matematici creativi non pensano computazionalmente. "La decisione circa la validità di un algoritmo -scrive Penrose- non è di per sé un processo algoritmico" (*ibid.* p. 524).

Dopo la seconda metà degli anni '60, il programma di ricerca dell'IA si rinnovò grazie ai *sistemi esperti*, un campo che finalmente riuscì ad entrare nel mercato. La generalità primitiva dell'IA fu abbandonata. I sistemi esperti compiono in certi campi (medicina, farmaceutica, geologia) i ragionamenti di uno specialista. Si svilupparono negli anni '70 e '80 seguendo l'architettura tipo *se...allora* nei loro algoritmi. Accumulano conoscenze in un ambito e stabiliscono percorsi adeguati nella risoluzione di problemi o questioni rilevanti in quel campo. I sistemi esperti ebbero successo, ma sono specifici per aree locali della conoscenza nelle quali si richiede una notevole competenza. Debbono accumulare una grande quantità di informazione, da dove nasce il problema della *rilevanza* e del *senso comune*, dato che tali conoscenze tendono ad essere enciclopediche e debbono aggiornarsi e risistemarsi continuamente, e vanno inoltre integrate in un

mondo reale dove ci sono numerosi presupposti ontologici e antropologici.

I limiti del pensiero computazionale, nonostante questi progressi, emergono dalla complessità del mondo reale e dalla fluidità del linguaggio, che va sempre interpretato in contesti semantici e pragmatici. Il linguaggio computazionale, invece, per essere utile dev'essere molto preciso. Quanto più complessa è una realtà, più bisogno abbiamo di un pensiero essenziale, comprensivo e qualitativo, e minore è il controllo ottenuto attraverso il puro trattamento computazionale dell'informazione. *La computazione è vincente soprattutto nelle competenze logico-formali, quantitative e organizzative.* Risulta invece meno utile nella politica, nella filosofia, nell'etica, nella formazione delle persone, campi in cui non basta l'informazione e la deduzione, ma sono in gioco valori, scelte di libertà, scambi tra gli esseri umani e criteri di comprensione essenziale (metafisica e antropologica) della realtà.

Dopo gli anni '80, la ricerca nell'IA cominciò ad affievolirsi nuovamente, e lo stesso nome di IA venne meno usato. Crebbe invece l'interesse per le reti neurali e per la base neurocognitiva del pensiero umano, mentre il progresso nella computazione si produce in aree più specifiche, ad esempio nella nuova robotica (R. Brooks, *Cambrian Intelligence. The Early History of the New AI*, MIT Press Cambridge (Mass.) 1999; A. Clark, *Being There. Putting Brain, Body, and World Together Again*, MIT Press, Cambridge (Mass.) 1997), o nell'interazione tra agenti intelligenti virtuali.

Il nuovo paradigma dell'IA non impiega computer "isolati" dall'ambiente e "individuali". L'intelligenza artificiale si distribuisce in una rete di molte "menti artificiali" o "agenti intelligenti" interattivi (non solo con l'uomo, ma anche tra di loro), situati in un certo ambiente (ad esempio, un ospedale, un aeroporto, un'azienda) sufficientemente conosciuto, da cui ricevono *input*, potendo "reagire" con adeguate risposte. Vengono così creati "sistemi multi-agenti", nei quali ogni "agente intelligente" (artificiale) è dotato delle seguenti caratteristiche (prendo questi punti da una conferenza presso la Pontificia Università della Santa Croce di Jordi Sabater Mir, esperto in IA e ricercatore nell'area dei sistemi multi-agenti):

a) *Autonomia*: funzionamento relativamente autonomo rispetto all'uomo, come ogni macchina, con un certo controllo sulla sua attività e "stati funzionali".

b) "*Socializzazione*": capacità di interagire con altri agenti intelligenti.

c) *Reattività*: capacità di percepire aspetti dell'ambiente, e di reagire opportunamente.

d) *Pro-attività*: capacità di “prendere iniziative teleologiche”, con domande ad altri agenti, esplorazioni in Internet, ecc. e proposte in tempi opportuni all’utente.

Ulteriori caratteristiche, non completamente essenziali, sarebbero la *capacità di apprendimento* (in una certa area ad ambiente), e l’*adattamento* alle variazioni dell’ambiente. Se il sistema multi-agente è robotizzato, va aggiunta la *mobilità* dell’agente. In definitiva, l’agente intelligente lavora in un ambiente e risolve problemi in un determinato campo. “Percepisce” attraverso *sensori*, e compie lavori o dà risposte attraverso *effettori*. Il nuovo paradigma, come nei sistemi esperti, non ha la pretesa di arrivare ad una IA “universale”. Ma è diverso dai sistemi esperti, a causa della nuova dimensione “sociale” (lavoro in rete con altri agenti) e della reattività e “iniziativa” in un ambiente conosciuto. Possono essere utili in ogni campo del lavoro dell’uomo (organizzativo, tecnico, scientifico, commerciale, medico, ecc.).

Sul tema dell’IA, cfr. E. AGAZZI, *Operazionalità e intenzionalità. L’anello mancante dell’intelligenza artificiale*, Jaca Book, Milano 1989; J. COPELAND, *Artificial Intelligence*, Blackwell, Oxford 1993; T. DE ANDRÉS, *Homo cybersapiens, La inteligencia artificial y la humana*, Eunsa, Pamplona 2002; H. DREYFUS, *What Computers Cant’do: The Limits of Artificial Intelligence*, Harper and Row, New York 1979 (*Quel che i calcolatori non possono fare*, Armando, Roma 1988); H. L. DREYFUS e S. E. DREYFUS, *Mind over Machine. The Power of Human Intuition and Expertise in the Era of the Computer*, The Free Press, New York 1988; PH. JOHNSON-LAIRD, *La mente e il computer. Introduzione alla scienza cognitiva*, il Mulino, Bologna 1990; J. SEARLE, *The Rediscovery of the Mind*, MIT Press, Cambridge (Mass.) 1992; R. SOKOLOWSKI, *Natural and Artificial Intelligence*, in *The Artificial Intelligence Debate*, ed. S. R. Graubard, MIT Press, Cambridge (Mass.) 1988, pp. 45-64); T. WINOGRAD e F. FLORES, *Understanding Computers and Cognition*, Ablex Pub. Corporation, Norwood (N. Jersey) 1987.

G. *Il primo cognitivismo*. Dopo il nostro *excursus* nella scienza computazionale, vediamo adesso gli sviluppi storici della psicologia cognitiva (cfr. per un quadro storico, M. Gardner, *La nuova scienza della mente*, Feltrinelli, Milano 1988, traduzione di *The Mind’s New Science*, 1985).

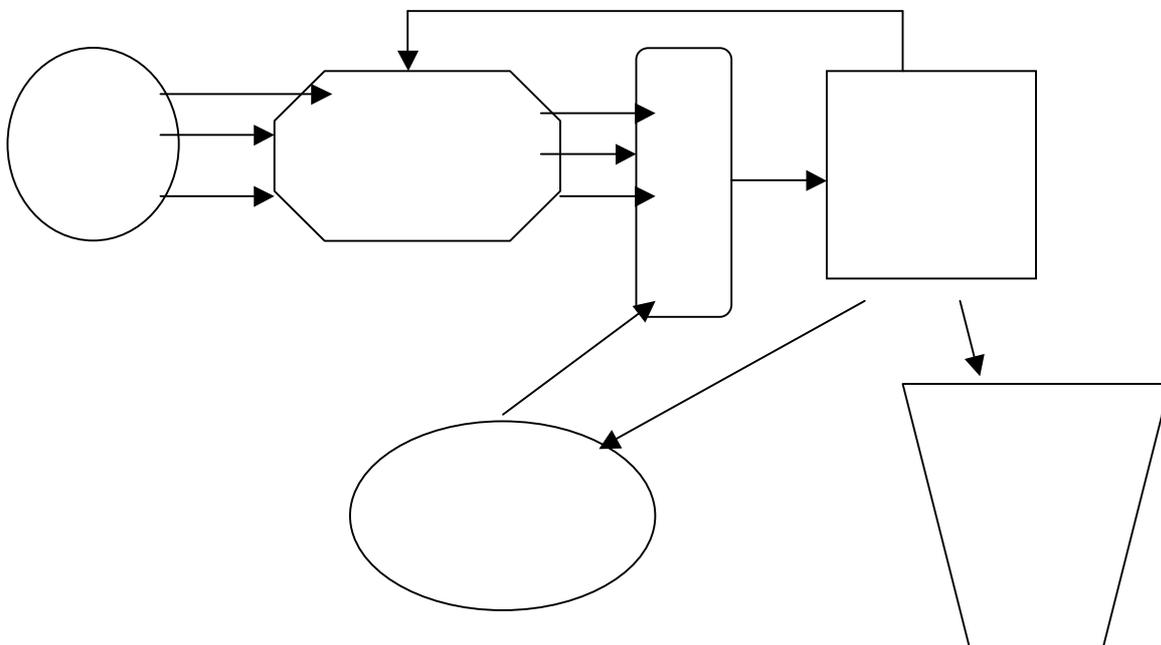
Il primo periodo della psicologia cognitiva si estende sin dalla fine degli anni ‘50 sino al termine degli anni ‘60. Una pubblicazione significativa a questo riguardo è stata *Cognitive Psychology* (1967), di U. Neisser, opera che diede nome alla nuova corrente. Nel 1970 nacque la rivista *Cognitive Psychology* e nel 1971 apparve la rivista *Cognition*.

In contrapposizione alla scuola comportamentista, il cognitivismo vede la mente umana come un insieme di sottounità interne tramite le quali l’informazione procedente dagli stimoli sensoriali viene processata. Per Kenneth Craik (cfr. *The Nature of Explanation*, 1943), uno degli iniziatori di questa linea di ricerca, il pensiero si costruisce un modello per rappresentarsi in scala ridotta la realtà e le proprie azioni possibili. La mente umana è in grado di elaborare simboli e di metterli in relazione, sia per esprimere la realtà che per passare all’azione. Questa prospettiva, analoga agli studi cibernetici e computazionali, si concentra quindi sul concetto di *processamento*

di dell'informazione mediante simboli (human information processing: HIP), con una certa indipendenza dagli elementi fisiologici che possono stare alla base dei modelli psicologici.

Numerosi studi psicologici empirici seguirono questa linea, sin dagli inizi degli anni '50. Utilizzando test in laboratorio, furono osservati, ad esempio, i tempi di reazione di certi compiti proposti, da cui sono emersi alcuni modelli teorici. Vengono così descritti i “flussi di informazione” che intercorrono attraverso la mente (con appositi “diagrammi di flusso” o *flow charts*).

Ecco un esempio:



I ricercatori presentarono modelli cognitivi fondati sull'idea di fasi successive lungo l'esecuzione delle operazioni cognitive: analisi fisica degli stimoli, selezione dell'informazione rilevante (compiuta da “filtri”), attribuzione di significato all'informazione filtrata e memorizzazione a diverse fasi. Una novità del cognitivismo fu la proposta, su base empirica e da confermare posteriormente dal punto di vista neurologico, di *diversi tipi di memoria* (vista all'inizio tramite la metafora del *magazzino*): memoria *iconica*, memoria *a breve* e a *lungo termine*, memoria *di lavoro*, memoria *dichiarativa (episodica o semantica)* e memoria *procedurale* (cfr. D. E. Broadbent, *Perception and Communication*, 1958. Altri autori su questa linea sono stati A. T. Welford, R. C. Atkinson, R. M. Shiffrin, W. E. Hick, G. Miller).

Il primo cognitivismo studiò empiricamente la *capacità limitata di trasmissione* di un determinata quantità d'informazione lungo i “canali informativi” della mente. I “filtri” selezionano i flussi in arrivo. La quantità d'informazione trasmessa può variare secondo l'intensità dell'attenzione, gli interessi, l'esistenza di “silenzio” o “rumori” e altri fattori. La capacità limitata di assimilare informazione si riflette sperimentalmente nei *tempi di reazioni*. La “legge di Hick” appunto mette in rapporto la quantità di informazione che un soggetto può elaborare con i tempi di reazione. Il numero 7, secondo *Miller*, sarebbe il limite di oggetti che la mente cosciente può ritenere alla volta (7 più o meno 2). Così, possiamo processare alla volta (e ricordare) al massimo 7 parole, 7 numeri, 7 lettere. Tuttavia, il nostro linguaggio è capace di concentrare in una sola parola una grande quantità di informazione, superando così i limiti sopraindicati (poiché una parola include una costellazione di significati). Secondo l'immagine di *Miller*, la nostra mente è come un borsellino che può contenere soltanto 7 monete, ma le monete possono essere da un centesimo o da un dollaro (cfr. R. Luccio, *La psicologia: un profilo storico*, p. 180).

Un'altra prospettiva del primo cognitivismo mira al comportamento in una maniera olistica. *G. A. Miller*, *E. Galanter* e *K. Pribram*, nell'opera *Plans and the Structure of Behavior* (1960), sostituiscono la vecchia “unità di analisi” del comportamentismo (S-R: stimolo e risposta), con l'unità TOTE (*test-operate-test-exit*). L'agente opera secondo un piano di comportamento per arrivare a uno scopo previsto, da cui segue una prova destinata a verificare la congruenza tra la realtà esterna e lo scopo (ad esempio, voglio fare un viaggio, quindi cerco un mezzo di trasporto). Se tale congruenza non si trova, si procede ad un'operazione mirante a superare la situazione. Quando la congruità è accertata, si passa all'uscita comportamentale. L'azione dell'individuo non viene vista esternamente, come faceva il comportamentismo, ma dal punto di vista del piano operativo dell'agente. Questo fatto costringe all'analisi delle informazioni che la mente deve processare per pianificare la condotta.

Un ulteriore versante del cognitivismo procede dal rinnovamento della linguistica operato da *Noam Chomsky* con la critica a *Skinner* già menzionata (1959), nonché la sua opera *Syntactic Structures*, pubblicata nel 1957. I parlanti nativi, anche da bambini, dominano la struttura della loro lingua come un tutto. Essi sono sempre in grado di riconoscere in modo “procedurale” (senza poterlo in modo esplicito) quando una frase possiede un senso sintattico, pur essendo assurda nel contenuto (ad esempio, “i fiori combattono contro i numeri” è sintatticamente corretta, mentre “di strano il cade” non ha senso grammaticale). Il linguaggio verbale, conclude *Chomsky*, è un comportamento specifico dell'uomo, fondato su strutture biologiche innate (del sistema nervoso).

Possiamo distinguere tra la *competenza* e l'*esecuzione* nel comportamento linguistico. La competenza non è la pura quantità di segni accumulati nella memoria, ma il saper produrre all'infinito frasi corrette (*produttività*), in quanto la mente possiede regole per *generare* frasi ben formate e per *trasformarle* in diverse forme (voce attiva e passiva, affermazione e negazione, ecc.). La teoria linguistica chomskiana, che è un capitolo dello strutturalismo, è denominata, infatti, *grammatica generativa e trasformativa* (con l'idea di una *Grammatica universale* della mente umana). Così Chomsky si spiegava la sbalorditiva capacità dei bambini di imparare perfettamente la lingua in un modo spontaneo in solo due o tre anni.

La linguistica chomskiana indicava un *a priori* mentale, muovendosi quindi in una direzione antiempiristica. La sua opera segnò la morte del comportamentismo. In sintonia con il cognitivismo, Chomsky portò l'attenzione sulla "scatola nera" della mente. La nostra mente gode di una padronanza indeterminata di una struttura linguistica geneticamente data, quindi in modo universale per tutti gli uomini. Tale struttura è la base per poter imparare bene le diverse lingue storiche. Siamo portati, dunque, a riconoscere l'esistenza di *rappresentazioni* mentali (considerate nella prospettiva linguistica). "Una persona che parla una lingua ha sviluppato un certo sistema di conoscenza, dotato di qualche rappresentazione all'interno della mente" (Chomsky, *Linguaggio e problemi della conoscenza*, il Mulino, Bologna 1991, p. 5). Da qui nasceranno, in psicologia cognitiva, le "teorie rappresentazionali" non solo nel campo linguistico, ma anche nel campo della visione, l'immaginazione, la memoria e il ragionamento.

Dopo la psicolinguistica chomskiana, le ricerche si sono orientate sul versante sia linguistico che psicologico. Questi ultimi studi cognitivisti, più vicini alla neurolinguistica, si concentrarono sulla *comprensione* del linguaggio (ascolto e lettura) e sulla *produzione* linguistica (parlare e scrivere). Gli sviluppi della psicologia del linguaggio, anche arricchiti con i modelli connessionisti, sono tuttora in corso e hanno davanti a sé un ampio orizzonte di ricerca.

*H. Nel secondo cognitivismo* i principi e la metodologia della nuova linea psicologica sono consolidati, come si può vedere nell'ulteriore opera di Neisser, *Cognition and Reality* (1976). Cresce la consapevolezza dell'approccio interdisciplinare della cognizione, per cui si comincia a parlare, in generale, di "scienze cognitive". Come risultato di alcune autocritiche, la psicologia passa adesso a occuparsi non soltanto di microprocessi cognitivi "di laboratorio", ma di un campo "macroscopico" più tipico della vita quotidiana. Nel 1977 venne fondata la rivista *Cognitive Science* e nel 1979 si tenne il primo convegno della *Cognitive Science Society* (La Jolla, California). Lo sviluppo delle scienze cognitive fa parte della storia della psicologia della seconda metà del XX

secolo negli Stati Uniti e in Inghilterra, luoghi centrali, anche se non assolutamente esclusivi, della visione storica che stiamo affrontando in queste pagine (cfr. Michael W. Eysenck, Mark T. Keane, *Cognitive Psychology. A Student's Handbook*, 1990; R. Luccio, *La psicologia: un profilo storico*, cit., pp. 206-222).

James J. Gibson (1904-1979), “vecchio” psicologo di formazione fenomenologica e gestaltistica criticò l’indirizzo cognitivista, con una visione “ecologica” della mente, più attento ai rapporti con l’ambiente. L’anno in cui morì apparve la sua opera più influente, *An Ecological Approach to Visual Perception*<sup>1</sup>. Un altro psicologo vicino alle sue tesi è lo svedese Gunnar Johansson, studioso della psicologia della percezione degli eventi (anziché di percetti statici), tramite il *motion analysis*.

Con la critica al primo cognitivismo (paradigma HIP), Gibson provocò la crisi che portò il movimento alla seconda fase. In una prospettiva di *realismo immediato*, Gibson ignora le rappresentazioni e i flussi informativi della mente. Il mondo ci si offre direttamente, con i messaggi luminosi provenienti dall’ambiente in moto (“assetto ottico”: *optic array*). Il mondo esterno è dato alla percezione diretta, una percezione esplorativa (non passiva), capace di coglierne le invarianze ambientali, nonostante i movimenti del nostro corpo. Johansson, a sua volta, sottolineò il ruolo del movimento nel riconoscimento di oggetti.

Secondo Gibson, l’ambiente circostante si presenta al soggetto in forma di certe disponibilità (*affordances*) in rapporto allo stadio evolutivo del soggetto. La sua impostazione è “antirappresentazionista”: vedere il mondo non è come vedere una fotografia. L’*affordance* (neologismo da lui coniato; *to afford*: offrire) è ciò che l’ambiente “fornisce” agli animali, qualcosa di reale ma significativo -positivamente o negativamente- per l’organismo (ad esempio, una superficie terrestre rigida e orizzontale “offre” -*affords*- un appoggio all’animale). La percezione coglie le cose, primariamente, non nelle loro qualità proprie ma nelle loro *affordances* per l’animale. Comunque non è esatto pensare, come suggerisce Gibson, che il soggetto percettivo non compie alcune elaborazioni delle informazioni ricevute, come hanno dimostrato le ricerche di David Marr -uno dei più importanti psicologi cognitivi- sulla percezione o *pattern recognition*.

La psicologia della percezione di Gibson ricorda la tesi tomista secondo cui la conoscenza intenzionale non ha come oggetto la rappresentazione come *oggetto quod*, ma solo a titolo di mediazione “silenziosa” per cui (*quo*) attraverso la quale la realtà è direttamente colta. L’*affordance* potrebbe essere integrata negli oggetti della facoltà “cogitativa” (o “estimativa”) tomista, in virtù

della quale le cose vengono colte nei loro significati pratici. Senza dare all'*affordance* un'interpretazione kantiana, come qualcuno ha proposto, direi che tale nozione è in armonia con il "realismo moderato", secondo cui il mondo ci si presenta in conformità con il modo di vedere le cose proprio della natura umana (nella gnoseologia tomistica, il *modus cognoscendi* non è identico al *modus essendi*). Il realismo cognitivo è compatibile con un certo "relativismo prospettico della percezione". Con la psicologia di Gibson, chiaramente non si cade nell'idealismo o nell'immanentismo, poiché si sostiene che conosciamo primariamente la realtà, non le nostre rappresentazioni. Ma la conosciamo dal nostro punto di vista, pur arrivando all'accertamento di alcune invarianze percettive.

Il secondo periodo del cognitivismo ampliò la panoramica di ricerca al mondo dei *concetti*, stimolato dalle nuove prospettive linguistiche e dalle tecniche dell'IA. Non continuò, in questo senso, l'immediatismo gibsoniano, bensì al contrario, sviluppò l'uso della categoria di *rappresentazione*, senza l'elementarismo empirista criticato da diverse istanze (come Gibson e Chomsky). Se il primo cognitivismo fu importante per la psicologia della percezione, dell'attenzione e della memoria, il secondo cognitivismo contribuì alla psicologia delle rappresentazioni concettuali. L'isomorfismo con i processi computazionali continuò ad essere sostenuto, spesso a livello di analogia e non necessariamente in un senso ontologico forte. Gli influssi vicendevoli tra la psicologia cognitiva e la ricerca di nuovi modelli di IA fu intensa negli anni aurei del cognitivismo.

La psicologia dei concetti fu dapprima approfondita con i modelli di *reti semantiche* (M. R. Quillian, E. Loftus, A. Collins, anni '60 e '70). Furono presentati, in questo senso, quadri logici che potevano rammentare le classificazioni dialettiche platoniche e aristoteliche (ad esempio: il canarino appartiene alla categoria di uccello, e le sue proprietà sono che canta, è giallo, ecc.), viste in un senso psicologico, come modi in cui le persone "categorializzano" gli oggetti in forma di "reti aperte", dove gli attributi e categorie non hanno sempre lo stesso "peso". Il lavoro qui consiste nella scomposizione "dinamica" (non logica, ma psicologica) di un concetto in sottoconcetti, nella sua collocazione in quadri più universali e nell'identificazione di rapporti con altri concetti. Il quadro generale dell'organizzazione dei concetti venne poi ampliato mediante la nozione di *schema* (R. Schank, U. Neisser, D. E. Rumelhart) o modello generale di una nozione (schemi, ad esempio, di "casa", "giocare al tennis", "leggere un libro", ecc., i quali possono includere, a loro volta, una serie di "sottoschemi"). Questi studi indagano su quanto la filosofia tradizionale conosceva sotto il nome di "astrazione", con i loro quadri categoriali, comunque senza chiusure "essenzialistiche".

---

<sup>1</sup> Cfr. J. J. Gibson, *Un approccio ecologico alla percezione visiva*, Il mulino, Bologna 1999.

All'interno della teoria degli schemi sorse il concetto di *script* (R. Schank, R. Abelson. *Script* significa *copione*, termine preso dal linguaggio cinematografico), che è lo schema di una situazione tipica a certe variabili o *slots*, riempite secondo i casi. Ad esempio, "andare al ristorante" è uno *script* che include azioni successive come guardare il menu, chiamare il cameriere, pagare alla fine dei pasti, ecc. Simili agli *scripts* sono i *frames* (cornici, scenari) di M. Minsky. Ad esempio, se faccio un viaggio a Firenze, nella mia mente si attiva il *frame* di viaggio, che include una serie di azioni stereotipate, prese per default in mancanza di altre connotazioni. Il *frame* di "festa di compleanno di un bambino" include cose come regali, dolci, torta, canzoni, giochi, ecc. Alcuni *frames* sono linguistici, come il frame "c'era una volta un principe...", che attiva l'aspettativa di ascoltare un racconto (cfr. Marvin Minsky, *La società della mente*). Un'altra proposta di rappresentazione cognitiva fu la teoria dei *modelli mentali* di Ph. Johnson-Laird. I modelli mentali contengono elementi e relazioni di un oggetto, rilevanti per le intenzioni, scopi, competenze e contesto dell'agente che li usa. Su questa base, quest'autore ha studiato certi aspetti psicologici del ragionamento, collegati all'impiego di modelli mentali (cfr. *Mental models*, 1983; cfr. *The computer and the mind*, 1988).

Le teorie degli schemi, *scripts*, *frames* e altre simili categorie concettuali rappresentarono un vero progresso nel campo della psicologia (con connotazioni utili anche per la teoria della conoscenza). In molti casi queste elaborazioni furono impiegate come modelli di IA (così per i *frames* di Minsky e gli *scripts* di Schank). Ma esse servirono pure in linguistica (significati e uso delle frasi), nella concezione della memoria semantica (o "dichiarativa") e in psicologia evolutiva. In queste ricerche c'è qualche continuità con quanto fu guadagnato negli studi di Piaget. Rumelhart propose un trattamento connessionistico degli schemi, rendendoli più fluidi a seconda delle connessioni più o meno attivate (in questo modo si possono studiare, ad esempio, i collegamenti a rete tra elementi di descrizione quali porta, muro, piccolo, ecc., e oggetti quali cucina, bagno, stanza da letto, ecc).

L'insieme degli elementi concettuali, con i diversi tipi di memoria (semantica, procedurale, di lavoro) e le regole che consentono di compiere inferenze e così passare all'azione, in funzione degli *input* ricevuti, viene spesso considerato come un'*organizzazione della conoscenza* ad ampio respiro, organizzazione dinamica denominata anche *architettura della mente*. Spesso questi quadri sono stati usati nella ricerca dell'IA (ad esempio, nei sistemi esperti), ma sono utili anche per la psicologia del ragionamento e delle decisioni. Seguendo una modalità usata in computazione, il passaggio dell'agente all'azione di solito è basato sull'esistenza di un *sistema di produzione* (sistema cognitivo di passare alla produzione di frasi e di comandi per l'azione). Il sistema di produzione

contiene “regole di produzione” del tipo *se...allora* (“se c’è un semaforo rosso, allora mi fermo”). Ancora qui la ricerca ha un risvolto computazionale e un altro psicologico.

La psicologia cognitiva portò avanti altresì ricerche sulle rappresentazioni in quanto *immagini* delle cose rappresentate. Due tipi di rappresentazioni simboliche sono oggetto d’indagine (indagine empirica, mediante test psicologici, ma anche col ricorso a modelli teorici): le immagini *pittoriche* o *analogiche*, come la “fotografia interna” che ci facciamo di una casa o di una persona, e le rappresentazioni *linguistiche*, scritte o verbali. I nostri processi cognitivi operano attraverso queste due modalità di “codifica”, spesso interagenti: il sistema verbale e il sistema immaginativo (A. U. Paivio). I due modi contengono tipi diverse d’informazione sulla cosa (così, una figura pittorica ci informa molto sulle relazioni spaziali). Su questa tematica c’è stata una controversia negli anni ‘70 e ‘80 tra *proposizionalisti* (ispirati a Fodor), come Zenon Pylyshyn, per il quale il sistema proposizionale è il codice fondamentale dei processi cognitivi, e gli *analogisti*, come Stephen Kosslyn e sua collaboratrice Martha Farah. Per questi ultimi, le immagini godono di un’autonomia funzionale nella mente (le archiviamo nella memoria, e siamo in grado di “scannarle”, ruotarle, trasformarle, ingrandirle, ecc., come facciamo con i file d’immagini nei computer).

Queste discussioni, non prive di sottigliezze, dimostrano la difficoltà del cognitivismo di concepire la natura del pensiero. Al momento di arrivare al livello intellettuale, infatti, gli autori spesso si fermano al livello linguistico. Guidati dal paradigma computazionale, il *software* del nostro cervello (il quale è visto come *hardware*) “contiene” il nostro mondo di credenze, opinioni, obiettivi, dotato di “penetrabilità cognitiva”. Questo mondo appare costituito da un certo linguaggio. Ma quale?

La psicologia cognitiva ha oscillato tra diversi elementi rappresentazionali, ad esempio tra i tre rilevati da Johnson-Laird: le *immagini*, le *rappresentazioni proposizionali* e i *modelli mentali* (si pensi, ad esempio, al modo in cui ricordiamo una storia, con elementi verbali e immaginativi, con *frames* che ci consentono di capirla meglio e cose simili).

Concludo questa sezione. Abbiamo visto un quadro storico sommario di alcune tematiche, problemi e impostazioni della psicologia cognitiva. Il campo della ricerca, come si è visto in parte, comprende punti quali la percezione e il riconoscimento di oggetti, l’attenzione, la memoria, le rappresentazioni e l’organizzazione delle conoscenze (architettura mentale), il linguaggio il ragionamento, le soluzioni di problemi, i processi decisionali. Anche se i modelli cognitivi sono ispirati, in alcuni casi, ai processamenti computazionali, si ricordi che piuttosto è il contrario, poiché l’organizzazione concettuale dei programmi computazionali ha seguito le vie cognitive percorse

naturalmente dalla mente umana. Solo che lo sviluppo informativo rese un grande servizio alla psicologia, consentendole di superare la strettoie del comportamentismo. Comunque il panorama delineato è sufficiente per capire le posizioni della filosofia della mente che sono emerse in occasione di questi sviluppi scientifici.

## 7. Il funzionalismo computazionale (mente=computer)

Torno adesso alle tematiche proprie della filosofia della mente. Pagine addietro abbiamo visto il dualismo cartesiano, il comportamentismo filosofico e le teorie dell'identità. La caduta del comportamentismo psicologico e il passaggio alla psicologia cognitiva, vicina alla scienza computazionale, suscitò la nascita di una nuova forma di filosofia della mente: il funzionalismo. Il suo fondatore potrebbe essere considerato Hilary Putnam, il quale poi si è discostato da tale posizione. In verità il funzionalismo è una cornice molto ampia in cui possono entrare tesi abbastanza diverse. In termini generali, questa corrente si manifesta come non riduttivista, contraria al fisicalismo e al comportamentismo. La sua tesi centrale è la considerazione degli stati mentali in termini di rapporti funzionali alle entrate sensoriali, ad altri stati mentali e alle uscite comportamentali. Il funzionalismo *spiega* dunque il comportamento indagando nella "scatola nera" delle funzioni cognitive. Gli stati mentali sono qui considerati come *stati funzionali*, come qualcosa di diverso sia dalla base neurobiologica, contro la teoria dell'identità, che dal comportamento esterno (contro il comportamentismo). Così il funzionalismo garantisce autonomia alla psicologia, sia concettuale sia nomologica (vi sarebbero "leggi" proprie degli stati mentali).

Normalmente il funzionalismo non va al di là di questa posizione. Quindi in linea di massima resta compatibile col materialismo, anche se talvolta potrebbe essere reinterpretato in un senso non materialistico (in questo senso non è troppo lontano dall'antico "emergentismo", almeno in alcuni tratti che vedremo dopo). Di solito il funzionalismo è "ontologicamente neutro". Non arriva ad accettare l'esistenza ontologica degli atti della coscienza come qualcosa di non riducibile agli atti fisici esterni, per paura di cadere nel dualismo cartesiano. Negli anni della grande fioritura della letteratura cognitivista, anni '70 ed '80, il funzionalismo fu la filosofia della mente più di moda tra gli autori coinvolti negli studi cognitivi. Oggi questa corrente è un po' in declino, in quanto le si è rimproverato di aver preso la mente umana troppo sul modello dei processi informatici e di aver tralasciato il suo rapporto all'organismo e all'ambiente. Nei momenti iniziali, come vedremo adesso, il funzionalismo stette molto vicino alla scienza computazionale (vedendo i processi mentali in termini di funzioni computazionali).

A. *Funzionalismo computazionale di H. Putnam.* Nel 1960, Putnam accennò nell'articolo

*Mind and Machines* (cfr. *Mind, Language, and Reality: Philosophical Papers*, Cambridge Univ. Press, Cambridge 1985 vol. 2, cit., pp. 325-440) alla possibilità di un'equivalenza tra la mente umana e i programmi di computer. Il dolore, ad esempio, forse non sarebbe da vedere come un semplice stato fisicochimico del cervello o del sistema nervoso, ma piuttosto come uno *stato funzionale* dell'intero organismo, analogo agli "stati interni" di una macchina di Turing. Qui si presuppone che uno stato fisico reale sorga da un "meccanismo", e il meccanismo più universale finora proposto sarebbe la macchina di Turing. Ovviamente questo presupposto è gratuito: così si decide *a priori* che il dolore o qualsiasi altro stato fisico o psichico dovrebbe essere il risultato di una computazione, il che è tutt'altro che ovvio.

La mente umana, di conseguenza, secondo questa visione del Putnam degli anni '60 sarebbe *una macchina di Turing* (in altre parole, un tipo di computer), naturalmente una macchina particolarmente complessa, con la capacità di interagire col mondo esterno (e con altre macchine) e di operare secondo azioni pesate per il loro valore di probabilità. Ciò spiegherebbe l'incertezza di molti dei nostri stati e decisioni. L'uomo sarebbe, dunque, *un automa finito probabilistico* (come matematico, Putnam aveva lavorato nella teoria degli automi finiti).

La teoria putnamiana si può vedere anche da un altro punto di vista. Dal momento che l'organizzazione biologica è in grado di suscitare "stati funzionali", diversi dalla pura materialità (ad esempio, un dolore o una credenza), si potrebbe ugualmente supporre, secondo Putnam, che un'organizzazione elettronica -un computer, un robot- sarebbe pure in grado di arrivare a tali stati funzionali. Questa era appunto la teoria dell'IA forte, balenata dallo stesso Turing: la mente umana sarebbe simulabile *in toto* (e superabile) da un computer adeguatamente attrezzato.

Da qui nasce anche la tesi putnamiana della *multipla realizzabilità materiale* degli stati mentali in quanto stati funzionali. Questa idea negli anni '60 colpì fortemente la teoria dell'identità (cfr. *I robot: macchine o vita creata artificialmente?*, articolo del 1964; *Psychological Predicates*, articolo del 1967, poi rieditato come *The Nature of Mental States*). Se un medesimo stato funzionale può essere implementato in diverse strutture materiali (ad esempio, un desiderio in molti stati cerebrali), segue che tale stato non può essere definito in termini materiali, come volevano i materialisti.

In quanto la teoria dell'identità postulava invece che qualsiasi *tipo* di stato mentale era equivalente a un *tipo* di stato fisico (identità o co-estensività *del tipo* o *type*, il che comunque non significava stabilire tale identità per definizione, come prima avevano fatto i comportamentisti), è stato detto che la teoria di Putnam sosteneva al massimo l'identità *dell'occorrenza* o *token*, cioè al

limite uno stato psicologico sarebbe realizzato *in un caso concreto*, in una struttura materiale data, ma sarebbe sempre possibile che fosse implementato in un'altra struttura materiale. Quindi i tipi di stato funzionali non coincidono con la base materiale, anche se potrebbero coincidere nei casi concreti (nelle "occorrenze", traduzione del termine inglese *tokens*). La teoria dell'*identità dell'occorrenza* appare come un indebolimento della teoria dell'*identità del tipo* e quindi come una posizione che diminuisce la forza del fisicalismo della tesi dell'*identità tout court*. Essa apriva uno spazio all'autonomia della psicologia come scienza, ma non è detto che in fondo non sia ugualmente riduttiva, anche se in grado minore. Diciamo che tale posizione è *epistemologicamente antiriduttiva*, ma non lo è *ontologicamente*.

Negli anni '70 e '80 Putnam trovò poco soddisfacente la teoria funzionalista computazionale. Cfr. *The Many Faces of Realism*, 1987 (*La sfida del realismo*, Garzanti, Milano 1991); *Representation and Reality*, MIT Press, Cambridge (Mass.) 1988 (*Rappresentazione e realtà*, Garzanti, Milano 1988); *The Threefold Cord* (la triplice corda): *Mind, Body, and the World*, Columbia Univ. Press, New York 1998. Putnam divenne così un filosofo antiriduzionista. "Il funzionalismo non funziona, cioè non è adeguato ai fenomeni" (*Representation and Reality*, p. 105). I concetti di verità, intenzionalità, significato, ragione, non possono venir ridotti a processi spiegabili dalla fisica, dalla chimica o della biologia, non possono cioè essere "naturalizzati" (*Why Reason Can't Be Naturalized*, in *Realism and Reason. Philosophical Papers*, vol. III, Cambridge Univ. Press, Cambridge 1983, pp. 229-247). Tuttavia, Putnam non dispone di una teoria alternativa consistente sulla questione mente-corpo. Possiamo dire che egli ha lasciato aperto il problema.

B. *Altri autori*. Nonostante le autocritiche putnamiane, il funzionalismo computazionale, pur senza questo nome, sin dai tempi di Turing sta presente in quasi tutti i sostenitori dell'IA forte o vicini ad essa. Questi autori sostengono, con diverse sfumature, la somiglianza strutturale della mente umana, spesso vista vagamente come l'insieme delle funzioni del cervello, con l'organizzazione (*software*) di un computer. In questa linea si colloca ad esempio D. R. Hofstadter (cfr. *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic, New York 1980; in ital. *Gödel, Escher, Bach. Un'eterna ghirlanda brillante*, Adelphi, Milano 1985).

Le critiche all'IA forte non impedirono a certi autori di prendere occasione dalle scoperte informatiche per sferrare nuove obiezioni contro le nozioni psicologiche fondamentali, come persona, io, coscienza, individualità. Ovviamente, se la nostra intelligenza è un computer sofisticato, queste nozioni non hanno senso. Di solito si usa in un modo retorico -finanche troppola solita immagine dell'*homunculus* cartesiano -una specie di folletto, spiritello o "demone"- per

squalificare allegramente *ogni accenno alla coscienza*, come se essa fosse una sorte di reduplicazione superflua di funzioni che dovrebbero essere assegnate alla macchina del nostro cervello.

Il funzionalismo computazionale veramente non si discosta troppo dalle teorie dell'identità. Se il primo funzionalismo computazionale di Putnam sembrava piuttosto euristico, quello di altri autori più decisamente materialisti appare invece ideologico. Le difficoltà per il *riconoscimento dell'identità personale e della coscienza* sono tradizionali nell'empirismo anglosassone sin dai tempi di Hume. Dietro queste difficoltà non solo si vede l'impatto dell'eredità empirista e comportamentista, ma si osservano pure le conseguenze del rifiuto della metafisica. Infatti, anche se l'esistenza del proprio io e dei fenomeni della coscienza può sembrare più che ovvia per una visione fenomenologica, non risulta facile "accettare" la realtà dell'io autocosciente -una realtà inspiegabile in termini fisici e comportamentali-se non si ha a disposizione un'ontologia metafisica.

Un rappresentante di quest'indirizzo fortemente antimetafisico è *Marvin Minsky* (cfr. *The Society of Mind*, Simon and Schuster, New York 1986; it. *La società della mente*, Adelphi, Milano 1989), uno degli esperti mondiali nel campo dell'IA. Il pensiero umano perde in Minsky unità e radicamento in un soggetto, riducendosi all'equilibrio tra una pluralità di elementi chiamati *agenti*. Il titolo dell'opera menzionata ne indica la tesi principale: il pensiero è "ciò che fa il cervello", e può concepirsi come una "società burocratica di agenti", non come un'unità monogovernata. Questi elementi sarebbero il prodotto dell'evoluzione biologica e culturale e si trovano in un rapporto dinamico reciproco, trasformando costantemente gli stati mentali. L'io e la libertà sarebbero dei miti (l'aspetto vero di questa teoria è che noi non siamo un'identità personale pienamente integrata. Abbiamo elementi neuropsicologici dotati di una certa autonomia, e che non sempre sono in armonia reciproca).

Così, Minsky ritiene sbagliato pensare che "dentro di noi vive una Persona che pensa per noi; questa Persona noi la consideriamo il nostro 'Sé'. Come funziona la visione? Semplice: il nostro Sé guarda attraverso i nostri occhi, allo stesso modo che noi guardiamo da una finestra. Come facciamo a parlare? Anche qui, tutto sembra venire da dentro: è il nostro Sé che ha qualcosa da comunicare. Come prendiamo le nostre decisioni? Il nostro Sé usa il suo 'libero arbitrio' (...) Ogni paragrafo di questo libro si propone di rispondere a queste domande, senza ricorrere al concetto di Sé (...) *La mente è semplicemente ciò che fa il cervello*" (*La Società della mente*, cit. pp. 19-20). Con queste parole si fa una facile caricatura del Sé (io, anima), visto come una sorta di "piccolo fantasma nascosto nella macchina", come una *virtus dormitiva* che tutto spiega saltandosi i

particolari. Si fa uso di metafore ridicole, intese a rendere facile il bersaglio da colpire.

*Daniel Dennett*, anche se può essere considerato un funzionalista “teleologico”, come diremo più avanti, coincide con Minsky nella sua critica dissolvente dell’io, visto in questo caso sotto la metafora burlesca dell’occhio invisibile che guarda un “teatrino cartesiano” (cfr. D. C. Dennett, *Consciousness Explained*, Little Brown, Boston 1991; it., *Coscienza*, Rizzoli, Milano 1993). Gli elementi molteplici a costituire la “mente” in questo caso sono chiamati *memi*. Il nostro cervello sarebbe sorto come conseguenza dell’evoluzione biologica e culturale, la quale alla fine produrrebbe unità autonome e prive di finalità (come i “geni egoisti” di Dawkins), i “memi”, talvolta in situazione conflittuale tra loro, tramite la solita concorrenza darwiniana. I memi sarebbero diffusi dappertutto, non solo nella mente -che in quanto tale non esiste, non essendo altro che un insieme di memi- ma anche in tutta la cultura.

Per non cadere nel famigerato dualismo cartesiano, non si dovrebbe quindi cercare un “dittatore” interiore o “capo” che cerchi di orchestrare la molteplicità dei memi in concorrenza,. Ciò che “noi” crediamo di essere viene ridotto all’atomizzazione dei memi di un corpo con un cervello (anzi due, quello destro e quello sinistro). L’io o la coscienza non sarebbe altro che una costruzione astratta irreali, una finzione utile, nata dalla narrazione di una certa storia costruita, che comunque si può ricostruire svariamente con diversi pezzi prescelti e messi insieme. Talvolta in noi si potrebbe parlare, anzi, di diversi io, contemporanei o successivi. La nostra identità, insomma, è un’identità narrativa convenzionale, così come si può creare un personaggio da fumetto, capace di essere risuscitato in nuovi contesti.

Si leggano, in questo senso, queste incredibili parole: “Se, invece, tu pensi te stesso come un centro di gravità narrativa, la tua esistenza dipende dal perdurare della narrazione (più o meno come *Le mille e una notti*), che potrebbe *in teoria* sopravvivere a infiniti cambiamenti del *mezzo*, essere teletrasportato così facilmente (in linea di principio) come il telegiornale della sera e immagazzinato indefinitamente come pura informazione. Se ciò che sei è questa organizzazione dell’informazione che ha strutturato il sistema di controllo del tuo corpo (o, per dirlo nel modo più usuale e provocatorio, se ciò che sei è il programma che gira nel tuo calcolatore cerebrale), allora potresti in linea di principio sopravvivere alla morte del tuo corpo, così come un programma può conservarsi intatto anche dopo che sia stato distrutto il calcolatore sul quale è stato creato ed eseguito per la prima volta” (D. C. Dennett, *Coscienza*, cit., pp. 478-9).

Queste conclusioni non hanno affatto una base scientifica e sono di una superficialità disgustante. Lo stesso si dica delle ricostruzioni evuzionistiche che narrano la “storia” di come

potrebbe essere nata la nostra mente così ridotta: veramente siamo davanti a racconti chimerici, che vengono vanno ripetuti molto spesso e che passano per scienza soltanto perché sono mescolati con elementi scientifici reali, ma subito interpretati in un certo modo a seconda dei presupposti *a priori* imposti. Non illudiamoci: è così come si fa scienza in certe aree. Ma i racconti inventati passano e le verità restano. Purtroppo siamo in un campo minato da ideologie.

È strano: malgrado l'antiplatonismo di questi autori, la concezione informatica dell'uomo sembra portare alla fine al disprezzo del corpo, la cui materia risulta irrilevante. Il programma che "noi siamo" si potrebbe "incarnare" in materia organica, nella particelle ecc. e perciò si pensa facilmente alla "reincarnazione". Ma è naturale che un approccio puramente tecnologico -la mente come macchina- sopprima il valore della natura fisica. La "realizzazione corporale" "qui è il risultato quasi accidentale del predominio delle funzioni delle macchine. Antonio M. Battro, neurologo argentino e membro della Pontificia Accademia delle Scienze, osservava in una conferenza a Roma come molte concezioni attuali della mente sono curiosamente "disincarnate". "Si somos coherentes con esta línea de 'desencarnación', lo único que necesitamos del cerebro, en realidad, es la información que tiene en su corteza, y no faltó un físico [si riferisce a Tipler, con la sua opera *La fisica dell'immortalità*, Mondadori, Milano 1995] que postuló que la resurrección equivaldría a ser simulado por una computadora en el ciberespacio, lo que sería, en definitiva, el más allá" (testo privato della conferenza, novembre 2001, Ambasciata Argentina presso la Santa Sede).

C. *Le difficoltà del funzionalismo computazionale*. Il primo funzionalismo computazionale (Putnam), come si è detto, colpì la teoria fiscalista dell'identità (*mente=cervello*), in quanto concepiva la mente umana non come materia fisica, ma sul modello delle macchine informatiche (*mente=computer* o, meglio, *mente=software del computer*). Tuttavia la tesi era confusa e non poteva durare a lungo, anche perché non era troppo difficile stabilire l'equazione cervello=un tipo speciale di computer (anche se oggi questo punto è rifiutato dai neurologi).

D'altra parte, *le nozioni di "realizzabilità" e di "funzione" sono un po' vaghe*. Come ho detto in altre occasioni, *non si possono mescolare atti tanto diversi quali i dolori fisici, le credenze o le promesse*. Dire che una promessa si può "realizzare moltepliciamente" in diversi stati cerebrali sembra chiaro, perché l'atto di promettere di per sé ci appare immateriale (in quanto atto, non è destinato a perfezionare un corpo). Invece è strano sostenere che le sensazioni di dolore o di fame si possano realizzare in vari modi, con indipendenza dalla base materiale. Mangiare, avere fame, ecc., sono atti fisici. Altri animali potranno avere una diversa struttura anatomica e fisiologica, ma ciò non toglie che la funzione di alimentarsi è intrinsecamente legata alla materialità (a livello di "tipo"

e non solo di “occorrenza”). La “multi-realizzabilità” putnamiana dei programmi informatici su diverse basi materiali (un film può essere registrato in un vecchio disco di plastica, in un video, in un CD o in un DVD) non indica necessariamente l’emergenza di atti psichici o vitali, ma al limite è solo un indice della semplice distinzione aristotelica tra materia e forma. Molte forme o strutture fisiche (una casa, un libro) possono essere realizzate su supporti fisici di vario tipo.

Di conseguenza, *l’argomento di Putnam della funzionalità come modo di superare il materialismo riduttivo era in realtà molto debole*. Un materialista non è tenuto a sostenere che qualsiasi cosa è un semplice oggetto fisico. Egli può facilmente riconoscere aspetti delle cose, come le relazioni o le funzioni, che non sono di per sé “cose fisiche”, e non per questo sarà costretto ad abbandonare il materialismo.

La debolezza del funzionalismo computazionale si vede poi nel fatto che *oscura la distinzione tra uomo e macchina*, facendo pensare che il nostro provare dolori o sentire la fame non è tanto diverso di quello che potrebbe avvenire in un robot elettronico dovutamente attrezzato. D’altra parte, la nozione di *funzione* è molto generale, dal momento che essa potrebbe interpretarsi in un senso non lontano dal comportamentismo, perdendo la qualità propria dell’*atto*. Quale sarebbe, ad esempio, la “funzione” del dolore? Se è quella di segnalare un danno all’organismo, ovviamente un computer potrebbe anche segnalare di avere danni analoghi. Anche la coscienza viene spesso ridotta alla funzione esterna di “monitoraggio” computazionale, cosa che ovviamente può fare un computer. In questo senso, oggi si distingue (per es., in N. Block) tra la *coscienza fenomenica* (o di prima persona) e la *coscienza di accesso* o “cognitiva”, la cui funzione sarebbe il monitoraggio degli elementi del sistema cognitivo. Tutto il problema risiede, in fondo, nel non voler accettare l’originalità dei *qualia*, in quanto non si conosce un’ontologia capace di darne ragione. Nessuna forma di funzionalismo è stata in grado di superare questa lacuna.

Di qui lo strano destino del funzionalismo. D’una parte, questa linea filosofica originò posizioni “disincarnate” della mente, come abbiamo visto citando Battro, in quanto la mente veniva ridotta a una pura funzione “realizzabile in qualsiasi materia”. D’altra parte, il funzionalismo provocò la reazione contraria di vedere la mente come intrinsecamente legata a un tipo di organismo (“naturalismo”). L’orbita di queste due posizioni è sempre materialista, come lo sono le altre forme di funzionalismo che ci accingiamo a vedere. Come dicevo all’inizio di questo corso, nell’attuale filosofia della mente spesso assistiamo a dispute di famiglia tra materialisti.

A questo punto può essere utile introdurre il significato di due termini molto usati nella *Philosophy of Mind*:

a) *Naturalismo* è la posizione o almeno il tentativo di spiegare le funzioni cognitive in termini di fisica, chimica e biologia.

b) *Mentalismo* sarebbe la linea filosofica che sostiene l'esistenza di una mente come qualcosa di diverso dal corpo, tesi riconducibile a Cartesio secondo gli autori (quindi, anche Aristotele sarebbe dualista in questo senso).

È scorretto, a mio parere, parlare della “vecchia teoria della mente”, come se si trattasse di una “teoria scientifica” superata, così come sarebbe molto inadeguato parlare della “vecchia teoria della realtà”. Né l'esistenza della realtà o della mente umana (ne quella del presidente dell'Italia) sono “teorie”, tanto meno scientifiche. Sono realtà incontrovertibili, da sempre riconosciute a livello di conoscenza ordinaria. Gli autori contrari all'esistenza dello spirito o dell'anima, se sono intellettualmente onesti sono tenuti ad ammettere che siamo in un campo filosofico e non scientifico (“ideologia”, invece, è far passare come acquisto scientifico ciò che è in realtà una tesi o un'opinione filosofica). D'altra parte, non è esatto opporre il mentalismo al naturalismo, nel senso di presentare la scelta tra “Cartesio o il materialismo”. Esiste un naturalismo aristotelico, compatibile con l'esistenza di una mente immateriale, una mente comunque radicata nella natura umana e anche nel nostro corpo.

D. *Alcune obiezioni particolari contro il computazionalismo*. Prima di passare ad altre forme di funzionalismo, presenterò certe obiezioni molto note che sono state sollevate contro la sua versione computazionale. In realtà, tutte le obiezioni contro il fisicalismo e in difesa dei *qualia* sono altrettanto valide nei confronti del funzionalismo computazionale, come quelle che abbiamo visto di Nagel e Jackson.

\* Un'obiezione molto classica è quella dei *qualia* invertiti: che cosa succederebbe se un soggetto A vedesse le cose divise nei colori rosso e verde, mentre un altro soggetto B vedrebbe alla rovescia, cioè quello che per A è verde, per lui sarebbe rosso, e viceversa?.

\* Argomento dei *superspartani*, proposto da Putnam contro il comportamentismo, ma valido contro chiunque neghi i *qualia*: si potrebbe pensare a individui “superspartani” che, pur avendo forti dolori, sarebbero in grado di nascondere qualsiasi manifestazione nei loro confronti.

Questi due esempi di fantasia non sono spiegabili in una concezione che voglia ridurre i *qualia* alle sue funzioni, senza riconoscere le sensazioni stesse come atti dell'esperienza soggettiva. In realtà, tutti questi esempi (anche i due seguenti) sono omogenei, in quanto tendono ad evidenziare il carattere irriducibile della coscienza.

\* Obiezione del *popolo cinese*, proposto da Ned Block (cfr. *Troubles with Functionalism*, 1978): supponiamo di simulare il funzionamento della nostra mente individuale in tutto il popolo cinese. Ogni cinese potrebbe essere incaricato di una funzione dei processi informativi di un gigantesco computer, secondo le diverse regole funzionali (stati di input, output, elaborazione informativa, ecc.). “Funzionalmente”, questi collegamenti sono tutto quanto si richiede per avere un computer o un cervello-computer. Ma ciò non renderebbe tutto il popolo della Cina “cosciente”, come se fosse una vera mente.

\* La possibilità degli *zombies* (Robert Kirk): lo *zombie* sarebbe una specie di uomo spettrale, simile agli uomini nel suo comportamento ma senza *qualia*, una sorta di uomo “comportamentista” o “funzionalista” (poiché soddisfa la versione cognitiva del funzionalismo). L’esistenza di *zombies* è una possibilità logica (non comporta contraddizione), ma che senso avrebbero tali creature?

\* Argomento della *stanza cinese*, proposta da John Searle nel 1980 (cfr. *Minds, Brains and Science: the 1984 Reith Lectures*, Penguin, Londra 1989; *The Rediscovery of the Mind*, MIT Press, Cambridge (Mass.) 1992): si potrebbe immaginare un uomo chiuso in una stanza che, come se fosse un computer, rispondesse a domande fatte in cinese, senza sapere una parola di quella lingua, ma comportandosi come un sofisticato programma, ad esempio consultando manuali d’istruzioni per saper rispondere bene alle domande poste. Bene, anche se ciò sarebbe possibile in teoria, bisogna dire che quell’uomo non capisce niente: domina una pura sintassi, ma non una *semantica*. È quello che fa un computer.

Più tardi Searle ha aggiunto che neanche si dovrebbe riconoscere alla macchina il ruolo di fare un vero computo *sintattico*, visto che essa, essendo soltanto un aggeggio elettronico, non può in realtà eseguire una vera operazione computazionale e sintattica. La sintassi non appartiene alla fisica del sistema, ma è sempre in rapporto ad un osservatore intelligente. Neanche il cervello, in quanto organo fisico, può computare in tal modo: “la sintassi non è intrinseca alla fisica. L’assegnazione di proprietà sintattiche è sempre relativa a un agente o a un osservatore che tratta certi fenomeni fisici come sintattici (...) la sintassi è una nozione essenzialmente relativa all’osservatore. La molteplice realizzabilità di processi computazionalmente equivalenti in diversi mezzi fisici non è un segno che quei processi siano astratti, ma che non sono intrinseci al sistema come tale. Essi dipendono da un’interpretazione dall’esterno” (J. Searle, *The Rediscovery of the Mind*, cit., pp. 208-209).

L’argomento di Searle è stato respinto dai funzionalisti, ma senza un vero motivo. Certamente il *software* del computer contiene un linguaggio, quindi possiede una sintassi e pure una semantica,

dunque una certa intenzionalità (un computer dotato di IA può indicare la persona che si trova davanti ad esso, conversare di politica, ecc.). Ma tale linguaggio e intenzionalità esistono nel computer nello stesso senso in cui un libro contiene idee (il computer, per certi versi, è come un “libro interattivo”). I simboli sono tali per chi li ha creati e sa interpretarli. Altrimenti si riducono a una pura materialità non intenzionale. Il significato sta nella mente dell’interprete, non nella macchina simbolica.

Questo fatto vale anche per il cervello. Abbiamo accennato sopra alle difficoltà che ha trovato la metafora del cervello come computer. Nel cervello come puro organo fisiologico non c’è alcun linguaggio, tanto meno un vero programma scritto. Dire che nel cervello vi sono dei “programmi” non oltrepassa la metafora, così come si può dire metaforicamente che il codice genetico è un linguaggio in codice. Queste metafore sono semplicemente utili, e hanno come fondamento *in re* l’esistenza di un ordine intelligibile nella natura fisica. Possiamo pure dire che la natura è “come un libro” scritto a caratteri matematici. Ma la verità è che nella natura non ci sono formule matematiche né computazioni. Credere il contrario sarebbe antropomorfismo. Il fatto che comprendiamo la natura attraverso i numeri non significa che nel mondo fisico ci siano dei numeri, come pensavano gli antichi pitagorici. Nella natura esiste soltanto un ordine razionale che noi comprendiamo attraverso numeri, computazioni e simboli, che sono creazioni mentali.

Ovviamente il cervello è anche *un organo sensitivo* che può elaborare informazione senza ricorrere al simbolismo astratto, e solo in questo senso compie delle “computazioni”. Il simbolismo astratto è un’opera della nostra ragione in quanto capacità che trascende il cervello, pur basandosi sulla sua struttura neurale per la rappresentazioni dei simboli sensibili. Perciò, l’uomo imita artificialmente il cervello non quando crea dei computer simbolici, bensì macchine “connessionistiche”.

## **8. Il funzionalismo rappresentazionale**

Dopo la caduta del comportamentismo, gli psicologi cognitivi spesso optarono per il funzionalismo, posizione filosofica aperta a molte sfumature e riluttante al riduttivismo forte. Un esponente di questa linea è *Jerry Fodor*, funzionalista “rappresentazionale” anziché “computazionale”.

Fodor studiò con Chomsky e ricevette da lui l’impatto del nuovo approccio linguistico, con il suo superamento dell’anticomportamentismo. Nato a New York nel 1935, fu professore al MIT, all’Università della città di N. York e alla Rutgers University (New Jersey). È conosciuto per la sua

teoria del “mentalese” come linguaggio del pensiero e per la sua teoria psicologica “modularistica”. Sono importanti tra altre, le sue opere *The Language of Thought* (1975); *The Modularity of Mind* (1983) (it. *La mente modulare*, il Mulino, Bologna 1988; *Psychosemantics* (1987) (it. *Psicosemantica*, il Mulino, Bologna 1990); *The Elm and the Expert. Mentalese and Its Semantics* (1994); *The Mind Doesn't Work that Way: the Scope and Limits of Computational Psychology* (it. *La mente non funziona così*, Roma-Bari, Laterza 2001).

La mancanza di una relazione di identità “del tipo” tra eventi mentali e neurofisiologici porta a Fodor, come prima a Putnam, a sostenere il carattere irriducibile dell’architettura delle funzioni mentali interne, fornendo così una base favorevole all’autonomia della psicologia rispetto alle neuroscienze. Egli segue in questo punto la scia delle scienze cognitive e della linguistica di Chomsky, nonché l’analogia putnamiana della mente con il *software* del computer. La mente è l’organizzazione funzionale del cervello, capace di *spiegare causalmente* il comportamento umano.

A questo punto emerge l’elemento originale fodoriano: i processi “computazionali” della mente, proprio per spiegare la *produttività* e la *sistematicità* linguistica rilevate da Chomsky, comportano che la mente sia in grado di combinare “sintatticamente” degli elementi interni che vengono individuati come *rappresentazioni* mentali (di qui il nome di “funzionalismo rappresentativo”). Fodor concepisce le rappresentazioni e le loro combinazioni secondo regole “grammaticali” come un vero *linguaggio*, vale a dire, come un sistema di simboli le cui combinazioni con valore ricorsivo creano una reale sintassi mentale. Esiste, quindi, un *linguaggio del pensiero*, il *mentalese*, diverso dai linguaggi convenzionali. Il codice mentale o mentalese è, chomskianamente, innato. Non si riduce a immagini mentali, ma ha una struttura proposizionale (da notare che Putnam criticò la teoria fodoriana del linguaggio del pensiero).

Le rappresentazioni del linguaggio del pensiero sono *proposizioni*. Non si confondono con le proposizioni delle lingue artificiali (italiano, inglese, ecc.). Che cosa sono esattamente queste “proposizioni mentali” in Fodor è un punto alquanto oscuro (i suoi esempi sono sempre di proposizioni in una lingua artificiale). Riteniamo che Fodor punti al pensiero concettuale, ma purtroppo non dispone di una gnoseologia capace di distinguere con chiarezza tra pensiero e linguaggio. Il linguaggio formale (strutture logiche e grammaticali) è una creazione della ragione umana destinata alla “canalizzazione” delle nostre oggettivazioni intellettive (ma Chomsky credeva all’esistenza di una grammatica universale innata). Fodor si ferma proprio qui, al livello linguistico, anche se ciò le consente di parlare con libertà degli elementi psicologici della mente umana, superando così alcune restrizioni del materialismo. Fodor parla di *rappresentazioni*, sì, ma non

La mano  
tocca un  
ferro da stiro  
scottante (A)

Mi  
brucia: A  
ritiro la mano

Credo

Desidero  
guarire

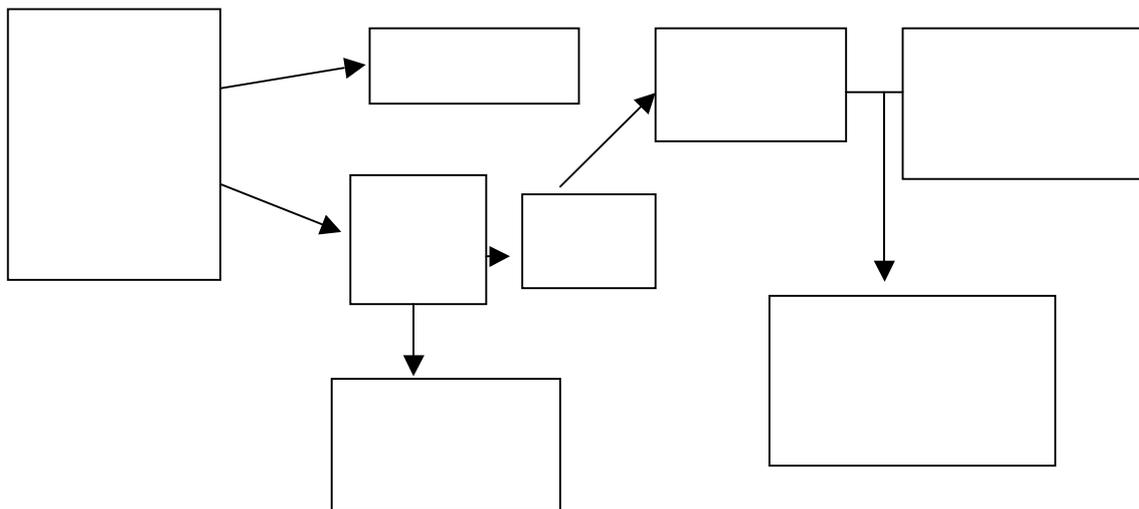
Condotta: Credo  
dirigo verso

mi  
scatola  
c'è una  
medicina

111

Elabora una vera gnoseologia di queste rappresentazioni. Egli è critico dell'IA forte, ma non riesce a liberarsi dall'analogia tra mente umana e software di un computer simbolico.

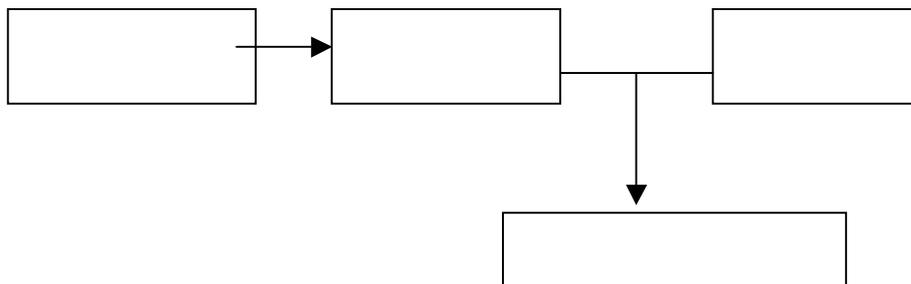
Schemi abituali della teoria funzionalistica rappresentazionale, molto affini al cognitivismo, sono come quello che segue (preso da Heil, *Philosophy of Mind*, cit., p. 100):



In quest'architettura cognitiva si vede come l'input sensoriale (tocco il ferro da stiro scottante) produce in me una sensazione di dolore (mi brucio), la quale, mentre provoca in me l'immediata reazione di ritirare la mano (comportamento esterno), mi conduce agli atteggiamenti proposizionali di ritenere che mi sono bruciato (credenza) e che la causa della mia scottatura è stata il ferro da stiro scottante che ho toccato (un'altra credenza). A sua volta, il sapere e sentire che mi sono bruciato provocano in me il desiderio di guarire (uno stato mentale, come il dolore, insieme a una credenza, conducono a un altro stato mentale: un desiderio). Tale desiderio, unito a un'altra credenza (so che c'è una medicina adeguata in una scatola), mi conducono all'ulteriore comportamento di muovermi verso la scatola delle medicine per prenderne una.

Ecco lo stile di analisi psicologica del funzionalismo fodoriano, molto usuale nelle scienze cognitive. Si veda sopra la nostra sezione sulla tematica delle rappresentazioni nel secondo cognitivismo. Abbiamo citato in questo senso la polemica di Zenon Pylyshyn, seguace di Fodor, contro i sostenitori dell'irriducibilità delle immagini a schemi proposizionali.

In un modo più generale, una spiegazione funzionalistica tipica sarebbe la seguente (presa stavolta da Heil, *ibid.*, p. 104):



Gli atteggiamenti proposizionali (credenze, desideri, speranze, atti di volontà, opinioni, ragionamenti) vengono così raccolti in “magazzini”. Si formano vincoli causali tra di essi, e tra gli input e gli output comportamentali. Forse non siamo molto lontani dalle spiegazioni cognitive della tradizione scolastica, dove si parlava di “facoltà” e di “abiti” per illustrare la struttura del mondo interiore e la sua causalità interna ed esterna. Ci pare pure adeguato, purché non sia riduttivo, prendere gli atti mentali non isolatamente ma nel loro intreccio causale con il resto dell’individuo. La visione olistica della mente è sempre più realistica. Del resto, il tipo di schemi cognitivi come quello di sopra è stato utile nelle scienze cognitive ed è vicino alla concezione computazionale della mente. C’è soltanto da obiettare l’eccessiva equiparazione che sembra stabilirsi tra pensiero e linguaggio, il che va incontro a difficoltà formidabili. D’altra parte, le obiezioni contro il funzionalismo, in favore di una risoluta affermazione degli stati coscienti, si mantengono in piedi.

Dal punto di vista filosofico, notiamo tre punti emersi molto chiaramente sia nelle scienze cognitive che nel funzionalismo: 1) *coscienza* in prima persona (il problema dei *qualia*); 2) *intenzionalità*, proprietà degli stati mentali di riferirsi (simbolicamente) a una realtà, che potrebbe essere pure inesistente, e di causare una condotta (“credo che *p*, e perciò agisco in questo modo”); 3) *razionalità*, o vincolo non fisico tra atteggiamenti proposizionali e tra forme della condotta umana, con valore esplicativo (se *p*, allora *q*, e allora agisco così). Questi tre punti, anche se erano conosciuti dalla tradizione filosofica, sono stati -e lo sono tuttora- gli elementi restii alla riduzione materialista o, come talvolta si dice, alla “naturalizzazione” della mente.

In difesa dell’autonomia della psicologia, Fodor ha seguito una concezione epistemologica “stratificata”, in cui una scienza di un livello più alto, astruendo dalle implementazioni di livello più basso, è capace di individuare leggi e proprietà valide di quello più alto, non riducibili alle leggi del livello inferiore. Quest’argomento ha originato numerose discussioni ed a che vedere con la questione della *riduzione* di una scienza ad un’altra (riduzione epistemologica con implicazioni

ontologiche). Come ho detto in altri momenti di questo corso, occorre distinguere con più chiarezza tra stati psichici sensitivi ed intellettivi, poiché il loro rapporto all'organismo (cervello) non è omogeneo. Il funzionalismo di Fodor, sulla base dell'idea che gli stati mentali possono essere implementati in qualsiasi tipo di sostrato materiale, occasionò un tipo di approccio psicologico troppo indifferente nei confronti della neurofisiologia. Gli stati psichici sensitivi, seppur non riducibili ai fenomeni fisici descritti dalla fisica, sono comunque fisici in un senso più elevato, e quindi non sono indipendenti dall'implementazione fisica (parlando aristotelicamente, si potrebbe dire che la forma non è indifferente alla sua realizzazione nella materia, anzi in un tipo concreto di materia, pur essendo diversa da essa).

Le teoria fodoriana dei *moduli*, infine, riguarda il modo in cui l'informazione è processata nella mente modellata sul paradigma computazionale. Così come i classici avevano parlato di facoltà psichiche, con la loro sede anatomica (si ricordino pure le funzioni cerebrali della frenologia di Gall), Fodor postulò l'esistenza nella mente di meccanismi computazionali specifici, specializzati, forse con una base neurologica, chiamati *moduli*. Alcune di queste funzioni sono state confermate nel campo neurologico, specialmente in rapporto a situazioni psicopatologiche con disfunzioni presumibilmente "modulari" (autismo, schizofrenia). I moduli, nella proposta di Fodor, sono indipendenti e hanno la proprietà dell'*incapsulamento informazionale*, cioè non accedono ad altre conoscenze all'infuori del loro *dominio specifico* e neanche si può ottenere l'informazione ad essi competenti se non grazie all'elaborazione informazionale adibita appunto al modulo specifico (si parla in questo senso di "impenetrabilità cognitiva" del modulo e della sua "obbligatorietà di uso"). Per la vista ci sarebbero, ad esempio, micromoduli deputati all'elaborazione dei colori, delle forme, delle relazioni spaziali. Nasce da qui la questione dell'individuazione di macromoduli e di micromoduli, cioè la scomposizione di una funzione in sotto-funzioni, fino ad arrivare ad alcune più elementari. Talvolta gli psicologi cognitivi impiegano la metafora dell'*homunculus* per riferirsi a tali funzioni, ma questo punto è aneddótico (l'*homunculus* sarebbe un piccolo "demone", pensato per prendere in giro il dualismo cartesiano, deputato ad una funzione psichica: *homunculus* che vede le forme, i contrasti, ecc.).

Il modularismo seguito da Fodor e da altri crea problemi non indifferenti di architettura cognitiva. Possiamo concepire la mente come composta di funzioni superiori (intelligenza) e inferiori (sensi), secondo una visione *piramidale* molto classica, oppure suddividerla in funzioni locali e autonome intrecciate, distribuite senza una precisa gerarchia, e soprattutto senza una funzione di "esecutivo centrale". Il "modularismo moderato" di Fodor combina entrambi gli aspetti. Secondo il suo indirizzo, il *modularismo* locale va assegnato soltanto a certe funzioni periferiche

(percezione, linguaggio, motricità), mentre si conservano le funzioni *olistiche* per processi superiori (credenze, desideri, ragionamenti, soluzione di problemi). In questo senso, viene rispettata una certa piramidalità e, soprattutto, si ammette l'esistenza di funzioni centrali, in grado di coordinare globalmente, almeno fino a un certo punto, l'insieme complessivo dei nostri atti mentali coscienti. Il problema è collegato a due punti: a) nel cervello non risulta esserci alcun'area deputata alla centralizzazione (non esiste una localizzazione dell'io); b) la modularità appare più legata a processi automatici, forse non coscienti, mentre le funzioni olistiche e centrali sembrano più in rapporto alla coscienza (che in noi è sempre coscienza di un io identico a se stesso). Certamente noi siamo unitari, ma anche divisi, e dobbiamo integrare la nostra personalità, nella misura del possibile. Un'illusione ottica continua ad essere tale, ad esempio, anche se sappiamo come sta la realtà dell'oggetto osservato: ecco un esempio di come una conoscenza globale non "accede" a un'informazione ricavata tramite un modulo locale; invece basta vedere una faccia per riconoscere una persona completa: abbiamo in questo caso la predominanza della globalità.

In una linea contraria a Fodor, altri psicologi cognitivi hanno sostenuto un modularismo "massivo", nel quale anche le funzioni ritenute superiori, come credenze, interpretazioni, modi di pensare, sarebbero dipendenti da "moduli". Si tentò di applicare questa tesi alla cosiddetta "teoria della mente", che tenta di studiare come siamo in grado di "leggere la mente altrui" (*mind-reading*), cioè di attribuire credenze, desideri, ecc. agli altri ("che cosa pensa lui di me? mente o dice la verità? sta egli attento a questo punto?, qual è la sua intenzione?"), anche in base a indizi come sono il rilevamento di stimoli a forma di occhio, interpretati come "vedere". Alcuni degli esperimenti fatti in questo senso non dimostrano necessariamente l'esistenza di un "modulo" biologico, ma indicano un certo abito mentale, forse acquisito in base a una certa educazione.

Un modularismo assoluto finirebbe per disintegrare la personalità in un insieme di meccanismi autonomi (come abbiamo visto per quanto riguarda Dennett e Minsky). Bisogna anche rendere ragione dell'identità personale e delle funzioni cognitive di controllo, attenzione e integrazione degli atti psichici, nell'unitarietà della propria coscienza. Probabilmente il funzionalismo rappresentativo, che comunque è anche computazionale, non è preparato per questo compito.

Il modularismo, infine, in quanto rappresentava una delle più tipiche espressioni dell'indirizzo cognitivista HIP (*human information processing*), si scontrò negli anni '80 e '90 con la concezione *connessionistica* dei processi mentali, che comporta un approccio più olistico e non basato sul simbolismo linguistico. Le polemiche non mancarono e sono ancora aperte. Personalmente

riteniamo che si possa parlare seriamente di modularità soltanto laddove ci sia una netta specializzazione organica destinata all'elaborazione cognitiva (i processamenti visuale o auditivo, ad esempio, sono chiaramente “modulari”).

I funzionalismi computazionale e rappresentazionale stanno nel cuore della scienza cognitiva. Le loro tematiche convergono, ad esempio con le indagini psicologiche sulle rappresentazioni, immagini e concetti. Non di rado le opinioni filosofiche e la trattazione psicologica si mescolano nelle pubblicazioni degli autori coinvolti in queste aree. La tesi dell'equiparazione tra cervello e macchina computazionale oggi è meno condivisa dagli autori. Il funzionalismo si è complicato in numerose altre versioni, come adesso vedremo, mentre la psicologia cognitiva continuava la sua storia. Le novità della scienza cognitiva negli ultimi anni sono le proposte connessionistiche e una nuova attenzione rivolta alla neuroscienza.

## 9. Altri tipi di funzionalismo

Altri autori hanno presentato versioni particolari del funzionalismo, con nuove sfumature, talvolta non prive di sottigliezza, nell'interpretazione degli stati mentali. Il punto comune a tutti è il rifiuto dei *qualia*, per evitare il “pericolo” del dualismo cartesiano, nonché la critica al comportamentismo, in quanto si vuole riconoscere l'esistenza di un ambito di “stati mentali” cognitivi. Spesso questi stati vengono presi “funzionalmente” nel loro *ruolo causale*. Il dolore, ad esempio, sarebbe uno stato psichico interno *causato* da un certo danno nei tessuti e *capace* di provocare una condotta. Gli stati psichici vengono visti nel loro rapporto causale, e se si arriva al punto di definirli in questo modo, allora siamo praticamente nella teoria dell'identità, come abbiamo detto a proposito di Armstrong e Lewis, talvolta chiamati “funzionalisti causali”.

Spesso la questione si pone in termini epistemologici, anziché ontologici. Così come una teoria adoperi concetti teorici (non osservabili) per poter spiegare i fenomeni osservabili, in modo analogo i concetti mentali sarebbero “teorici”, nei confronti delle scienze biologiche, più limitate al campo empirico. Quindi i concetti “mentalistici” per alcuni apparterebbero o alla cosiddetta “psicologia del senso comune” (ad esempio, nozioni di dolore, desiderio, ecc.), oppure ad una teoria scientifica psicologica e non puramente neurologica.

Il ruolo causale dei concetti mentali sarebbe, inoltre, *nomologico*, in quanto consente di stabilire un *nesso condizionale universale* tra uno stato interno e il suo ruolo causale (ad esempio, ci sarà “dolore” ogniqualvolta ci saranno queste o quelle manifestazioni o vincoli causali). Il modo di distinguere tra i diversi stati mentali sarebbe la specificazione di tali leggi (ad esempio, per

distinguere la sensazione di dolore da quella di bruciore o di prurito). D'altra parte, il funzionalismo concepisce i ruoli causali degli stati mentali come *contingenti* (quindi con “molteplice realizzabilità” o identità *token-token* (non *type-type*) tra stati mentali e fisici.

Abbiamo già accennato alla vaghezza della nozione di *realizzabilità multipla*. Alcune “funzioni psichiche” sono intrinsecamente collegate a un tipo di struttura materiale, o forse sono un'attività intrinsecamente materiale. Così, la fame è una sensazione fisica collegata all'alimentazione organica. Se per astrazione prendiamo la fame “funzionalisticamente”, ad esempio solo in rapporto al fatto di aver bisogno di energia per continuare a funzionare, allora potremmo parlare anche della necessità di “alimentare” una macchina. Vale a dire, una funzione materiale, presa in un certo grado di astrazione, può essere collegata ad un tipo di entità materiale presa in un corrispondente grado di astrazione. Ogni lavoro fisico ha bisogno di energia, ma tale bisogno negli organismi si realizza come alimentazione organica, e negli organismi dotati di sensazione è accompagnato dalla sensazione di fame o sete. Dire la verità, invece, non è di per sé un atto materiale (non è una funzione destinata a perfezionare un corpo), eppure contiene essenzialmente un lato materiale, che è l'espressione linguistica. Senza la comprensione della verità, che è del tutto immateriale, nessuno capirebbe il senso di tale espressione linguistica, che sarebbe ridotta a puri grafici o rumori. Purtroppo i funzionalisti ignorano queste sfumature, collegato all'uso analogico di concetti in base ai diversi modi in cui una medesima caratteristica può realizzarsi nei diversi modi di essere.

A. *Davidson*. Uno degli autori più noti del funzionalismo “causalista” è Donald Davidson (cfr. *Essays on Actions and Events*, 1980; it. *Azioni ed eventi*, il Mulino, Bologna 1992). La sua teoria filosofica, denominata da lui stesso “monismo anomalo”, è pensata soprattutto per gli *atteggiamenti proposizionali* o “*stati intenzionali*” (desiderare, sperare, credere, avere paura di), non per le sensazioni o atti di coscienza. Il termine *monismo* indica che si sostiene, d'una parte, che gli eventi mentali sono fisici, ma si dice *anomalo* nel senso che *non è possibile correlare i due tipi di eventi mediante leggi strette*. Così, non si rinuncia alla validità e autonomia delle spiegazioni psicologiche, solo che le leggi della psicologia non sono strette, ma sono piuttosto generalizzazioni approssimative tratte dal senso comune. Le leggi non strette sempre ammettono delle eccezioni se cambiano le circostanze, cioè sono leggi cui si aggiungono clausole -sempre problematiche- del tipo *ceteris paribus* (“mantenendo invariato il resto”).

Un esempio di questo tipo di leggi sarebbe: “se Luigi desidera rinfrescarsi, e se crede che aprendo la finestra si rinfrescherà, allora -*ceteris paribus*- egli aprirà la finestra”. Abbiamo qui un

*desiderio* e una *credenza* che causano un'azione fisica. Il *ceteris paribus* previene contro le possibilità -veramente infinite- che ci siano altri elementi in gioco, per cui io potrei comunque non aprire la finestra (per esempio, per non contrariare un mio compagno, perché posso trovare altre soluzioni per rinfrescarmi, o perché mi accorgo che se apro la finestra entreranno zanzare, ecc.).

Gli eventi “desidero rinfrescarmi”, “percepisco una finestra”, “penso che aprire la finestra mi rinfrescherà” trovano una corrispondenza, secondo questa teoria, con eventi cerebrali che causeranno i miei movimenti fisici destinati ad aprire la finestra. Allora, se prendiamo la serie degli eventi fisici corrispondenti a queste azioni, dovremo andare a leggi strette di natura *fisica* e biologica. Se invece prendiamo il loro risvolto mentale (vedere la finestra, credere che aprirla mi rinfrescherà, desiderare rinfrescarmi, ecc.), andiamo ad una spiegazione della condotta in termini razionali (finalità, normatività, deduzioni), non fisici. Ma i criteri d'intelligibilità delle azioni umani non sono assumibili dalle leggi scientifiche naturali, per cui la spiegazione psicologica non sarà riducibile alla spiegazione fisica. Eppure, ogni evento mentale troverà una corrispondenza nell'occorrenza (*token-token*), ma non secondo il tipo (*type-type*), con certi eventi cerebrali. Questa corrispondenza è chiamata *sopravvenienza* (*supervenience*), termine improntato da Moore. Però, come abbiamo detto, non vi saranno leggi strette per collegare gli eventi mentali agli eventi fisici (leggi psicofisiche). Il concetto di legge *stretta* appartiene esclusivamente all'ambito fisico, nel quale la causalità viene considerata chiusa.

Nella teoria davidsoniana si possono considerare due punti. Uno si riferisce all'*interpretazione della condotta degli altri*. Secondo Davidson, possiamo capire il comportamento linguistico delle persone se accogliamo i loro atteggiamenti proposizionali (desideri, credenze, intenzioni, ad esempio di parlare seriamente o di scherzare). Dobbiamo capire il significato delle loro frasi e interpretarle in base a ciò che gli altri credono, sostengono come vero e intendono dire. Si aggiunge in questo senso il cosiddetto “principio di carità”: in linea di massima, dobbiamo interpretare con benevolenza gli errori, espressioni contorte o modi strani di dire degli altri (pensando, ad esempio, che vogliono dire qualcosa che abbia un senso, che sia vero, ecc.). Così, se qualcuno ci dice “lo farò e non lo farò”, presupponiamo che non si contraddice e cerchiamo di vedere “in che senso” sta facendo tale affermazione (evitando di pensare che “è un pazzo”). Questi criteri ermeneutici dei rapporti intersoggettivi sono ragionevoli e fanno parte della “teoria della mente” (metacognizione) con cui possiamo “penetrare” nella mente dei nostri interlocutori.

Un secondo punto riguarda la *causalità degli eventi mentali*. Davidson sostiene che uno stato mentale può essere la *ragione* (causa) per cui noi facciamo qualcosa (“ho preso il giornale perché

volevo leggerlo”). Ciononostante, come abbiamo visto, non ci sono leggi che possano stabilire un collegamento rigido tra eventi mentali e neurali. Così Davidson si è avvicinato, almeno in un modo epistemologico, alla teoria aristotelica dell’azione (cfr. l’opera di S. Brock sull’azione umana in San Tommaso). Purtroppo, la mancanza di un’adeguata elaborazione ontologica di questi punti non gli consente di superare il materialismo.

*B. Kim.* Il “monismo anomalo” di Davidson è ambiguo, per cui ha provocato numerose critiche e nuove proposte, non è nemmeno facile seguire le sfumature delle posizioni davidsoniani nei suoi scritti lungo gli anni a proposito dei problemi indicati. L’argomento più discusso è quello della causalità. Menzionerò, ad esempio, la tesi al riguardo di Jaegwon Kim (cfr. *Philosophy of Mind*, cit., pp. 125-154).

Il problema è vedere in che senso un atto mentale può essere detto “causa” di eventi mentali o fisici. La linea materialista di Kim (anche di Davidson) comporta l’esclusione automatica di una possibile causalità tra atto psichico e atto fisico -vista come interazionismo cartesiano-, anche perché per questi autori lo statuto ontologico dell’atto psichico è funzionale e non ontologico (nel caso di Davidson, come abbiamo appena visto, si riconosce agli eventi mentali solo un’autonomia epistemologica ed ermeneutica, vista l’impossibilità di una riduzione fisicalistica in tal senso).

Il problema della “causalità mentale”, classico nella *Philosophy of Mind*, resta però ristretto ad una visione angusta della causalità, di tipo humeana, dove non si concepisce la nozione di causa come separata dall’idea di legge, almeno statistica. Di conseguenza è improponibile sin dall’inizio la nozione di *causa libera*. Sorprende che nella visione classica, invece, *il vero agente sia l’agente libero*, il che non ha niente a che vedere con leggi fisiche. Ciò concorda, tra altre cose, con la pratica ordinaria dei processi giudiziari, dove la causa di un crimine non è una legge fisica, bensì la cattiva volontà del criminale. La tesi materialistica, in quanto suggerisce la riduzione della causalità degli atti umani a eventi neurali, condurrebbe alla conclusione che il criminale ha agito in un certo modo perché è determinato dai suoi meccanismi cerebrali, il che dissolve la responsabilità propria degli atti personali. Ogni ulteriore appello alla responsabilità, alla moralità, alla scienza, ecc., se viene fatto da questi autori, non è che un esercizio di retorica.

La soluzione di Kim al problema sta nel suo concetto di *sopravvenienza*. Egli stabilisce tre principi per definire il suo materialismo, più esplicito o radicale di quello di Davidson:

1. *Principio di sopravvenienza*: “ciò che è mentale sopravviene a ciò che è fisico, nel senso che due cose qualsiasi (oggetti, eventi, organismi, persone, ecc.) che siano esattamente uguali in

tutte le loro proprietà fisiche, non possono differire nelle loro proprietà mentali. Vale a dire, l'indiscernibilità fisica comporta l'indiscernibilità mentale" (*Philosophy of Mind*, cit., p. 10).

2. *Principio anticartesiano*: "non possono esistere esseri puramente mentali (ad esempio, anime cartesiane). Cioè, niente può avere una proprietà mentale senza avere qualche proprietà fisica e senza essere, di conseguenza, una cosa fisica" (*ibid.*, p. 11).

3. *Principio di dipendenza mente-corpo*: "le proprietà mentali di una data cosa dipendono e sono determinate dalle sue proprietà fisiche. Cioè, il carattere psicologico di una cosa è totalmente determinato dal suo carattere fisico" (*ibid.*, p. 11).

Se è così, ovviamente ci si può domandare in che cosa differisca veramente una proprietà mentale da una proprietà fisica. La sopravvenienza implica almeno una distinzione, cui si aggiunge una corrispondenza biunivoca (nel mantenere ancora la distinzione fisico/mentale, siamo nel funzionalismo, non nella teoria dell'identità). Ma non è solo una semplice corrispondenza -come accade nel parallelismo fisico-, bensì una dipendenza totale dalla base fisica. Basta provocare un evento fisico (neurale) per produrre la sopravvenienza, e invece non si ammette il contrario, in quanto implicherebbe interazionismo. Probabilmente siamo vicini al epifenomenismo o all'emergentismo. Il "sopravvenire" corrisponde propriamente all'evento psichico. L'evento fisico in fondo è ritenuto più reale. Gli eventi mentali "sopravvengono" agli eventi fisici ("emergono", "appaiono"), non viceversa. Ma non si dice che cosa sono gli eventi psichici (apparenze? funzioni?). La sopravvenienza, così come è intesa dalle precedenti premesse, tende a dissolvere la realtà psichica.

C. *Valutazione della sopravvenienza*. La tesi esposta, a nostro parere, va *ridimensionata*. In modo rigoroso sarebbe valido *solo riguardo agli atti psichici organici elementari*, come il dolore fisico o le sensazioni di fame o sete. I tre principi menzionati si possono accettare solo in questo senso. Ovviamente non c'è dolore fisico senza un'eccitazione neurale, che causa la sensazione dolorosa (e non ha senso dire il contrario). Sarebbe assurdo pensare a un dolore fisico puro "di un'anima" (ma anche di un robot), senza una base nervosa.

Invece, *negli atti psichici sensitivi più complessi* (immaginazione, emozioni) tale corrispondenza non è del tutto chiara. Probabilmente è vero che ad ogni situazione psichica concreta corrisponde una base neurale specifica, ma il discorso non può finire in termini di pura "corrispondenza". Ciò che conta è il dinamismo. L'animale dominato dalla paura a causa della percezione di una cosa ritenuta pericolosa sta suscitando nel suo organismo, simultaneamente,

poiché si tratta di un unico atto, un'alterazione neurale. La causalità procede, in questo caso, dall'alto verso il basso, e non viceversa. Per provocare la paura in modo normale nell'animale, non dobbiamo agire fisicamente sul suo corpo (come nel caso in cui vogliamo causare in lui un dolore fisico), ma dobbiamo agire "intenzionalmente" sul suo corpo sensibilizzato, ad esempio, mostrandoli un animale che gli incute terrore.

Il discorso invece è, infine, totalmente diverso per quanto riguarda *le qualità intellettuali e morali delle persona*. Secondo il principio della sopravvenienza di Kim, bisogna dire che "se Socrate è una buona persona, allora chiunque sia esattamente come lui, in rapporto a tutte le sue proprietà descrittive non etiche (ad esempio, lo stesso carattere e tratti personali, le sue disposizioni di comportamento, ecc.), sarà altrettanto una buona persona" (Kim, *The Philosophy of Mind*, cit., p. 223). Questa frase è ambigua. Chiunque fosse uguale a Socrate solo esternamente, con esclusione della dimensione interna, come un robot, *non sarebbe una buona persona, perché neanche sarebbe una persona*. Invece, se una persona fosse veramente indiscernibile da Socrate, nei suoi atti esterni in quanto manifestazioni di atti interni, allora tale persona semplicemente sarebbe Socrate: il principio, in questo caso, sarebbe del tutto banale. Se il principio è interpretato in un senso completamente exteriorista, allora siamo piombati nel comportamentismo.

Consideriamo adesso il caso di un'opera artistica. Se due opere di arte sono fisicamente indiscernibili, saranno esteticamente indiscernibili (cfr. *ibid.*, p. 222). Una volta determinate le caratteristiche di un'opera d'arte, o modificate, determiniamo o alteriamo le sue caratteristiche estetiche. Il cambiamento fisico, quindi, "determina" le proprietà superiori (mentali, morali, estetiche, ecc.). Queste affermazioni confondono, come quelle di sopra, la causalità materiale con la causalità formale. È vero che esiste una esatta corrispondenza tra struttura fisica e qualità artistica di qualsiasi opera artistica o letteraria. Basta cambiare una lettera della *Divina Commedia*, per alterare il suo valore letterario. Tuttavia, questa considerazione è statica, e perde il dinamismo della realtà. Se io cambio le lettere della *Divina Commedia* non a caso, ma con criteri estetici ("causalità dall'alto"), allora inciderò sul valore estetico dell'opera, come ogni artista che modella la materia secondo le esigenze della forma. Ma se io cambio le lettere della *Divina Commedia* solo con criteri tipografici, ad esempio affinché quelle lettere si vedano meglio, allora io esercito *una causalità materiale, non formale*, che migliorerà una concreta edizione dell'opera del Dante dal punto di vista materiale, non estetico. Ugualmente, posso stare attento alla conservazione materiale di un libro, per preservarlo dalla sua distruzione. Ma se il libro è di matematica, per scriverlo meglio, anche fisicamente, devo sapere matematica e agire con criteri matematici. La nostra conclusione è che il "principio di sopravvenienza" è banale.

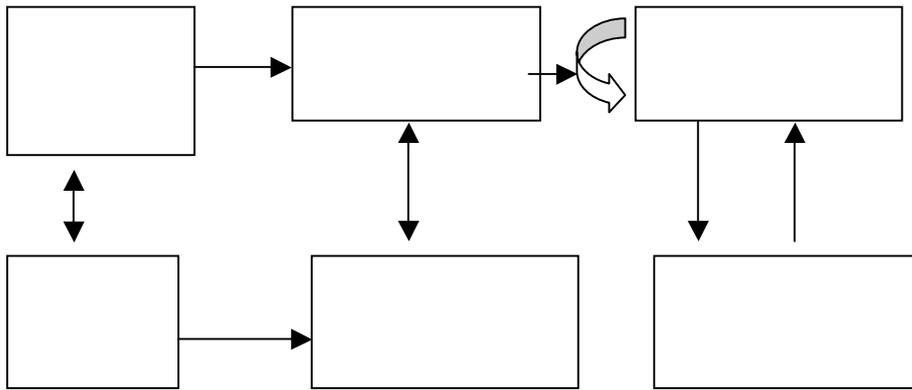
Nel mio articolo *La mente immaterial en Tomás de Aquino*, comunicazione presentata al Congresso Internazionale *L'umanesimo cristiano nel III millennio: la prospettiva di Tommaso d'Aquino*, 21-25 de septiembre de 2003) penso di essere stato troppo indulgente con questo principio, quando l'ho ritenuto accettabile solo per la dimensione sensitiva. Dopo ulteriori riflessioni, la mia opinione si è fatta più severa al riguardo. Ritengo fuorviante parlare di "sopravvenienza". Bisogna invece domandarci sul *grado di proporzione tra la dimensione formale e materiale degli atti*, diverso nel contesto dei gradi della vita. Ci sarà sempre qualche "corrispondeza" tra il lato psichico e il lato materiale degli atti sensitivi e intellettivi, ma ben diversa a seconda dei casi. La sopravvenienza di Kim semplifica la realtà in un modo univoco. Solo può convincere apparentemente nel caso elementare del dolore fisico (e atti simili).

I principi sulla corrispondenza tra la realtà psichica e fisica, seguendo l'ispirazione tomistica, sono i seguenti a nostro parere:

1. *Atti sensitivi*. L'operazione sensitiva, nella filosofia di Aristotele e Tommaso d'Aquino, è *un unico atto psicofisico* e non due atti correlati (cfr. *S. Th.*, I, q. 75, a. 3). Possiamo parlare qui di un'*identità elevata*, secondo cui *il psichico=il fisico elevato*. Non è che l'attivazione neurale "produca dolore", ma è dolore. Ma noi possiamo vedere tale passione dal punto di vista materiale o formale, così come una coppa di cristallo si può vedere nella sua forma o nella sua materia. *La distinzione tra linguaggio psichico, neurologico e comportamentale è adeguata secondo il nostro modo di conoscere*. "Ira" non si riduce funzionalmente a "rossore della faccia con certe smorfie", né all'attivazione di certi centri cerebrali. *L'ira comporta l'unità di questi aspetti, in un modo relativamente contingente*, dal momento che gli aspetti materiali di certe formalità superiori possono realizzarsi materialmente ed esprimersi con una certa flessibilità, in una maniera molteplice, ma entro certi limiti.

2. *Atti intellettivi*. I nostri atti spirituali (moralì, intellettivi, volontari: i pensieri di uno scienziato o di un mistico; le nostre opinioni, convinzioni e ragionamenti) hanno sempre una base organica (nervosa), *inadeguata ma necessaria, e anche essenziale fino a un certo punto*, dal momento che le nostre operazioni spirituali sono sempre basate su qualche esperienza sensibile e normalmente si esprimono in simboli sensibili. Ogni volta che formiamo un pensiero -almeno, "quasi" sempre- attiviamo il nostro linguaggio, ed esiste sempre qualche riferimento a un'esperienza sensibile (secondo la teoria tomistica della *conversio* alle immagini e alla sensibilità superiore (cfr. *S. Th.*, I, q. 84, a. 7). Di conseguenza, tutte le volte che pensiamo, qualche elemento del nostro cervello è attivato.

Lo spazio del problema si può capire con l'aiuto del seguente schema:



Ogni atto mentale trova in qualche maniera un corrispondente evento neurale. Concretamente, M<sub>1</sub> ed N<sub>1</sub>, così come M<sub>2</sub> ed N<sub>2</sub>, sono un identico atto sensitivo (psicofisico), con un lato psichico-formale e un lato fisico-materiale. N<sub>1</sub> determina M<sub>1</sub> l'ò viceversa? La domanda è fuori posto, poiché si tratta di un unico atto con una dimensione formale e un'altra materiale, causato dalla presenza dell'oggetto visibile di fronte alla capacità visiva, insieme alla capacità psiconeurale di riconoscere tale oggetto come "gelato". D'altra parte, analiticamente possiamo seguire la serie causale dal lato psicologico, e allora si dirà che M<sub>1</sub> causa psicologicamente o suscita M<sub>2</sub> (la visione suscita il desiderio), e anche che in qualche modo M<sub>2</sub> "causa" parzialmente M<sub>3</sub> (un desiderio attira l'adesione della volontà, che comunque resta libera di decidere). Veramente la volontà, assistita dalla ragione, decide di acconsentire al desiderio: non solo "mangio il gelato perché lo desidero" (come farebbe un animale), ma anche perché l'ho deciso (*ragione* della mia condotta, detta appunto "razionale"). Posso decidere freddamente, o sotto la pressione dei desideri, o con l'aiuto della moderazione proveniente dalla virtù (temperanza). Dal lato neurologico, possiamo anche dire che N<sub>1</sub> causa fisicamente N<sub>2</sub>, ma presupponendo che N<sub>1</sub> è un tutt'uno con M<sub>1</sub>, e che N<sub>2</sub> è un tutt'uno con M<sub>2</sub>.

L'atto volontario è libero, quindi non è propriamente *causato* dalla base neurale: dire il contrario sarebbe negare la libertà. Anzi una scelta provoca sicuramente nuovi atti psiconeurali. Ad esempio, quando mi decido a mangiare il gelato in un determinato momento futuro, continuo a vederlo e ad immaginarlo volontariamente, e inoltre suscito il comportamento esterno destinato a mangiare il gelato a una data ora. Ma si può riconoscere che l'atto volontario abbia pure un sostegno neurale nella sensibilità superiore: la scelta razionale richiede attenzione, coscienza, immaginazione, dominio linguistico. Una lesione cerebrale può diminuire la capacità di decidere (come lo fa l'ubriachezza, la droga, ecc.), così come l'intemperanza indebolisce l'energia del dominio volontario sulle emozioni e i desideri.

Da queste poche osservazioni si può capire quanto sia difficile la tematica della causalità psicofisica umana. Il nostro quadro, peraltro, è semplificato. Abbiamo saltato la causalità esercitata tra l'intelligenza e la volontà (molto studiata dai tomisti), e abbiamo parlato di semplici "corrispondenze" tra atti psichici ed eventi neurali, quando in realtà non si tratta di semplici eventi, ma di situazioni neurologiche molto complesse. E abbiamo ignorato l'influsso delle causalità esterne (ambiente, idee, pressioni di altre persone, ecc.).

Per concludere, possiamo considerare due punti relativi agli atti spirituali (intellettivi e volitivi):

i. *La corrispondenza degli atti spirituali (intellettivi e volontari) con la base neurale (sensibilità superiore) non è del tutto adeguata*, ma è pure necessaria e sicuramente contingente nei particolari. Il grado e modo di dipendenza tra questi atti appartiene al problema causale (vedere il seguente paragrafo). Esiste una corrispondenza (contingente, variabile, non univoca), ad esempio, tra ciò che Einstein pensa con le sue formule e determinate attivazioni correlative del suo cervello.

Veramente non sappiamo se ad ognuno dei nostri pensieri -ad esempio, il progetto di Michelangelo di modellare "La Pietà"-corrisponda un'attivazione neurale complessa e *specificata*, diversa da qualsiasi altra, anche se contingente (in quanto essa si potrebbe elaborare nel cervello in molteplici modi). Probabilmente no. Ma anche supponendo che fosse così, non ci sono motivi per il materialismo monistico. Tale attivazione neurale sarebbe al massimo l'espressione o *rivelazione* di un pensiero, il quale può essere "contenuto" in una forma materiale, pur trascendendola, così come un'idea può essere contenuta in una parola o in una frase pronunciata con senso, e anche in un gesto della faccia. Tale espressione sarebbe capita soltanto da chi fosse in grado di leggerla. Siamo di fronte ad un *corpo elevato, personalizzato, manifestativo dello spirito*. In termini antropologici, così come il linguaggio manifesta il pensiero, il corpo umano rivela la persona.

ii. *Il problema della causalità va affrontato ontologicamente*. Nel concetto di sopravvenienza di Kim, le proprietà fisiche determinano quelle mentali e non viceversa. Tutta la causalità ricade sulla parte materiale. Questa prospettiva nasce da una visione ristretta della causalità, presa univocamente dalle scienze naturali. La causalità mentale di tipo interazionista è respinta, ma in fondo viene pensata, come faceva Cartesio, sul modello della causalità (humeana, kantiana) delle scienze. Ma se è così, ci si colloca contro una convinzione incancellabile del senso comune, secondo la quale noi siamo padroni dei nostri atti, con libertà di determinazione. Si osservino le grandi opere dell'agire libero dell'uomo. È seriamente proponibile che le opere scientifiche e di arte, sempre variabili, sempre creative senza limiti, siano il risultato di leggi fisiche? Se qualcuno

così lo sostiene, bisognerà rispondergli che la sua opinione non è avallata da nessun argomento e da nessuna esperienza scientifica, ma resta soltanto un'idea, un postulato e in fondo una petizione di principio. Tale idea è fortemente controintuitiva, ed è ricavata semplicemente dal preconetto secondo cui le leggi fisiche dovrebbero per forza coprire tutta la realtà, anche il mondo umano della libertà.

*La causalità dei corpi è materiale e formale (“ilemorfismo”). Le loro alterazioni materiali possono indurre perfezioni formali (possono anche distruggere). Negli organismi sensitivi, la formalità superiore è un atto “attivo” insieme alla base materiale, in unità sostanziale. La causalità fisiologica è guidata dalla dimensione formale superiore inerente al corpo o alle loro parti specializzate (non in un modo “isolato”, bensì nella ricezione degli *input* ambientali o puramente fisici). Il cane assettato, beve dell'acqua perché ha sete, o perché qualcosa del suo sistema nervoso è attivato? Le due cose insieme, visto che la sete è “atto-di-una parte-del-corpo”, così come il sistema nervoso del cane è “materia-specializzata-di-una-forma”. L'organismo del cane percepisce l'acqua che soddisfa la sua sensazione di sete in quanto corpo animato. Alcuni filosofi legati al funzionalismo parlano, per casi simili, di *sovradeterminazione*, come se ci fosse un ricorso a troppe cause, a scapito di una visione causale “più semplice”. Ma questo termine vuole semplicemente mettere al riparo un'intenzione riduzionistica. La realtà invece è complessa e comporta una causalità complessa.*

Nell'uomo, inoltre, la forma superiore (“atteggiamenti proposizionali”, idee morali, religiose, ecc.) è ancora più attiva, potendo modificare -in un certo senso senza limiti- tutto l'ambito corporeo. Con la ragione, l'uomo è padrone del mondo materiale, del cosmo e del proprio corpo, come si evince dalle opere “infinite” della tecnica e dell'arte. *L'attività razionale (e volitiva) segue interiormente la linea della causalità formale che opera all'interno della materialità.* In questo senso, l'uomo muove naturalmente e con libertà molte parti del suo corpo:

- 1) Per compiere alcune funzioni biologiche naturali (mangiare, spostare il corpo).
- 2) Per comunicare con gli altri attraverso atti significativi e simbolici (sorridere, salutare, parlare).
- 3) Per modificare altri corpi (arte e tecnica), tramite i movimenti del proprio corpo e grazie alla conoscenza intellettuale della natura.

*Come può un atto di volontà muovere il corpo?* (“voglio muovere la mano e la muovo; voglio parlare e parlo”). In quanto la formalità superiore dell'uomo (anima) informa il corpo, l'intelligenza

volontaria è in grado di attivare naturalmente la sensibilità organica superiore inerente al cervello come organo proprio, purché tale sensibilità sia adeguatamente predisposta. Questa causalità dell'anima umana sul proprio corpo *non è fisica* (è metafisica, così come è metafisica la causalità di Dio sul mondo), ma è naturale. Non è accertabile dalle scienze naturali e non interferisce minimamente nelle leggi naturali. La dimensione spirituale dell'uomo apporta progetti, finalità, intenzionalità (voglio scrivere un libro e lo scrivo). In quanto tale dimensione formalizza e domina una parte dinamica del corpo, trova disponibili alcune delle sue forze e le utilizza. In quanto dipende dalla materialità, può trovare indisponibile il corpo (“voglio muovere la mano e non ci riesco”).

Le localizzazioni cerebrali sono indici specifici di queste disponibilità in rapporto a determinate funzioni (emotive, linguistiche, rappresentazionali, mnemoniche, attenzionali, di coscienza). Il cervello di Einstein, possiamo dire, sicuramente comportava per lui una maggiore disponibilità del suo corpo per fare scoperte intelligenti, ma non era la sua causa adeguata o completa.

Inoltre, dal momento che alcune opere della ragione si possono automatizzare (aspetti formali o computazionali: una deduzione, una traduzione), l'uomo è anche in grado di farsi aiutare da meccanismi esterni (*computer*) per arrivare a risultati razionali. In questo senso, si può dire che l'uomo “pensa con il suo cervello”, in quanto organo della sua sensibilità superiore, e inoltre che egli “pensa coi libri e coi computer”, in quanto strumenti esterni della ragione che consentono di agevolare e di ampliare la potenza di calcolare.

*D. Atti e funzioni.* Il concetto di funzione veramente è alquanto vago. Una risposta precisa alle problematiche suscitate dal funzionalismo richiederebbe l'uso di una serie di distinzioni ontologiche. In questo senso, è utile tener presente i seguenti punti, ispirati ad Aristotele e San Tommaso:

1. *La funzione primariamente è il compito delle parti di un organismo che serve per un determinato fine.* In questo senso, il concetto è pure usato analogicamente per qualsiasi tipo di *attività strumentale*, ad esempio per le operazioni degli *strumenti* (la funzione di una matita è di scrivere), per le *macchine* e le loro parti (la funzione della tastiera di un computer), o per gli atti degli organismi *istituzionali* (la funzione del Parlamento), oppure per gli atti di una persona in quanto servono a qualche fine (“la mia funzione in questa riunione è di stimolare il dialogo”).

2. *Gli atti o le operazioni, in senso stretto, sono ciò che fa un soggetto.* Ora, i “soggetti” in senso forte sono gli *individui sostanziali*, come un animale o una persona. Quindi le *azioni* in senso forte sono le attività attribuibili a soggetti discernibili, individuali e sostanziali, specialmente quanto tali attività dipendono da loro in una misura forte (altrimenti parliamo di *passioni*). Per questo motivo, parliamo di “attività” atomica, molecolare, ecc., ma non diciamo, ad esempio, che “una pietra agisce” (quando cade), mentre diciamo che “un cane agisce” quando perseguita un gatto.

La *prassi* è l’attività del vivente, soprattutto del vivente sensitivo. L’insieme delle attività dell’animale si chiama *comportamento*, particolarmente quando tali attività sono esterne, anche se implicano atti interni. Il soggetto che agisce normalmente è chiamato *agente*. Siccome l’azione si dice dell’agente in quanto *procede attivamente* da esso, l’azione per antonomasia è l’azione libera dell’uomo. L’uomo è l’essere “più attivo” dell’universo.

Nella filosofia di Aristotele e San Tommaso, le azioni degli animali e dell’uomo sono denominate *immanenti* quando sono perfezioni del soggetto “a se stanti, senza altri scopi” (così come un uomo può vedere un film “per vederlo, per contemplarlo”), anche se tali azioni possono pure avere altre utilità diverse dall’atto stesso (vedere un film può “servire” per stare con gli amici, o per riposarsi, ma lo scopo specifico di vedere un film è la contemplazione estetica “in se stessa”). L’atto immanente (detto anche *operazione*) è “fine a se stesso”, atto cercato per il suo valore intrinseco, e non “mezzo per un fine”, anche se può essere simultaneamente ordinato ad atti diversi o più alti (così, l’amore è un atto immanente, ma di per sé perfeziona anche colui che è amato). L’atto immanente può avere un *oggetto*, incluso nel contenuto stesso di tale atto: “conoscere” è un atto immanente che comporta la realtà conosciuta, così come “amare” è un atto immanente che comporta volere il bene di un altro soggetto. L’atto immanente non significa, dunque, che sia “chiuso”.

Gli atti che servono esclusivamente per il perfezionamento di un’altra cosa sono detti, nell’aristotelismo, *atti transitivi* (ad esempio, spingere, tagliare, riscaldare). Normalmente gli atti transitivi sono le *azioni materiali* di carattere inorganico, che servono per produrre effetti in altri corpi. Il concetto di *funzione*, esaminato sopra, si addice di più agli atti transitivi dei viventi, delle macchine e delle istituzioni, a causa della loro strumentalità. Non a caso i funzionalisti tendono a capire più facilmente tutto ciò che sia strumentale, mentre si dimostrano lontani dal comprendere la nozione di “atto immanente”.

Parliamo anche di *atti interni* ed *esterni*. Chiamiamo “atti esterni” quelli che sono *osservabili dai sensi esterni* e quindi stanno “alla disposizione di tutti gli osservatori”. Come abbiamo visto

pagine addietro, gli atti esterni spesso *vanno interpretati*, perché possono essere informati da una dimensione più alta, non propriamente osservabile con i sensi esterni. Così, il volto di un animale pericoloso è “esterno”, ma il riconoscimento che tale volto è “pericoloso” non dipende dalla sola vista, ma da una percezione in un senso più alto (neanche il riconoscimento che è “il volto di un animale” dipende esclusivamente dagli occhi, i quali di per sé colgono soltanto luci e colori). “Atti interni”, invece, sono quelli che *non risultano accessibili ai sensi esterni* (sensazioni, pensieri). *Gli atti transitivi di solito sono esterni*, anche se la loro funzionalità richiede un’interpretazione percettuale o intelligente. *Gli atti immanenti, invece, sono interni*. In questo senso, il comportamentismo e il neurologismo, nella linea dell’empirismo, tendono a ridurre ogni attività animale e umana agli atti transitivi esterni, con un inevitabile impoverimento dell’antropologia.

E. *Il funzionalismo teleologico*. Il funzionalismo è una posizione filosofica molto ampia e un po’ ambigua. Al suo interno si possono presentare molte tesi e sfumature sempre nuove, ma anche “vecchie”, nel senso che antiche soluzioni tornano a riapparire nel contesto delle nuove problematiche, ad esempio in rapporto al linguaggio, ai contenuti mentali o alla causalità degli atti mentali sulla condotta. Non è facile offrire una classificazione chiara delle posizioni denominate “funzionalistiche”.

Parlerò brevemente soltanto del funzionalismo “teleologico”. Come è suggerito dal nome, quest’indirizzo prende la categoria della *finalità*, vista in questo caso come una funzione biologica che può dare un senso ai “contenuti mentali” o all’apparente presenza di una “mente”. Autori vicini a questa posizione sono William Lycan, Elliott Sober, Ruth Millikan, Fred Dretske, D. Papineau. Personalmente riteniamo che concentrarsi sulla finalità, omettendo la questione della coscienza, anche se consente di ridimensionare la psicologia animale (il dolore, ad esempio, si può capire come qualcosa che punta a modificare la condotta in un certo senso), alla fine crea gli stessi problemi dell’approccio “causalistico” che abbiamo visto sopra. È vero che esiste una causalità negli atti “interni”, ma non basta concepirli solo in quella dimensione di “transitività” funzionale.

*Daniel Dennett* viene considerato talvolta un funzionalista teleologico (cfr. *L’atteggiamento intenzionale*, il Mulino, Bologna 1993; *Coscienza*, cit., 1993; *L’idea pericolosa di Darwin*, Bollati Boringhieri, Torino 1997; *La mente e le menti*, Sansoni, Milano 1997; *Brain Children*, MIT Press, Cambridge (Mass.) 1998). Secondo quest’autore, laddove vediamo lo svolgimento di un sistema razionale rivolto a fini, ad esempio un animale alla caccia di insetti, potremmo parlare di “atteggiamenti proposizionali”, con credenze e desideri. Possiamo dunque assegnare delle *menti* a piante, animali, uomini, robot, praticando nei loro confronti un “atteggiamento intenzionale”

(*intentional stance*). Ad esempio, se gioco a scacchi contro un computer, è bene - cioè comodo e vantaggioso- assumere nei suoi confronti una *intentional stance*, pensando che “mi vuole vincere”, “fa quella mossa perché pretende questo scopo”, ecc. Se vedo un gatto accanto al frigorifero, posso pensare che “vuole mangiare qualcosa”.

Di conseguenza, Dennett accoglie l'intenzionalità e il linguaggio psicologico come una forma di linguaggio utile e anche necessario nella pratica, pur nella consapevolezza della sua totale mancanza di riferimento ontologico. La teoria di Dennett, influito da Ryle e Quine, è una forma di *neocomportamentismo funzionalistico* che prende i concetti mentalisti secondo un'*epistemologia strumentalistica*. Lycan vede in Dennett una sorta di *omuncolarismo funzionale*, ispirato alle ricerche di IA. Le attività mentali vengono considerate qui in modo finto e strumentale, ma “utile”, come risultato del lavoro di un insieme di “omuncoli”, suddivisi a loro volta in piccoli omuncoli che svolgono compiti sempre più elementari e stupidi. Si ricordi la teoria di Dennett, sopra menzionata, dei molteplici memi in competizione darwiniana che costituiscono il cervello, senza alcun io centralizzatore. Secondo Sandro Nannini (*L'anima e il corpo*, cit., pp. 166-182), Dennett è tornato al comportamentismo di Ryle, solo che spostandolo dalla condotta alla struttura fisica, sia neurologica che meccanica.

*F. La questione del realismo delle rappresentazioni.* Nei dibattiti sulle tesi del funzionalismo filosofico sono emerse questioni gnoseologiche tradizionali, spesso relative alla questione del realismo cognitivo. In questo senso, alcune discussioni si sono riferite alla nozione di *significato concettuale*, un problema segnalato nella filosofia della mente col nome di “contenuto mentale” (*mental content*).

Putnam, in alcuni dei suoi lavori (*Il significato dei significati*, 1975), rilevò il lato “esterno” dei significati dei termini, evidenziando l'insufficienza della primitiva teoria di Fodor, che semplicemente si limitava a situare i significati nel linguaggio mentale. L'argomentazione ricorre a uno dei soliti *esperimenti mentali*: secondo un esempio putnamiano, chiamiamo *acqua* una certa sostanza con caratteristiche fenomeniche, conosciuta posteriormente come H<sub>2</sub>O. Eventuali terrestri gemelli, di una fantascientifica Terra gemella, potrebbero parlare anche di *acqua* rispetto a una simile sostanza fenomenica, che comunque non sarebbe H<sub>2</sub>O, ma avrebbe un'altra struttura chimica. Quindi non tutto dipende dagli stati mentali, che nell'esempio dei gemelli sono identici, ma anche dal *riferimento*, legato ad un contesto sociale e ad una serie di rapporti causali con la cosa cui si fa riferimento (*teoria causale del riferimento*, sostenuta anche da Saul Kripke). I significati dunque non stanno semplicemente “nella testa”, osservò Putnam. In un modo simile, Tyler Burge propose

l'esempio di una persona che usa il termine *artrite* secondo il suo vero significato (infiammazione delle articolazioni), mentre in un contesto contrafattuale una persona in una situazione simile impiega *artrite*, secondo un peculiare uso della sua comunità linguistica, come riferito anche a dolori reumatici.

Questi esempi si muovono in una linea un po' "antirappresentazionale" e spingono, in un certo senso, verso un maggiore realismo epistemologico. Difatti, Fodor presentò più tardi una teoria semantica dove il riferimento dei concetti è determinato in base al rapporto causale con gli oggetti riferiti. Il simbolo mentalese "cavallo" sarebbe attivato in presenza di cavalli reali, nelle "condizioni normali", poiché ovviamente possiamo sbagliare e confondere cavalli con asini.

Questi pochi cenni comportano problemi gnoseologici, non specifici della problematica "mente-corpo" (condizioni di verità e di riconoscimento del vero, pensieri su cose inesistenti, ecc.). "Il contenuto di un pensiero dipende dalle sue relazioni esterne, dal modo in cui esso è correlato al mondo, non dal modo in cui è correlato ad altri pensieri" (Fodor, *The Elm and the Expert. Mentalese and Its Semantic*, MIT Press, Cambridge (Mass.), 1994, p. 4). Si presenta, così, una tesi *realistica intenzionale*, come si vede nella seguente frase: "i pensieri sui cani vertono sui cani perché sono il tipo di pensieri che si può ritenere che siano causati dai cani" (*ibid.*, pp. 4-5). In questa linea, afferma pure Fodor: "le stesse proprietà delle rappresentazioni mentali che determinano i loro ruoli computazionali portano pure informazione sul mondo" (*ibid.*, p. 86). Quindi "la rappresentazione è un tipo di relazione causale col mondo" (*ibid.*, p. 87). Altri autori, invece, cercarono di delimitare il valore semantico dei concetti in base al *ruolo concettuale* giocato all'interno delle relazioni inferenziali che una rappresentazione intrattiene con altre in un sistema complessivo di pensiero (teoria del *ruolo concettuale* o del *ruolo funzionale*).

Rinasce così l'intera problematica gnoseologica del *realismo cognitivo* (cfr. il mio volume *Introduzione alla gnoseologia*, Le Monnier, Firenze 2003, cap. 3, sui significati concettuali). È naturale che questo avvenga qualora si voglia dar conto del significato delle rappresentazioni, un punto che la scienza cognitiva è costretta ad affrontare quando parla di reti semantiche, schemi, rappresentazioni linguistiche e immagini.

L'*esternalismo* noetico vede i contenuti mentali in rapporto ad elementi esterni di tipo causale, ambientale, sociale. L'*internalismo*, invece, fa dipendere i significati dai contenuti interni del pensiero (proprietà della mente o del cervello). Con un altro dei suoi esempi fantascientifici, Putnam l'eventualità in cui le nostre idee non fossero che l'effetto dell'azione di uno scienziato pazzo che avrebbe messo il nostro *cervello in una vasca*, collegandolo ad un computer, per produrvi

soltanto allucinazioni (un esempio molto simile a quello del genio maligno di Cartesio): questo sarebbe un esempio di “internalismo” estremo.

Tuttavia, non basta parlare di riferimento causale per sostenere il realismo noetico. L’istanza “esternalista” di tipo semantico potrebbe anche derivare verso un relativismo del “contesto sociale”. Il “realismo interno” di Putnam (contrario al “realismo metafisico”) sostiene, in questo senso, che i significati e la verità sono collegati alle molteplici *cornici concettuali* (o punti di vista pluralistici) degli uomini o dei sistemi scientifici. La cornice concettuale consentirebbe di accedere al mondo reale in un modo peculiare, sia pure valido.

Su questo problema potrebbe dare luce la distinzione tomista tra *modus essendi* e *modus cognoscendi*. Conosciamo la realtà trascendente tramite i nostri modi di rappresentarci le cose e di giudicare, con la possibilità di arrivare a giudizi veri o falsi. La verità del giudizio può essere relativa a contesti concettuali, semantici o pragmatici. Ma siamo capaci di avere un discernimento su tali contesti, per cui possiamo distinguere tra ciò che appartiene alla realtà o al nostro modo di conoscerla e di nominarla. Le nostre rappresentazioni concettuali spesso ci offrono un quadro imperfetto ma reale del mondo conosciuto. Tali quadri possono essere, talvolta, veri ma del tutto adeguati, o troppo vaghi e semplificati, per cui possono essere perfezionati, a seconda degli aspetti che intendiamo conoscere meglio. La conoscenza umana può essere vera o falsa. Ma insieme alla verità, nella conoscenza contano altri valori, quali la profondità, l’essenzialità, la completezza, la rilevanza, la fondatezza.

## 10. Altre posizioni nella filosofia della mente

A. *Eliminativismo*. Considerò in seguito altre posizioni della *Philosophy of Mind*, per completare il panorama storico su cui ci siamo soffermati in questo corso. Una posizione drastica e fortemente materialista nella filosofia della mente, che riprende la vecchia teoria dell’identità e la radicalizza epistemologicamente, è il cosiddetto eliminativismo. Come il termine suggerisce, la nuova prospettiva intende non fermarsi alla riduzione della sfera del mentale a quella fisica, ma piuttosto di *eliminare* l’ambito mentale, considerandolo come falso e fuorviante. La tesi ricorda l’antica posizione neopositivista che cercò di scardinare il senso della metafisica in favore della pura e semplice fisica. Infatti, l’eliminativismo è completamente scienziista. La scienza prescelta per eliminare la psicologia è, in questo caso, la neurofisiologia.

L’eliminativismo non è una novità. L’operazione di eliminare i concetti metafisici e psicologici è presente dappertutto nella filosofia contemporanea, ad esempio in Quine o Rorty, per

il quale la stessa questione mente-corpo nacque dall'impostazione cartesiana e quindi sarebbe uno pseudo-problema, che va *dissolto*, non risolto, nella linea anche del decostruzionismo di Jacques Derrida.

L'eliminativismo in senso tecnico del termine appartiene alla teoria preconizzata dai coniugi Churchland, Paul M. e Patrizia S. (cfr. P. M. Churchland, *La natura della mente e la struttura della scienza. Una prospettiva neurocomputazionale*, il Mulino, Bologna 1992; P. M. Churchland, *Il motore della ragione, la sede dell'anima*, Il Saggiatore, Milano 1998; P. S. Churchland, *Neurophilosophy. Toward a Unified Science of the Mind-Brain*, MIT Press, Cambridge (Mass.) 1986; P. S. Churchland e T. J. Sejnowski, *Il cervello computazionale*, Bologna, il Mulino 1996; R. N. McCauley (a cura di), *The Churchlands and their Critics*, Blackwell, Oxford 1996).

I Churchland si sono radicati a San Diego, California, vicino a un gruppo scientifico impegnato nelle ricerche connessionistiche (D. E. Rumelhart). Una caratteristica del loro indirizzo - Patrizia Churchland ha coniato il termine *neurofilosofia*- è la critica della cosiddetta *psicologia del senso comune* o popolare (*Folkpsychology*). Dal punto di vista delle polemiche di scuola, questa corrente è in contrapposizione al funzionalismo e al "mentalismo" di Fodor, più legato al MIT del Massachusetts. Un altro autore eliminativista è Stephen Stich, il quale poi ha moderato le sue posizioni, senza però uscire dal quadro pragmatistico e materialista (*Deconstructing the Mind*, Oxford University Press, Oxford 1996). Gli eliminativisti in realtà non sono molti e hanno ricevuto parecchie critiche.

Secondo i Churchland, bisognerebbe abbandonare in modo "kuhniano" il paradigma della psicologia del senso comune, su cui i filosofi computazionali (Fodor) cercano di costruire le loro teorie cognitive. L'unico paradigma valido è offerto dalla neurologia. La psicologia popolare (con i suoi concetti come credenza, desideri, coscienza, ecc.) sarebbe una teoria sbagliata, un programma di ricerca degenerativo, statico e inutile per la spiegazione, ad esempio, del sonno, delle malattie mentali, della creatività, del coordinamento senso-motorio, della differenza dei gradi di intelligenza delle persone. Cercare di mantenerla sarebbe come voler continuare ancora con vecchi concetti superati come flogisto, etere, o mantenere il geocentrismo tolemaico. Bisogna notare che la critica della psicologia del senso comune talvolta confonde il bersaglio, specialmente in Patrizia Churchland, con la versione funzionalista fodoriana, in particolare con la sua tesi del mentalese o modello "proposizionale" della mente.

La "riduzione" epistemologica in senso classico (Nagel) consisteva nella spiegazione dei concetti e principi di una teoria  $T_1$  in base ai principi di una teoria  $T_2$  riducente (in linea di massima,

tramite vincoli deduttivi). Ad esempio, la termodinamica fenomenologica TF fu ridotta alla meccanica statistica MS, cioè TD si può dedurre da MS, se aggiungiamo condizioni al contorno e alcune “leggi ponte”, e in questo senso MS “spiega” TF. In modo simile, i Churchland prospettarono una “coevoluzione riduzionista” della psicologia nei confronti della neuroscienza. Gli eventi psicologici dovrebbero a poco a poco essere correlati a eventi neurologici. Se una categoria psicologica concreta fosse restia a tale correlazione, andrebbe ravveduta, riconfigurata o eventualmente sostituita. La revisione forse condurrà all’eliminazione del concetto psicologico in questione. Così, dalla “coevoluzione riduttiva” si dovrebbe passare ad una “eliminazione” più decisa e sistematica. Ne risulterebbe una nuova psicologia ridimensionata, che potrebbe essere facilmente dedotta dalla neurologia.

Per avvalere il loro progetto, la coppia Churchland ha riposto molte speranze nel connessionismo, che si allontana dai modelli computazionali “enunciativi”, basati sul linguaggio. Le connessioni neurali costituiscono *pattern* o configurazioni. Seguendo gli attuali modelli fisici dinamici, potremmo rappresentare in uno spazio degli stati i *pattern* neurali come punti, anzi come un vettore a  $n$  dimensioni (dove  $n$  è il numero di neuroni o di gruppi neuronali). I rapporti causali tra i *pattern* sarebbero riconducibili a problemi di calcolo vettoriale, “mappando” (mettendo in corrispondenza) i punti dello spazio degli stati (ad esempio, correlando *pattern* sensoriali e motori, per spiegare il coordinamento sensomotorio).

B. *Osservazioni critiche.* Possiamo distribuire le nostre osservazioni critiche sull’eliminativismo nei seguenti punti (cfr. Josefa Toribio Mateas, *Eliminativismo y el futuro de la psicología popular*, in Fernando Broncano (ed.) *La mente humana*, ed. Trotta, Madrid 1995, pp. 245-272; Garth Fletcher, *The Scientific Credibility of Folk Psychology*, Lawrence Erlbaum Ass., Mahwah (N. Jersey) 1995).

i. *Portata della spiegazione neurologica.* A causa della vicinanza della psicologia alle neuroscienze, molte osservazioni dei Churchland sono giuste, in quanto esplorano la tematica classica della base neurologica dei fenomeni psichici. Può essere utile, senz’altro l’impiego nelle neuroscienze della fisica dei sistemi dinamici, con la rappresentazione  $n$ -dimensionale di spazi corrispondenti alle diverse modalità sensoriali (spazio dei colori, dei sapori, ecc.), e di spazi delle fasi per gli input sensoriali e motori. L’insufficienza dell’approccio neurologico per capire il comportamento umano è un fatto comunque scontato. Le nostre osservazioni critiche sono quelle che abbiamo avanzato nei confronti della teoria dell’identità. Conosciamo la base neurale dei nostri fenomeni psichici sensitivi, anche superiori. Così, la memoria a lungo termine in parte è basata,

molto probabilmente, su una modificazione permanente della permeabilità chimica delle sinapsi, e la memoria a breve termine può essere correlata alle reti neurali ricorrenti, così come la coscienza sensitiva appare collegata al coordinamento di varie attività corticali tramite il nucleo intralaminare del talamo. Questi e tanti altri esempi non sono una prova dell'identità tra gli aspetti fisici e psichici. Soltanto dimostrano che i nostri atti psichici hanno una base neurologica, che è solo parzialmente esplicativa.

ii. *Vincoli tra eliminativismo e connessionismo*. Il ricorso al connessionismo non dimostra che l'eliminativismo sia vero, così come il ricorso ad una scienza empirica non dimostra che tutto si debba ridurre a tale scienza. Il nuovo approccio connessionista può essere interpretato anche con una visione antropologica e cognitiva più ampia, non necessariamente riduttiva.

iii. *Eliminativismo e modelli proposizionali*. Nelle ultime fasi della letteratura eliminativista si osserva uno spostamento dalla critica della psicologia popolare verso la critica della psicologia fodoriana del mentalese. L'eliminativismo segnala certamente aspetti deboli del funzionalismo proposizionale, il quale sottovaluta l'importanza dei livelli subsimbolici e neurologici della conoscenza. Non è necessario, è ovvio, che tutte le nostre rappresentazioni siano proposizionali (non lo sono le nostre conoscenze procedurali). Abbiamo anche una dimensione animale delle nostre capacità cognitive, una dimensione che non usa il linguaggio, e neanche tutta la nostra vita intelligente si esaurisce negli usi linguistici.

iv. *Critica della psicologia popolare*. Gli eliminativisti prendono la conoscenza comune prescientifica come se fosse un blocco unico e indifferenziato. Innanzitutto bisognerebbe distinguere un livello fondamentale dei primi principi universali (esistenza del mondo, dell'io, delle altre persone, principio di non-contraddizione, valore dei ragionamenti, esistenza della verità e della menzogna), ben diverso da opinioni particolari o da credenze popolari, ad esempio sulla natura degli astri o degli elementi fisici o sull'esistenza della telepatia. La "psicologia popolare" non è una teoria, bensì un insieme di conoscenze condivise, alcune delle quali sono fondamentali e altre sono più particolari e contingenti e possono essere false, vaghe o ambigue (ad esempio, sappiamo di sognare, ma la conoscenza comune niente ci dice sulla natura e le cause dei sogni). È vero che con il sapere psicologico comune non si spiegano tante cose (come le cause delle malattie nervose), ma molti aspetti della vita umana sono ben spiegati con le luci ordinarie della ragione. Così, l'enunciato "Olga entrò nella farmacia perché voleva comprare medicinali per sua sorella, che è malata" può essere vero e spiegare veramente un comportamento razionale.

L'intreccio tra la psicologia scientifica e la conoscenza psicologica ordinaria non è di per sé conflittuale, ma di approfondimento, allargamento e complementarità. Quest'ultima è una piattaforma solida, malgrado la sua portata limitata in tanti aspetti. La scienza non si sostituisce semplicemente alla conoscenza ordinaria, ma l'approfondisce, spesso in una linea di convergenza e conferma, anche se talvolta la corregge. L'eliminativismo, in realtà, prende in modo troppo unilaterale la prospettiva neurologica, e così non soltanto squalifica la conoscenza ordinaria, ma anche gli apporti della psicologia scientifica.

C. *Emergentismo*. La tesi emergentista fu proposta da alcuni autori di lingua inglese della prima metà del XX secolo, quali Samuel Alexander, A. O. Lovejoy, C. D. Broad e C. Lloyd Morgan, allo scopo di rendere ragione delle proprietà *nuove* che compaiono nelle entità fisiche complesse, particolarmente negli organismi, rispetto alle loro parti e componenti più semplici. I viventi, ad esempio, hanno nuove proprietà rispetto ai non viventi, eppure la loro costituzione fisica è condivisa con gli essere inanimati, cioè non dispongono di un materiale diverso nei confronti degli esseri inanimati. La più complessa organizzazione di tale materiale *fa emergere* nuove proprietà. L'emergenza si può comprendere secondo gradi, in corrispondenza con i livelli di complessità degli esseri dell'universo, e allora dà luogo a una visione stratificata della realtà fisica. Si può anche collegare all'evoluzione, per cui si è parlato di *evoluzione emergente* (Lloyd Morgan).

L'emergentismo tradizionalmente si è presentato come una tesi antiriduzionistica e antimeccanicistica. Privilegia l'autonomia ontologica ed epistemica della totalità sulle parti, delle funzioni complessive sui componenti, quindi è contrario a vedere il tutto come costituito dalla somma delle sue parti. Lo si potrebbe ritenere, in questo senso, vicino alla teoria aristotelica della forma, che nell'ilemorfismo emerge o viene "edotta" dalla materia come qualcosa di diverso, così come la formalità della statua non può ridursi al marmo di cui è fatta. La nozione di proprietà emergente, comunque, può essere filosoficamente imprecisa. È semplicemente collegata alla totalità di un sistema (come un'automobile ha la proprietà di avere porte, il che non si comprende dalla prospettiva delle piccole parti), o è qualcosa di più? Candidati all'emergentismo sono, nell'universo, la vita rispetto alla non-vita e la mente rispetto agli esseri senza mente.

L'emergentismo non è stato molto di moda nella *Philosophy of Mind*. Spesso viene respinto proprio per il suo antiriduttivismo, ad esempio quando si pretende che le *proprietà emergenti* siano fatti reali nuovi e inspiegabili e non prevedibili dal basso, quindi in un modo "poco scientifico". Ma le nozioni di "totalità" o di "proprietà olistica emergente" non sono inspiegabili e imprevedibili, come se fossero "misteriose". L'Italia, ad esempio, è una "nazione" con una serie di caratteristiche

intelligibili, non qualcosa di inspiegabile. È chiaro che, analizzando solo gli individui, non arriveremo mai a “prevedere” o meglio a capire la realtà *nazione*, così come l’unità della *Divina Commedia* non si capisce senza lo sforzo per cogliere l’unità dell’opera, al di sopra dei versetti, delle parole e delle sillabe. L’emergenza certamente non “spiega” così come spiegano le leggi fisiche. Piuttosto spiega nel senso di far capire una realtà complessa a partire da un principio di unità in rapporto a una molteplicità di elementi. L’emergentismo va più in profondità rispetto al funzionalismo. Quest’ultimo vede il candidato all’emergenza solo nelle relazioni funzionali (strumentali, causali). L’emergenza indica, invece, un’autentica nuova qualità. Gli emergentisti considerano che tale qualità comunque è fisica, solo che appartiene ad un livello superiore e non può essere ridotta al tipo di qualità fisiche dei livelli inferiori.

*i. Emergentismo olistico.* Un tipo di emergentismo fa leva sulla *novità delle caratteristiche olistiche di un sistema. Riguardo ai processi cognitivi ed emotivi, tuttavia, non basta parlare di una semplice “totalità superiore”*. Evidentemente ci sono molti sensi di “emergenza”: non sono uguali una struttura collettiva ordinata, l’unità essenziale di una sostanza, un’operazione o un soggetto. Questo punto ci indica *la difficoltà di accogliere una nozione di “emergenza di una proprietà” all’infuori di una sistemazione metafisica della realtà*. Alcuni autori si trovano a disagio con la nozione di emergenza perché utilizzano soltanto il modo scientifico-naturale di capire le cose, nel quale non ha senso parlare di “cose che agiscono” (così si comprende più facilmente la nozione di *proprietà*, meno quella di *azione di un soggetto*). Ora, nel caso della vita psichica (conoscenza e emotività), una vera emergenza non solo significa l’apparizione di una nuova qualità, come accade negli organismi, ma anche la comparsa di un *soggetto* di attribuzione di operazioni e passioni immanenti. Queste operazioni e passioni sono *interne* (immaginazione, desiderio, dolore) e non solo semplicemente il risultato globale del coordinamento di una serie di azioni esterne (chimiche, elettriche, ecc.), un risultato che rimane sempre sul piano degli atti esterni. Proprio per questo nell’animale che sente appare per la prima volta il concetto di soggetto in senso forte (soggettività). Il modello degli atti olistici serve per capire certe caratteristiche macroscopiche di insiemi di microsostanze, ma non è valido per la comprensione degli *atti soggettivi*, che il comportamentismo e il neurologismo si ostinano a “non capire” solo perché si vuole capire soltanto ciò che appartiene alle categorie descrittive della fisica e della chimica.

L’emergentismo olistico non può superare il materialismo. Infatti, *Mario Bunge* definisce *materialismo emergentista* la sua soluzione “olistica” del problema mente-corpo (cfr. *The Mind-Body Problem. A Psychobiological Approach*, Pergamon Press, Oxford 1980; in spagnolo, *El problema mente-cerebro. Un enfoque psicobiológico*, Tecnos, Madrid 1985). La sua posizione,

benché antifunzionalista, non è così lontana dalla teoria dell'identità e infatti egli la vede come una varietà di monismo, contrapposta alle teorie dualiste. L'attività "mentale" è l'attività del cervello o del sistema nervoso. L'emergenza non è altro che l'attività nervosa considerata nelle sue proprietà *sistemiche*, cioè olistiche, non corrispondenti alle cellule prese in modo isolato. Di conseguenza, i rapporti "psicofisici" sono interazioni tra sottosistemi del cervello o tra il cervello e i sistemi muscolari, endocrini, digestivi, immunologici, ecc. La "mente" è l'insieme degli eventi mentali, e questi ultimi sono, come abbiamo detto, proprietà fisiche e organiche, non qualcosa di diverso.

A causa del suo emergentismo, Bunge non intende però negare la validità della psicologia né delle categorie psicologiche, tranne quelle che non abbiano una radice neurologica. La riduzione della psicologia alla neurologia è parziale. La psicologia impiega concetti che non compaiono nella neuroscienza.

La distanza tra Bunge e gli altri materialisti che abbiamo considerato è soltanto di sfumatura. L'eliminativismo ritiene inutili le categorie della psicologia comune. Bunge invece le accetta, ma ricordandoci che indicano atti organici. Ammette il senso normale di frasi come "Giovanni ama Maria", segnalando comunque che si tratta di un'azione del cervello.

*Emergentismo aperto.* Altre forme di emergentismo sembrano meno riduttive, senza però arrivare alla vera dimensione spirituale. Potrebbe essere la posizione di *John Searle* (cfr. *La riscoperta della mente*, Bollati Boringhieri, Torino 1994; *Il mistero della coscienza*, Cortina, Milano 1998; *Mente, linguaggio, società. La filosofia del mondo reale*, Cortina, Milano 2000), considerato da molti come un emergentista, anche se egli chiama la sua posizione filosofica "naturalismo biologico" e non si ritiene un materialista.

Il punto che più preme a Searle è la polemica contro il riduzionismo computazionale della mente, ed è celebre in questo senso il suo argomento della stanza cinese, che abbiamo già menzionato. Searle sostiene che i nostri pensieri sono *intenzionali*, non in un senso husserliano o brentaniano ma realistico: con i nostri pensieri ci riferiamo agli oggetti reali del mondo. Sostiene la realtà della coscienza e degli atti soggettivi. Intenzionalità e coscienza, quindi, sono irriducibili, non funzionali, e nascono o "emergono" dai poteri causali del cervello (quindi sono delle proprietà biologiche che emergono in determinati sistemi fisici). Il modo in cui si produce tale emergenza non è stato precisato da Searle. Questo punto rende la sua posizione ambigua, anche se è indubbiamente antiriduzionistica.

In una prospettiva tomistica, ciò che viene indicato come “proprietà emergente” nella posizione esaminata si comprende in un modo più ontologico alla luce di una metafisica dei gradi dell’essere. “Emerge”, “nasce”, “spunta” indica ovviamente l’apparizione di qualcosa di nuovo. Le cose del mondo hanno in comune la base materiale, ma non sono formalmente identiche. Alcune cose, sulla base di una migliore organizzazione della “materia seconda”, presentano strutture essenziali veramente nuove, così come l’uomo emerge come qualcosa di radicalmente nuovo rispetto alla vita infraumana. Nel tomismo, l’emergenza si comprende alla luce dei rapporti tra atto e potenza, cioè tra modi di essere, anche stratificati: l’atto emerge sulla potenza, la forma emerge dalla materia e sulla materia. L’emergenza di forme superiori di vita, nel caso di Popper, che adesso vedremo, s’inquadra in un dualismo forse non troppo lontano dall’aristotelismo.

E. *Dualismo interazionista (Popper, Eccles)*. In queste lezioni ci siamo riferiti soltanto al dualismo cartesiano. Nella filosofia della mente, come si è visto, il dualismo è la posizione sempre avversata, non solo a causa dell’*a priori* materialista, ma anche per la forza dello scientismo, con la conseguente ignoranza della dimensione metafisica. Questo non significa suffragare il dualismo. Ma bisogna anche riconoscerli i suoi diritti e non semplicemente cacciarlo via come “antiscientifico”. Un certo dualismo esiste sempre, altrimenti non si parlerebbe di mente-corpo, mentale e fisico. Questo certo “dualismo” o “mentalismo”, come dicono alcuni, impregna il nostro linguaggio e la nostra conoscenza psicologica ordinaria. Per questo motivo l’eliminativismo trova sconcertante il linguaggio psicologico popolare, che comunque i funzionalisti accettano e interpretano in modo “funzionale”, forse non del tutto chiuso ad una possibile interpretazione metafisica. Nella filosofia moderna i dualisti in senso classico sono pochi. Uno di questi pochi è Karl Popper.

La filosofia dell’uomo di Popper (cfr. K. Popper, J. C. Eccles, *L’io e il suo cervello*, Armando, Roma 1981, 3 voll.; *La conoscenza e il problema mente-corpo*, il Mulino, Bologna 1994) è chiaramente dualista: “mi descrivo a me stesso come un dualista cartesiano” (*Knowledge and the Body-Mind Problem. In Defence of Interaction*, Routledge, Londra e N. York 1994, p. 5). E afferma più avanti: “sono decisamente fuori moda. La moda prevalente nella filosofia è il monismo, e lo è stata per lungo tempo” (*ibid.*, p. 8). Secondo Popper, nel corso dell’evoluzione emersero reali novità qualitative, come la *vita*, la *coscienza animale* e *l’uomo*, animale dotato di intelligenza creativa e di autocoscienza, la cui sede è il cervello. Niente al di sotto dell’uomo è simile a lui. La superiorità e la grandezza dell’uomo in tutto l’universo è dovuta alla sua intelligenza, capace di capire e di costruire infinite teorie, e alla sua libertà.

La parte superiore degli animali contiene organi di controllo dell'organismo. Popper si ispira in questo senso ai sistemi cibernetici (cfr. *Conoscenza oggettiva*, Armando, Roma 1975, sezione *Nuvole ed orologi. Saggio sul problema della razionalità e della libertà dell'uomo*, del 1965). Gli organismi viventi più complessi sono sistemi gerarchizzati di "controlli non rigidi" del corpo. In questi viventi, un sistema superiore, dotato di meccanismi di prove e di correzioni, modera o controlla fino a un certo punto una serie di sottosistemi inferiori (i quali a loro volta hanno i propri meccanismi di prove). Così le piccole deviazioni di un sistema inferiore vengono controllate complessivamente dal sistema superiore. A sua volta si verifica un retro-influsso (*feed-back*) del sottosistema inferiore sul sistema di controllo.

Questo modello si può applicare all'uomo. I nostri stati mentali controllano i nostri movimenti fisici, con un'interazione e nuovi meccanismi di *feed-back*. Il controllo "non rigido" ("plastico") si contrappone al dominio "ferreo" o automatico, e in parte ricorda l'idea aristotelica del dominio "politico" e non "despotico" dell'anima sulle sue passioni. Questo concetto è collegato all'indeterminismo in cui operano i sistemi inferiori, per cui è possibile orientarli in un certo senso anche se possiedono una propria autonomia dinamica. Popper critica l'idea cartesiana di causalità "rigida" della mente sul corpo, concepita secondo un modello meccanico. Le difficoltà per spiegare le modalità dell'interazione mente-corpo sono dovute al fatto che, anche se abbiamo superato la teoria cartesiana della causalità come impulso meccanico, non disponiamo di una teoria generale della causalità.

L'io o soggetto -persona- per Popper è attivo, anzi è la realtà più attiva dell'universo. Non è un puro flusso di processi, né un osservatore passivo. L'immagine platonica dell'io come "pilota" della nave del nostro corpo è positiva. Una delle funzioni dell'io è l'*integrazione* continua della condotta umana, funzione anche svolta dal sistema nervoso a diversi livelli gerarchici. L'integrazione operata dall'io consiste nella selezione degli elementi del passato, ma indirizzati verso il futuro, e anche la direzione dell'attenzione (scegliendo oggetti pertinenti e ignorandone altri). L'io arriva in un certo senso a modellare il proprio cervello.

Una tesi tipicamente popperiana è che l'io è influito dal mondo oggettivo che egli stessi crea o scopre (linguaggio, cultura, tecnica). L'io cresce proprio nella misura in cui si rende "oggetto" in tale mondo. Più che di dualismo, si potrebbe parlare in Popper di un certo "trialismo". I suoi celebri *tre mondi* (mondo fisico, soggettivo e oggettivo) interagiscono a vicenda e spiegano non solo la condotta e la personalità umana completa, ma anche il dinamismo culturale.

La filosofia di Popper sulla tematica mente-corpo è ricca in sfumature e non può essere

riassunta in poche righe. Egli ha dedicato numerose pagine alla critica delle filosofie materialistiche (comportamentismo, teoria dell'identità), anche se non ha conosciuto bene gli sviluppi del funzionalismo e delle scienze cognitive. Siamo di fronte a un vero filosofo -come Wittgenstein-, uno che indaga, che ha senso del mistero e si ammira sinceramente di fronte alla realtà. Nei suoi confronti, gli autori materialisti che dogmatizzano con arroganza sulla conoscenza diventano piccoli. Questi autori, dominati dallo scientismo, sono come "ingegneri della conoscenza" che ignorano che cosa sia il filosofare.

*John C. Eccles* (cfr. *L'io e il suo cervello*, cit., vol. 2; *Come l'io controlla il suo cervello*, Rizzoli, Milano 1994), neurofisiologo australiano, accoglie integralmente la teoria mente-corpo di Popper, aggiungendo la convinzione cristiana dell'immortalità della mente e della sua possibile infusione divina in alcuni animali superiori, nel corso dell'evoluzione.

Nelle sue ricerche filosofiche, Eccles tentò di elaborare un modello scientifico che, in un modo alquanto simile a Cartesio, incorporasse l'attività di controllo della mente immateriale su alcune parti della corteccia cerebrale. L'unità della cognizione dell'uomo e del suo dominio sul corpo -azioni intelligenti e volontarie- si spiega soltanto grazie ad un'azione unitaria e continua della mente immateriale sul cervello. Eccles si serve della meccanica quantistica per spiegare il modo in cui la mente, caratterizzata dall'autocoscienza, potrebbe agire su gruppi di cellule corticali piramidali (moduli), secondo campi quantici di probabilità. Questa spiegazione potrebbe aggirare la vecchia difficoltà della violazione del principio di conservazione (cfr. le nostre pagine su Cartesio, dove ci siamo riferiti a questo punto). La mente sarebbe impegnata nella lettura selettiva -fenomeno collegato all'attenzione e all'intenzione- di numerosi centri di attività corticale (aree di collegamento), integrando costantemente opportuni elementi nervosi e modificando i loro schemi dinamici spazio-temporali (*L'io e il suo cervello*). In definitiva, le funzioni della mente sarebbero la *lettura* dei messaggi e il *controllo* della dinamica cerebrale superiore.

Nell'opera *Come l'io controlla il suo cervello* Eccles aggiunge una nuova ipotesi. L'unità degli eventi psichici si potrebbe denominare *psicone*. Quest'unità dinamica sarebbe in grado di collegarsi a un fascio di dendriti apicali di cellule piramidali, denominato *dendrone*. Il dendrone quindi, concepito come un'unità anatomica cerebrale, sarebbe l'unità nervosa fondamentale della corteccia. "L'interazione io-cervello può essere dunque considerata sulla base dell'interazione unitaria di uno psicone con il proprio dendrone" (p. 117). Anche i mammiferi avrebbero psiconi collegati alla loro coscienza sensitiva.

Le teorie di Popper ed Eccles sono dualistiche (forse ancor di più quella di Eccles). Esse cercano di spiegare come la parte spirituale dell'uomo potrebbe interagire col cervello, ma non sono sostenute da una concezione metafisica e antropologica elaborata. L'idea dell'intervento della fisica quantistica per spiegare l'insediamento della mente nel cervello, condivisa in un modo simile da Roger Penrose in rapporto all'attività della coscienza, nasce dall'intento di prospettare una base fisica indeterministica nelle aree controllate dalla mente.

Oggi il dualismo è respinto energicamente da quasi tutti gli autori coinvolti nelle indagini cognitive. Un aristotelico potrebbe obiettare che il dualismo non riesce a spiegare bene l'unità dell'uomo. Ma il motivo del rifiuto di ogni forma di dualismo è la scelta pregiudiziale di lavorare esclusivamente con le categorie naturalistiche delle scienze empiriche. Forse il dualismo è inadeguato, ma i motivi che spingono al dualismo, in mancanza di migliori spiegazioni ontologiche, sono comprensibili. L'uomo non può essere ridotto a puro corpo. Esiste una dimensione della realtà superiore al corpo. L'esistenza di Dio, la fede cristiana negli angeli, la verità religiosa e anche filosofica dell'immortalità dell'anima e del valore della persona umana, sono elementi fondamentali della concezione della realtà, che il monismo materialista semplicemente elimina, svuotando così la vita umana di ogni valore e di ogni possibilità di trascendenza.

E. *Panpsichismo. David Chalmers.* Potrà sembrare strano, ma nell'ambito della filosofia della mente c'è un autore anomalo che sostiene una posizione panpsichista: David J. Chalmers, australiano (cfr. *The Conscious Mind*, Oxford Univ. Press, Oxford 1996). Inizialmente egli intende sostenere il carattere irriducibile della *coscienza soggettiva*, contro il comportamentismo e il funzionalismo. Dal momento che è possibile concepire un mondo di *zombies*, si conclude che la coscienza è irriducibile alla realtà fisica esterna e che non esiste una sopravvenienza logicamente necessaria di ciò che è mentale su ciò che è fisico. D'altra parte, nessuna corrispondenza degli eventi mentali con gli eventi fisici spiega che cosa siano gli eventi psichici. Quindi non sappiamo come nascono gli stati mentali. Tutto questo dimostra la falsità del materialismo: ci sono aspetti del mondo che stanno al di sopra delle proprietà fisiche e che sopravvivono su di esse in una maniera contingente.

In un primo momento Chalmers sembra sostenere un dualismo forte. Nel mondo ci sono due aspetti, *il lato fisico* e *il lato soggettivo*. A un certo punto, tuttavia, egli sostiene che l'organizzazione o informazione produrrebbe "coscienza" anche in soggetti con basi materiali diverse dall'uomo (ad esempio un robot), da cui si potrebbe concludere la corrispondenza o isomorfismo tra *stati coscienti* e *stati fisici informativi*. La coscienza quindi nasce

dall'informazione. Ogni oggetto dotato di informazione dovrebbe possedere un certo grado di coscienza. Pure un termostato o una pietra avrebbero qualche forma di coscienza, benché “minima”. Un computer dotato di IA potrebbe avere, invece, una mente tanto ricca quanto la nostra. Questa posizione, ammette Chalmers, è una varietà del panpsichismo (ci sembra una posizione vicina a di Leibniz).

Nell'esposizione di Chalmers non si scorge un vero spiritualismo. Le sue argomentazioni restano ancorate alla prospettiva funzionalista, solo che egli non riesce a ridurre la coscienza ad un'altra cosa. Di qui la strana idea di assegnare coscienza a qualsiasi realtà dell'universo che possieda informazione (in questo senso, anche l'universo dovrebbe avere una “mente”). L'assegnazione di una forma di coscienza a cose come una pietra o un robot sono, a mio avviso, un sintomo di materialismo. Il panpsichismo di Chalmers è un materialismo larvato.

F. *Teorie unitarie*. Nei circoli accademici della filosofia della mente non si considerano in genere altre posizioni all'infuori di quelle che abbiamo considerato finora. È vero, comunque, che negli altri approcci della filosofia, ad esempio nella fenomenologia o nell'esistenzialismo, il problema mente-corpo non si pone, di solito, o almeno non si presenta come un problema ontologico, come invece succede nella *Philosophy of Mind*.

I filosofi ispirati ad Aristotele e Tommaso d'Aquino non hanno elaborato ancora una visione sistematica della conoscenza adeguata al livello delle discussioni contemporanee della filosofia della mente. Tuttavia, come abbiamo detto all'inizio, la concezione tomasiana “naturalistica”, con l'importanza che assegna al cervello come sede organica della sensibilità superiore, intimamente collegata all'intelligenza, potrebbe essere un buon interlocutore nei problemi affrontati in queste pagine.

G. *Basti* forse è uno dei pochi autori tomisti che ha proposto alcune riflessioni sul problema mente-corpo in un modo vicine alle scienze cognitive (cfr. *Il rapporto mente-corpo nella filosofia e nella scienza*, Edizioni Studio Domenicano, Bologna 1991, e numerosi articoli anche più recenti). Basti trova nella filosofia di Aristotele e soprattutto in San Tommaso un approccio né monista, né dualista, ma *duale*, capace di rendere conto dell'unità operativa e ontologica tra l'intelligenza e la percezione sensitiva. La “mente” andrebbe concepita piuttosto come una forma animante il corpo organico dell'uomo, in modo tale da costituire con esso l'unità di una sola persona. Ma la mente esiste anche come una capacità cognitiva superiore, pienamente immateriale e capace di trasmettere informazione al cervello. La mente umana è dotata di autoreferenzialità e libertà nei confronti dei dati e della memoria, per cui può sempre ridefinire lo spazio dei problemi.

La causalità dell'intelletto sul corpo andrebbe capita, secondo Basti, non come scambio di energia, bensì come scambio di informazione tra mente e cervello. Il cervello può essere soggetto di questo scambio ed è guidato dall'intelligenza grazie alla sua enorme complessità, che va capita secondo i canoni della fisica dinamica della complessità, con comportamenti caotici non-lineari. La rivoluzione fisica del non-lineare e della complessità, più radicale della rivoluzione quantistica, ci consente di capire meglio la vita sensitiva, meglio del ricorso troppo ipotetico alla meccanica quantistica (come hanno fatto Eccles e Penrose). I sistemi dissipativi aperti (cioè non conservativi) generano informazione in un modo imprevedibile, secondo un flusso informativo che va nella direzione opposta a quella dei flussi energetici, vale a dire dai microstati ai macrostati (invece il sistema che solo dissipa energia va dal macrostato verso il microstato).

Il cervello, in questo senso, appare come un sistema complesso (il più complesso dell'universo), che manifesta un comportamento altamente biforcuto, incapace di essere descritto con l'analisi matematica classica. In questo modo si potrebbe spiegare meglio il funzionamento "connessionistico" del cervello. Questo punto, lontano dal determinismo e continuismo della fisica classica, se inizialmente si apre a una filosofia della vita costituita da "forme discontinue" e dal comportamento "auto-organizzativo", consentirebbe pure di comprendere il cervello come organo molto atto per l'apprendimento di *patterns* percettivi. In definitiva, una biologia filosofica aristotelica si trova in una particolare sintonia con gli sviluppi recenti della fisica, capaci di dar conto della dinamica cerebrale, dando così un quadro ontologico superiore alle teorie funzionalistiche, che sono state meno attente alla base neurologica dei nostri processi cognitivi.

Credo che bisognerebbe incorporare anche al dibattito la nozione di persona, assente nella filosofia della mente. Basta vedere la ricchezza delle analisi di K. Wojtyla (cfr. *Persona e atto*) sulla cognizione, la coscienza, le emozioni e il corpo umano, analisi incentrati sul concetto fondamentale di *persona umana*. In questo senso, la filosofia della cognizione dovrebbe essere elaborata in vicinanza all'antropologia, e non dovrebbe restare come una pura "filosofia della mente" isolata.

La nozione di persona umana, ovviamente, non è afferrabile a partire dallo studio operativo delle funzioni cognitive, sia a livello psicologico che neuropsicologico. Non si può arrivare a una "fondazione" della persona partendo da tale prospettiva presa in modo esclusivo (ma neanche la si può criticare validamente). Pur essendo facilmente conoscibile nell'esperienza comune, la persona umana non è analiticamente afferrabile. Essa non s'identifica con la totalità del corpo, in quanto è qualcosa di più di una "totalità" o insieme di elementi in rapporto (non risponde a un concetto "olistico"). Il miglior modo di afferrare il concetto di persona è pensarla come un

*soggetto unitario costituito da un atto spirituale, nucleo della libertà e dell'intelletto, intrinsecamente unito ad un corpo organico proprio.* Tale atto spirituale è capace di controllare il corpo in qualche misura, rispettando il suo dinamismo proprio. Ogni tentativo di scoprire lo spirito, la libertà, l'io o la persona attraverso la via scientifica è destinato al fallimento. Le scienze possono invece -anzi debbono- presupporre l'esistenza della persona nel loro studio dell'uomo. Questa presupposizione metafisica è anche una responsabilità etica.

## 11. Connessionismo

A. *Connessionismo e informazione.* Lungo la nostra esposizione ci siamo riferiti occasionalmente alle reti neurali. Vorrei adesso farne un cenno allo scopo di una loro valutazione filosofica in rapporto ai problemi "mente-corpo". Le cosiddette "reti neurali" (o "neuronal") sono un modo di processare informazione diverso dalla computazione eseguita dai computer costruiti secondo l'architettura di von Neumann o Turing. La ricerca cominciò negli anni '40 del XX secolo, come abbiamo già detto, si fermò di fronte al sopravvento dell'IA classica, dopo gli anni '50, ma riprese negli anni '80 con relativi successi, con autori di rilievo come McClelland e Rumelhart. Il modello basato sulla costruzione di reti neurali interattive è chiamato *connessionistico*. Si parla anche di *parallelismo*, nel senso che il nuovo metodo di processamento dell'informazione è basato sulla distribuzione lungo una rete e non concentrata su alcune unità. In quanto la rete funziona olisticamente in tutte le sue interconnessioni prese simultaneamente, tale modalità è caratterizzata dal "parallelismo", in contrapposizione alla "serialità" o successione passo dopo passo del flusso informativo dei computer tradizionali.

L'idea che il nostro cervello elabori informazione mediante collegamenti a rete tra i suoi miliardi di neuroni, e non come una semplice sequenza seriale che procede dall'input all'output, è quasi naturale. Del resto, non si osserva nel cervello alcuna unità centrale di processamento (CPU: *Central Processing Unity*), ciò che invece è essenziale nei computer simbolici. In definitiva, il nostro cervello -come quello degli animali superiori- non manifesta per nulla un'architettura tipo Turing. Il suo modo di "computare" è diverso. Ma qui bisogna far notare *il senso analogico dei verbi "computare" o "calcolare"* (anche le cellule, ad esempio, "computano" o "calcolano" i tempi, con i loro "orologi biologici", e così ricevono e inviano segnali in modo preciso e tempestivo). La computazione naturale eseguita dal cervello non si basa sul simbolismo vero e proprio, cioè sulla creazione di un linguaggio astratto (una questione diversa è che noi, col nostro pensiero, siamo in grado d'inventare un linguaggio e anche metodi di computare tipo Turing e di implementarli in macchine). Nel cervello non ci sono programmi né regole di processamento. Le

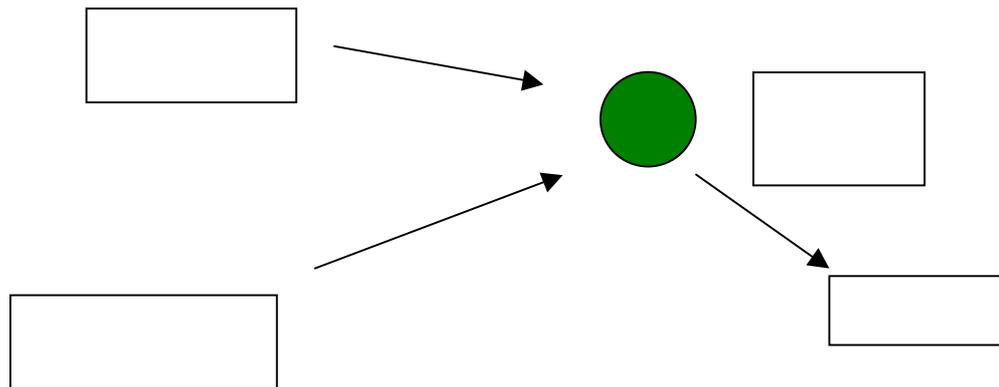
“regole concrete” sono contenute e modificate nelle interconnessioni che a poco a poco si stabiliscono tra le diverse aree cerebrali e neuroni.

Questo punto risulta specialmente chiaro se pensiamo alla percezione, all'apprendimento e alla memoria, ma può anche valere per compiti più alti, come la soluzione di problemi e l'apprendimento linguistico. Così si potrebbe spiegare, ad esempio, che un giocatore di tennis sia in grado di calcolare in modo rapidissimo e molto flessibile la traiettoria della palla in arrivo, tenendo conto del suo effetto, per decidere in pochi istanti il tipo di colpo con cui rispondere, per cui deve badare alla sua posizione nel campo da tennis e al posto dove vuole inviare la palla, controllando al contempo la sua posizione corporale e numerosi movimenti coordinati dei muscoli. Un computer “sequenziale” dovrebbe risolvere tutti questi problemi con una quantità enorme di calcoli molto complessi e quindi lo farebbe più lentamente. Eppure la velocità delle trasmissioni nervose è di gran lunga inferiore a quella dei circuiti elettronici, solo che l'incredibile velocità computazionale del nostro cervello è dovuta alla simultaneità dei miliardi di interconnessioni esistenti e operanti tra le cellule nervose (il cui numero, secondo alcune stime, sarebbe di 100 miliardi, vicino al numero delle stelle della Via Lattea). Ogni neurone è connesso a migliaia di neuroni da cui riceve segnali, e con altrettanti neuroni cui invia i suoi messaggi. Ogni connessione può essere eccitatoria o inibitoria.

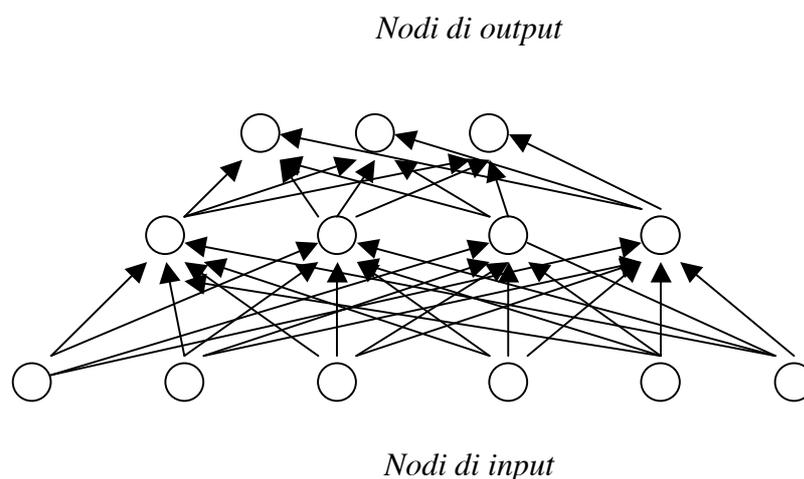
*La ricerca nel campo delle reti neurali consiste nella creazione di reti artificiali, ad esempio in circuiti elettronici o mediante simulazione al computer. A partire da alcune unità input o “nodi”, equivalenti ai “neuroni”, l'informazione si propaga lungo più connessioni, fino ad arrivare a un certo output o risultato finale che dipende dalle configurazioni dinamiche assunte dalla rete. Tali configurazioni dipendono dai valori di attivazione e d'inibizione di ogni singola unità in collegamento “sinaptico”. I nodi hanno ad ogni istante una soglia di attivazione. Il loro collegamento con gli altri elementi della rete viene determinato dai “pesi” (misurabili) delle connessioni. Questo dipende da una serie di fattori ed è determinante per l'effettiva attivazione o inibizione del nodo, cioè per la trasmissione del segnale e la costruzione di una rete. Ogni nodo prende la “somma pesata” di tutti i nodi da cui riceve segnali.*

“L'apprendimento” delle reti consiste nella modificazione dei valori di questi pesi, finché essa si stabilizza in una certa configurazione o *pattern*, dopo una serie di cicli di “addestramento”. I pesi determinanti delle interconnessioni sono alterati a seconda della variazione degli *input*, in funzione del conseguimento di una configurazione *target* nell'*output* (riducendo progressivamente i margini di errore dell'uscita).

Il seguente disegno illustra l'idea basilare della rete. Un neurone A, ad esempio, può avere una *soglia* di 2, cioè 2 è la minima quantità di *input* che causa la sua scarica. "A" può ricevere la spinta da parecchi neuroni, ciascuno dei quali ha un "peso". Se il neurone A riceve una spinta soltanto da 1, non potrà sparare. Si aggiungano poi i neuroni inibitori, con una forza negativa, come -1, -2, -3, ecc. In reti neurali più recenti questi valori non sono discreti ma variabili continui, incorporati in una dinamica statistica: la rete neurale in questo caso è analogica e non digitale.



Un modello molto semplificato di rete, con solo 13 unità disposte in tre *strati* (strato degli input, strato di "unità nascoste" e strato degli output), con propagazione in una sola direzione, è il seguente:



Il comportamento complessivo della rete dal punto di vista fisico è soggetto alla dinamica caotica non-lineare e quindi può essere seguito dalla matematica dei processi non-lineari. Si può impiegare in questo senso uno spazio degli stati dove un punto può rappresentare il *pattern* di simultanei livelli di attivazioni, oppure si può ricorrere a *vettori* di attivazioni.

Dal punto di vista filosofico, il problema di fondo è questo: in che senso la rete contiene,

elabora e trasmette informazione? *Il termine “informazione” è analogico.* Abbiamo detto prima che anche l'operazione di computare può essere intesa in parecchi sensi analogici, anche se il *primum analogatum* dell'informazione e della computazione è il suo significato logico-astratto, proprio del pensiero umano. Esiste una classe di informazione non legata propriamente alla conoscenza (ad esempio, l'informazione contenuta nel codice genetico), un'altra collegata alla conoscenza sensibile, e un'altra ancora che può rapportarsi alla conoscenza intellettuale, dove interviene il linguaggio. Il funzionamento a rete è una forma di associazionismo “induttivo”, che procede “dal basso” e non secondo regole “dall'alto”. La rete può essere in parte guidata da istanze normative “superiori”, per esempio, da un supervisore. È così come si costruiscono reti neurali artificiali “sotto supervisione” (anche se possono anche essere pure lasciate nelle mani di un dinamismo spontaneo, “senza supervisione”). Ma le reti possono anche essere guidate dagli istinti animali, o dal sapere superiore della persona, nel caso di reti particolari nel dinamismo del cervello. L'informazione non è localizzata in un'unità particolare, ma è distribuita, come abbiamo detto, ed è rappresentata dalla matrice dei pesi dei legami tra i nodi. L'evoluzione dell'informazione quindi è l'evoluzione di tale matrice.

*Siamo di fronte a due modelli di processamento dell'informazione.* Il modello *connessionistico* si basa sulle associazioni e può esistere in modo naturale nei processi vitali, in funzione dell'adattamento a fini biologici, ma può anche svolgersi in un ambito di rappresentazioni (immagini, parole, concetti), e può essere simulato artificialmente, contribuendo così alle tecniche computazionali dell'uomo. Il modello connessionistico è più vicino, nel campo cognitivo, all'ideazione in base all'induzione e all'analogia, ed è più simile alla strategia della vita e in qualche modo al pensiero spontaneo (alla “psicologia popolare”). Invece il modello *sequenziale* è piuttosto logico-deduttivo. Nasce dal pensiero razionale dell'uomo (regole e inferenze) e può essere implementato in macchine simboliche, che imitano appunto il pensiero linguistico e deduttivo a partire da regole scritte. Il modello simbolico risponde pure al nostro pensiero comune, poiché l'uomo ragiona normalmente in base a regole date.

Ciascuno di questi modelli serve a certi scopi meglio dell'altro in campi specifici, così come per alcune questioni è meglio ragionare con calcoli deduttivi, precisi ma rigidi, mentre per altre conviene lasciarsi guidare dal “peso” dei diversi *input* che abbiamo ricevuto, in definitiva dall'esperienza globale. Ad esempio, se certa soluzione è più efficace in certi contesti, tendiamo a ripeterla e a migliorarla, anche col rischio di sbagliare. “Perché è tanto difficile far riconoscere un albero a un calcolatore? La risposta a queste domande dipende in ultima analisi dal fatto che i problemi di riconoscimento delle forme non possono essere definiti in modo conciso. Per

riconoscere gli alberi è necessaria una definizione esauriente di cos'è un albero, e una definizione siffatta sarebbe equivalente a una descrizione di tutte le varianti concepibili. I problemi posti dal riconoscimento delle forme costituiscono un sottoinsieme dei cosiddetti problemi aleatori, problemi cioè la cui soluzione richiede sostanzialmente la conoscenza di ogni stato possibile del sistema (...) La soluzione di un problema di computazione classico, come la moltiplicazione, può essere in genere espressa succintamente con un algoritmo, cioè con una successione di istruzioni precise che indicano come si debba operare sui dati di ingresso per giungere alla soluzione del problema proposto” (Y. S. Abu-Mostafa e D. Psaltis, *Il calcolatore ottico neuronico*, “Le Scienze”, n. 225, maggio 1987, in *Quaderni*, 66, giugno 1992, p. 55). La strategia connessionistica consente di aggirare le difficoltà dei problemi algoritmicamente intrattabili, al prezzo di rinunciare all'impostazione analitica. Certe volte tendiamo a pensare secondo regole, mentre in altre occasioni ragioniamo piuttosto in base all'esperienza. Entrambi gli aspetti sono complementari e sono sfaccettature della conoscenza umana.

B. *Connessionismo e conoscenza*. Le reti neurali artificiali possono svolgere compiti come il riconoscimento di volti, di tipi di minerali o di diversi *patterns* visivi, acustici, ecc. Rumelhart e McClelland riuscirono a far imparare a una rete il passato di alcuni verbi in inglese (con risultati piuttosto modesti). In questi casi, non è che le reti abbiano vere immagini o rappresentazioni interne, dal momento che sono soltanto macchine. Il problema in questo caso è analogo a quello dei computer tradizionali. La rete neurale contiene l'informazione necessaria per poter adempiere compiti “cognitivi” senza simbolismo e senza rappresentazioni interne. Comunque, se la rete lavora con elementi di per sé simbolici (ad esempio, per imparare a riconoscere lettere, immagini, configurazioni), siamo noi che ci serviamo dei suoi risultati, interpretandoli nella nostra dimensione veramente cognitiva. D'altra parte, il nostro cervello funziona a rete nell'ambito veramente rappresentazionale e concettuale.

Prendiamo ad esempio il caso della *memoria*. Il connessionismo corrisponde al modo di agire della memoria associativa. Aristotele si era già accorto che il dinamismo della memoria correva a carico dell'associazione dei ricordi. La memoria dei computer simbolici è basata sull'immagazzinamento in un settore fisso (“domicilio” o archivio). La memoria a rete si costruisce, invece, sulla base dell'associazione tra molti elementi. Individuiamo un amico grazie ai ricordi associati del suo nome, la sua faccia, un incontro personale, un regalo ecc.). Gli schemi percettivi e mnemonici non sono “schede” situate in qualche punto fissi del cervello, ma piuttosto sono costituiti, nella loro base neurologica, da circuiti cerebrali a rete.

Si pensi al *riconoscimento di parole o lettere*. Riconosciamo facilmente la lettera A, ad esempio, anche se viene scritta in modi diversi (A, a, A, **A**), grazie ai contesti associativi. Se leggiamo la parola *arm\*dio*, al posto dell'asterisco tendiamo a vedere una *a*. Le reti si dimostrano flessibili e anche in grado di superare “rumori” informativi, pure erronei o contraddittori. Anche se vediamo una A semicoperta, il peso delle connessioni può aiutare a decidere che è una A. Se vediamo scritto *arm*, per fare un altro esempio, si attivano i nodi *arm* come braccio in inglese, *armadio*, *arma*, mentre invece il vincolo con la parola *letto* sarà inibitorio. Il *pattern* della rete deciderà alla fine di quale parola si tratta.

La strategia connessionistica può servire anche per *la rappresentazione dei concetti* tramite reti neurali. Il concetto, ad esempio, di *stanza*, può essere rappresentato da “microtratti”, cioè proprietà prototipiche (di una stanza) di livello inferiore. Il concetto *stanza* si attiverà quando saranno eccitate simultaneamente e in modo interattivo i microtratti di “stanza”, ad esempio, “avere porta”, “avere finestre”, “avere mura”, “avere soffitto”, ecc. Rumelhart ha costruito una rete di questo tipo. Ad un livello più specifico, si possono considerare *tipi di stanza* in collegamento a *patterns* di attivazione di altre unità in rapporto al concetto di stanza, come “avere un letto”, “avere tappeti”, “avere un televisore”, “avere un sofà, e così via.

Ma alcune reti rappresentano oggetti in una maniera non semantica, ad esempio, per distinguere le mine dalle rocce. In questo senso, la discriminazione non è concettuale, in quanto è basata sui “profili” degli echi provenienti da questi oggetti e rilevati da un sonar, così come un animale, senza avere concetti, può effettuare discriminazioni cognitive (su questi punti, cfr. Marraffa, *Scienza cognitiva*, cit., pp. 130-157).

Il connessionismo tocca qui il problema del significato concettuale, su cui le scienze cognitive si sono occupate (reti semantiche, concetti, *scripts*, problemi di intenzionalità e di riferimento, ecc.). I concetti predefiniti in modo simbolico e secondo regole logiche corrispondono per lo più al modello computazionale classico. Il connessionismo può rapportarsi all'acquisto di concetti tramite l'esperienza.

*C. Significato e utilità del connessionismo.* Dal punto di vista *pratico*, le reti neurali sono utili strumenti appartenenti alla ricerca dell'IA, come una via alternativa a quella dei computer tradizionali. I vantaggi o svantaggi pratici di queste due vie costituiscono un problema tecnico. Le reti possono arrivare con più successo laddove l'IA classica si rivela meno efficace, ad esempio nel riconoscimento di configurazioni. Ma esiste la possibilità di combinare le due metodologie. Difatti oggi si costruiscono alcuni computer basati sul parallelismo.

Dal punto di vista *teoretico*, le reti neurali sono modelli parziali di come funzionano il nostro cervello e i nostri processi cognitivi. In questo senso, il connessionismo affievolì il prestigio del funzionalismo e screditò la metafora del cervello come computer. Il connessionismo getta una luce sul modo reale di funzionare del nostro cervello come organo di elaborazione psicologica dell'informazione, aprendo così nuovi orizzonti al problema mente-cervello, orizzonti che erano chiusi con l'approccio eccessivamente computazionale delle scienze cognitive.

Quando si sente dire che la mente umana “è come un computer”, o che il nostro cervello “è come una rete neurale”, sarebbe utile ricordare che è piuttosto il contrario. Aristotele aveva detto che l'ingegnosità artificiale della ragione umana imita i processi della natura. Questo antico principio si è avverato in un modo sorprendente con la scoperta delle macchine informatiche. I computer simbolici sono simili al modo di pensare dell'uomo proprio perché è stato l'uomo a crearli, prendendo spunto dai processi cognitivi naturali, trovati nel funzionamento ordinario e scientifico della nostra intelligenza. Le reti neurali si sono ispirate a ipotesi sull'elaborazione cerebrale dell'informazione.

Ovviamente la ricerca nel campo informatici può anche aiutarci a migliorare i nostri modi di pensare e anche può dare luce agli studi neurologici. Possiamo creare artificialmente situazioni virtuali per trovare soluzioni ai nostri problemi reali e per confrontarle con ipotesi neurologiche. Ad esempio, certe disfunzioni delle reti neurali suggeriscono analogie con patologie del cervello. Le lesioni cerebrali spesso indeboliscono il rendimento intellettuale, ma non lo bloccano di un colpo. Analogamente, la perdita naturale di neuroni con l'età non comporta di per sé patologie specifiche di tipo neuropsicologico. Questo fenomeno è possibile soltanto nei sistemi a rete. La loro “degradazione” è blanda o graduale (e si può superare, talvolta, con nuove connessioni). Invece nei computer tradizionali normalmente un guasto blocca tutto l'apparecchio.

D. *Complementarietà tra connessionismo e simbolismo.* Nella ricerca cognitiva degli ultimi anni si è verificato un dibattito tra gli autori connessionisti (Smolensky) e la linea rappresentazionista o simbolica (Fodor, Pylyshyn, McLaughlin). E più giusta l'immagine *connessionistica* della mente o quella *sintattica* (con la tesi del linguaggio della mente o mentalese)?

Nei tempi di oro del funzionalismo sembrava vincente il “modello sintattico della mente”. Ma il connessionismo spiega meglio la capacità adattativa della nostra mente/cervello, la sua incredibile velocità e semplicità nel dare risposte, il suo rapporto agile ed efficace con i contesti rilevanti (selezionando ciò che è giusto), l'apprendimento a partire da esempi, in modo induttivo e con una

generalizzazione progressiva, e il fenomeno menzionato della degradazione graduale dei poteri mentali. I computer simbolici non hanno queste capacità, poiché debbono affrontare i loro problemi in una maniera “deduttiva”. I seguaci della linea simbolica hanno replicato che il connessionismo non spiega il carattere strutturato, produttivo e sistematico del linguaggio, ad esempio, la capacità di costruire enunciati semplici e complessi attraverso procedimenti ricorsivi, e hanno cercato di difendere l’immagine “modulare” della mente sostenuta da Fodor.

Probabilmente le due parti della contesa hanno ragione parzialmente. Le loro proposte dovrebbero essere integrate e adeguate a compiti specifici della mente. I nostri processi cognitivi possiedono aspetti connessionisti e modulari, simbolici e subsimbolici. La percezione, la memoria e certe abilità motorie sono meglio spiegate dai modelli connessionisti, ma anche qui bisogna ammettere un certo modularismo. Peraltro, non basta parlare di associazioni tra neuroni per capire il funzionamento nervoso. Bisogna tener conto dei livelli con cui affrontare il suo funzionamento. Molecola, membrana, cellula, circuito, rete, mappa, sistema cerebrale, linguaggio, concetti: ciascuno di questi livelli opera “computazioni” in vario modo.

In generale, *il connessionismo è un modello più credibile per le funzioni cognitive non simboliche*, corrispondenti alla sensibilità e all’esperienza. Il simbolismo astratto è una creazione della mente razionale, al di sopra delle capacità organiche del cervello. Il simbolismo astratto comporta universalità e una completa astrazione, quindi si colloca sul piano dell’immaterialità dell’intelligenza. La base organica di tale simbolismo sono le immagini, l’esperienza, i ricordi, gli schemi percettivi, ma in particolare le immagini linguistiche, e tutto ciò è regolato da canoni neuroassociativi. L’astrazione logico-linguistica rende l’uomo idoneo per la costruzione di macchine di computazione (macchine di Turing, reti neurali artificiali). Ma il modo di fare tipico del connessionismo può incorporarsi nei processi cognitivi più alti, al servizio dell’induzione, della generalizzazione e del pensiero analogico ed emotivo.

Se accettiamo la distinzione tradizionale di Tommaso d’Aquino tra *intellectus* (visione immediata delle cose essenziali) e *ratio* (*noús* e *lógos*, in Aristotele), forse si potrebbe dire che il connessionismo può far parte del dinamismo euristico e contemplativo del *noús*, mentre il puro simbolismo delle macchine calcolatrici apparterebbe alla funzione del *lógos* formale (logico-matematica).

Forse il connessionismo potrebbe pure aiutare a capire meglio alcuni aspetti del nostro modo inconscio di “pensare”, basato su uno sfondo di abiti e rappresentazioni non sempre richiamata alla coscienza. Osserva Penrose che l’unità della nostra percezione cosciente “mi pare del tutto in

disaccordo con l'immagine di un computer parallelo. Questa immagine potrebbe, d'altra parte, essere più appropriata come modello per l'azione *inconscia* del cervello. Noi siamo in grado di compiere simultaneamente e in modo più o meno automatico varie azioni indipendenti -come camminare, abbottonare, respirare e persino chiacchierare- senza essere coscienti di *alcuna* di esse!" (R. Penrose, *La mente nuova dell'imperatore*, cit., p. 505).

La sequenzialità dei nostri pensieri coscienti, invece, probabilmente è dovuta alla sequenzialità del linguaggio e ai tempi del dinamismo delle attivazioni cerebrali. Tale temporalizzazione cognitiva è unificata grazie al fenomeno dell'attenzione, con la distinzione tra foco dell'attenzione e conoscenze concomitanti e di sfondo. Non possiamo pensare molte cose alla volta e nemmeno possiamo prestare attenzione a più di una cosa per volta (sequenzialità), ma possiamo concentrare l'attenzione su un campo aperto implicitamente a molte relazioni e a certi sottocampi. Ma il vero centro della nostra percezione e delle nostre rappresentazioni che vanno e vengono siamo *noi stessi*, cioè l'unità della persona, padrona di se stessa nei suoi momenti coscienti.

## 12. Approcci neurofisiologici recenti

Il relativo declino del funzionalismo suscitò un rinnovato interesse verso le neuroscienze, anche se a volte mescolato con nuove tentazioni materialistiche. La neurobiologia e il connessionismo hanno contribuito ad una migliore comprensione del funzionamento psicofisico del cervello. Questo fatto è coerente con il vincolo essenziale esistente tra pensiero e cervello, secondo la concezione antropologica di Aristotele e di Tommaso d'Aquino. Non è questa la sede per la considerazione delle ricerche neurofisiologiche attuali e il loro impatto sulla filosofia della mente. Mi limiterò, per finire questo corso, a fare solo un cenno su alcuni autori contemporanei di riguardo nel campo della neuroscienza (ometto, quindi, la visione del russo Luria, interessante ma situata nei decenni vicini alla metà del XX secolo). Le teorie neurologiche di questi autori sono collegate alle tematiche cognitive esaminate in queste pagine.

*Antonio R. Damasio*, portoghese studioso negli USA (Iowa), non appartiene a nessuna corrente filosofica concreta. I suoi studi e le sue pubblicazioni divulgative (*L'errore di Cartesio*, del 1994, e *The Feeling of What Happens. Body and Emotion in the Making of Consciousness*, 1999; it. *Emozione e coscienza*, 1999) si sono occupate delle emozioni e delle basi neurofisiologiche della coscienza. In quanto segue mi riferisco esclusivamente a quest'ultima opera.

Damasio ha proposto la distinzione tra la coscienza *nucleare* -diversa dal semplice stato di

veglia- e la coscienza *autobiografica* o *estesa*, distinzione parallela in qualche modo a quella tra memoria a breve termine e a lungo termine. Egli ha evidenziato, soprattutto, il ruolo delle emozioni, sentimenti e sensazioni relative al proprio corpo (il “proto-sé”, di natura biologica e preconsua) nella formazione della coscienza di noi stessi (il *Sé*). D’altra parte, la coscienza non si riduce al linguaggio. Nella formazione della coscienza contano le immagini, diverse dalle rappresentazioni linguistiche. Le immagini sono dinamicamente collegate tra loro tramite la formazione di *mappe* o di *configurazioni neurali* (circuiti e reti tra neuroni). Gli oggetti esterni, nonché i diversi settori della nostra corporeità, vengono rappresentati da *mappe neurali in costante interazione*. Il cervello si costruisce una mappa dell’ambiente esterno, usando i propri parametri, e anche mappe dello stato interno dell’organismo.

La posizione di Damasio in *The Feeling of What Happens* non è riduzionistica. Il suo studio si sofferma al livello fondamentale, senza entrare nelle forme superiori della coscienza morale, artistica o scientifica. Egli riconosce l’esistenza di due livelli descrittivi non equiparabili: uno per la mente e l’altro per il cervello. Di conseguenza, critica il comportamentismo e il funzionalismo, sostenendo addirittura la convenienza di fare psicologia contando con le descrizioni in prima persona, come egli ha fatto con i suoi pazienti.

La linea scientifica inaugurata da Damasio rappresenta uno sviluppo positivo nelle scienze cognitive, che consentirà di esaminare meglio i rapporti tra pensiero e cervello. Egli considera superato il rifiuto di considerare la realtà fenomenica della coscienza, così come ugualmente è superata l’idea che i programmi di computer comportino un modello adeguato per la mente umana.

L’impostazione neurologica di *Gerald Edelman* (cfr. *The Remembered Present*, 1989, it. *Il presente ricordato*, Rizzoli, Milano 1991; *La materia della mente*, Adelphi, Milano 1993), ugualmente contraria al funzionalismo, è più incline al riduzionismo naturalista. Egli sostiene che si potrebbe fare una teoria completa dei fenomeni cognitivi dell’uomo basata sullo studio del cervello. In realtà, quello che veramente fa è uno studio della possibile base neuronale degli atti e funzioni psichiche fondamentali, presupponendo la loro esistenza e il loro senso per la vita umana. Ovviamente questi studi sono utili, *ma sono sempre parziali*. Non spiegano tutto della mente umana, e neanche i suoi aspetti essenziali, che non possono essere toccati dalla prospettiva neurologica. La neuroscienza si riferisce soltanto alla base materiale della mente, presupponendo la mente affinché tale spiegazione abbia senso e interesse per noi. Credere il contrario è una posizione filosofica, simile a quella delle filosofie “fiscalistiche”, la cui tesi era che solo la scienza naturale forniva una conoscenza valida. Tale tesi non appartiene alla scienza naturale: è metascientifica, e

quindi si auto-contraddice.

La proposta scientifica di Edelman si basa sull'idea di una *formazione progressiva selettiva di numerose mappe neurali interattive*, variabili negli individui e nei tempi, a seconda della loro storia, tenendo conto anche di fattori ereditati. Il carattere selettivo si comprenderebbe alla luce di concetti biologici come la competizione selettiva nelle popolazioni di neuroni e l'adattamento all'ambiente. Nell'ambito della percezione, le diverse mappature riuscirebbero a "categorializzare" gli oggetti in assenza di programmi o istruzioni, cioè in base a selezioni che condurrebbero a un'auto-organizzazione. La percezione sarebbe, quindi, una discriminazione adattativa di un oggetto o evento rispetto ad altri fenomeni, in uno sfondo dato. Ulteriormente si procederebbe a una generalizzazione, tramite la quale gli individui di una collezione sarebbero trattati come equivalenti.

Uno dei meccanismi con cui Edelman spiega i rapporti tra le mappature è il *rientro*, vale a dire *lo scambio di segnali tra molteplici mappe* che ricevono input diversi (visivi, come bordi, sagome, colori, o altro), per cui si creano di continuo nuove correlazioni parallele. Verrebbero così a stabilirsi rientri continui e paralleli tra le mappature. Grazie al rientro e al principio selettivo, le operazioni delle mappature sarebbero sempre meglio collegate. La "rappresentazione" degli oggetti risulta dunque distribuita in molte aree del cervello, grazie alle mappature globali consentite dai rientri tra le mappature locali. Con questi elementi, Edelman spiega il coordinamento tra la percezione e l'azione, la memoria come continua ricategorizzazione, l'apprendimento e la coscienza. La sua distinzione tra coscienza *primaria* e coscienza *superiore* è analoga a quella di Damasio tra coscienza nucleare e autobiografica.

L'importanza in Edelman della coscienza fenomenica e l'approccio non simbolico per la spiegazione dei processi cognitivi comportano ovviamente una visione antifunzionalistica. Egli ha criticato l'idea che il cervello possa essere paragonato a una macchina di Turing, e non è neanche soddisfatto con il connessionismo come modello utile per la neuroscienza. Il cervello come organo cognitivo andrebbe visto in chiave biologica, e anche il pensiero andrebbe visto come radicato o "incorporato" in un organismo.

In una linea simile, che privilegia le connessioni interattive tra neuroni e gruppi o "coalizioni" neuronali, si collocano le indagini sulla coscienza visiva di *F. Crick* -autore fortemente riduzionista- e *Christof Koch* (cfr. *Il problema della coscienza*, "Le Scienze", n. 291, novembre 1992).

La consapevolezza di ciò che si sta vedendo è legata, secondo Crick, all'attenzione e ai movimenti oculari, e troverebbe la base neuronale non tanto in una zona particolare della corteccia,

quanto in una distribuzione di attivazioni su certe aree dove si formano reti, forse anche grazie alla sincronizzazione dei ritmi delle scariche neuronali. Si cerca così una comprensione della base nervosa dello *stato di coscienza e dell'attenzione*, in questo caso limitato alla vista. “Nel cervello deve essere immagazzinata anche una rappresentazione latente di un volto, probabilmente sotto forma di un particolare insieme di connessioni sinaptiche tra neuroni (...) Per esempio ognuno di noi ha probabilmente una rappresentazione della Statua della Libertà nel suo cervello, rappresentazione di solito inattiva. Se si pensa alla statua, la rappresentazione diventa attiva, e si ha la scarica dei relativi neuroni” (F. Crick - C. Koch, *Il problema della coscienza*, cit., p. 128).

Le connessioni neurali sono quindi, il punto più interessante per la ricerca della base neurale dei processi cognitivi. Parlando della percezione, Damasio segnalava, nella stessa linea, che una tazza di caffè, ad esempio, potrebbe “evocare le rappresentazioni visive e tattili riguardanti forma, colore, materiale di cui è fatta, nonché profumo e gusto della bevanda o traiettoria che mano e braccio devono compiere per portare la tazza dal tavolo alle labbra. Tutte queste rappresentazioni vengono formate in regioni distinte del cervello, ma la loro ricostruzione avviene in modo sostanzialmente simultaneo” (A. R. Damasio e H. Damasio, *Cervello e linguaggio*, in *Le Scienze*, n. 291, novembre 1992, p. 70).

Siamo arrivati alla fine del nostro percorso storico. La collaborazione tra psicologia e neurologia sarà la chiave per arrivare ad una visione più adeguata dei nostri atti cognitivi e affettivi. In questo lungo cammino vanno evitate le posizioni riduttive. Il materialismo offre una visione piatta dell'uomo e lascia in ombra gli aspetti più importanti della persona umana. Invece lo spiritualismo, il personalismo e soprattutto la filosofia della conoscenza, ancorata al realismo metafisico e antropologico di San Tommaso, hanno tutto da guadagnare dalle ricerche cognitive. Solo rimane in dietro è l'unilateralismo, che non è scienza, ma mancanza di scienza.