

NATURA UMANA NATURA ARTIFICIALE

a cura e con introduzione di
Maria Cristina Amoretti

Scritti di
Evandro Agazzi, Michele Di Francesco,
Marcello Frixione, Giulio Giorello, Roberta Lanfredini,
Massimo Marraffa, Maurizio Martelli, Simona Morini,
Sandro Nannini, Pietro Perconti e Silvano Tagliagambe

E

Epistemologia

FrancoAngeli

Epistemologia, collana diretta da Evandro Agazzi

Comitato scientifico:

prof. Marco Buzzoni, ordinario di Filosofia della Scienza, Università di Macerata
prof. Fabio Minazzi, ordinario di Filosofia Teoretica, Università dell'Insubria, Varese
prof. Gino Tarozzi, ordinario di Filosofia della Scienza, Università di Urbino

La collana intende venire incontro a quell'esigenza, ormai generalizzata, di conoscenza epistemologica che si riscontra a li vello di cultura medio-alta e che corrisponde, in senso lato, alla diffusa aspirazione a prender coscienza critica della complessa varietà della nostra civiltà scientifico-tecnologica. Aspirazione che si accompagna, altresì, al desiderio di venire in chiaro circa lo statuto epistemologico di molte discipline le quali solo di recente hanno rivendicato l'impegnativa qualificazione di «scienza», pur riguardando ambiti di ricerca non inclusi nell'alveo delle discipline scientifiche tradizionali.

Rispetto ad analoghe collane già esistenti, questa si propone anche di allargare l'ambito delle scuole e tradizioni epistemologiche finora più correntemente conosciute in Italia, e che si ispirano in prevalenza al filone analitico anglosassone, portando l'attenzione su opere e autori afferenti ad altre aree culturali, come ad esempio quelle di lingua francese, tedesca, polacca.

Verranno quindi pubblicati, sia in traduzione che in opere originali, alcuni testi base di carattere istituzionale relativi all'epistemologia generale e alle diverse branche della filosofia della scienza. Per altro verso, verrà dato uno spazio più cospicuo del solito all'epistemologia delle scienze «umane», alla filosofia della logica, alle tematiche etiche che di recente si sono aperte nei riguardi della scienza. Pur senza rinunciare ad opere di carattere tecnico, l'accento generale verrà posto piuttosto su quei tipi di trattazione epistemologica nei quali è più presente un taglio specificamente filosofico.

La collana si propone di essere utilizzabile anche per corsi universitari: a tale scopo, oltre alle opere di carattere istituzionale cui si è fatto cenno, annovererà anche alcuni «readings» antologici, sia a carattere miscelaneo che monografico.

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

NATURA UMANA NATURA ARTIFICIALE

a cura e con introduzione di
Maria Cristina Amoretti

Scritti di
Evandro Agazzi, Michele Di Francesco,
Marcello Frixione, Giulio Giorello, Roberta Lanfredini,
Massimo Marraffa, Maurizio Martelli, Simona Morini,
Sandro Nannini, Pietro Perconti e Silvano Tagliagambe

FrancoAngeli

Questo volume è stato pubblicato con il contributo del MIUR - PRIN 2007-2009 *Credenza, giustificazione e natura. Problemi, prospettive e applicazioni del naturalismo filosofico contemporaneo* (inserito nella ricerca nazionale *Filosofia e naturalizzazione del mentale. Nuove prospettive sulla conoscenza e la natura umana*).

Copyright © 2010 by FrancoAngeli s.r.l., Milano, Italy.

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

Indice

Introduzione. Sul confine tra umano e artificiale di <i>Maria Cristina Amoretti</i>	pag.	7
Atteggiamenti proposizionali e modularità della mente di <i>Massimo Marraffa</i>	»	43
Operazionalità e intenzionalità: l'anello mancante dell'intelligenza artificiale di <i>Evandro Agazzi</i>	»	63
Riferimento, scienza naturale, modelli artificiali di <i>Marcello Frixione</i>	»	79
<i>Qualia</i> privati e pubblici. Cosa si prova a sentire che una mano è la propria di <i>Pietro Perconti</i>	»	99
Fenomenologia naturalizzata e fenomenologia naturale. Il problema degli stati qualitativi di <i>Roberta Lanfredini</i>	»	117
Orologi, menti e cervelli: riflessioni preliminari su tempo reale e tempo fenomenico tra fisica teorica e filosofia della mente di <i>Sandro Nannini</i>	»	135

Natura & cultura 2.0. Il soggetto tra neuro-riduzionismo e mente estesa di <i>Michele Di Francesco</i>	pag.	155
Naturale e artificiale di <i>Simona Morini</i>	»	181
La relazione naturale/artificiale tra rappresentazione, ibridazione, organizzazione di <i>Silvano Tagliagambe</i>	»	195
Natura e ideologia di <i>Giulio Giorello</i>	»	217
Epilogo di <i>Maurizio Martelli</i>	»	231
Gli autori	»	235
Indice dei nomi	»	237

Introduzione.

*Sul confine tra umano e artificiale**

Maria Cristina Amoretti (Università degli Studi di Genova)

1. Mente e natura

La dicotomia tra umano e artificiale¹ può essere efficacemente pensata in relazione al pensiero di René Descartes. Secondo il filosofo francese, infatti, tutti gli esseri viventi sono “macchine” biologiche più o meno sofisticate. Anche gli esseri umani, in quanto esseri viventi, sono quindi macchine, ma, a differenza di qualunque altra creatura naturale o artificiale, essi possiedono una mente, o con le parole dello stesso Descartes: «Sono dunque in poche parole solamente una cosa pensante, cioè una mente, o un animo, o un intelletto, o una ragione» (1641, p. 167). L’idea che sia la nostra mente – o per lo meno alcuni aspetti della nostra mente – a renderci esseri *umani*, e a differenziarci pertanto sia dagli animali non umani sia dalle macchine artificiali, appare plausibile anche ai giorni nostri, almeno al livello del senso comune². Ma se per Descartes la mente (*res cogitans*) era una sostanza diversa e indipendente rispetto al corpo e, più in ge-

* Desidero ringraziare Niela Vassallo, che ha letto una precedente versione di questa introduzione, fornendomi indicazioni e osservazioni preziose.

1. Nonostante la distinzione, peraltro confusa, tra umano e artificiale possa essere applicata a molti campi diversi della filosofia (come, per esempio, alla filosofia della conoscenza, della scienza, della biologia ecc.), è opportuno premettere che in questa introduzione mi concentrerò soprattutto su questioni legate alla filosofia della mente.

2. Sebbene il dualismo ontologico sia una concezione presente già nella filosofia antica, molti ritengono che il concetto contemporaneo di mente possa essere ricondotto proprio a Descartes (cfr., per esempio, Nannini 2002).

nerale, alla materia (*res extensa*), molti studiosi rifiutano oggi l'idea che esistano due domini diversi, quello fisico e quello mentale³, sposando una qualche forma di "monismo". Come spiegano Jaegwon Kim, «quando usiamo espressioni come "avere una mente" [...] non c'è bisogno di supporre che ci siano cose nel mondo chiamate "menti" che noi abbiamo [...]. L'avere una mente può essere pensato semplicemente come una *proprietà*, una *capacità*, una *caratteristica* che gli esseri umani e alcuni animali più evoluti possiedono, a differenza di cose come le matite e le rocce» (1998, p. 5). In altre parole, sebbene la mente sembri "speciale" e apparentemente estranea al mondo fisico, essa non sarebbe ontologicamente distinta dalla materia. In questo senso, quindi, la mente sarebbe parte della natura e di conseguenza potrebbe essere indagata con i metodi offerti dalle varie scienze empiriche.

La questione naturalismo/anti-naturalismo occupa buona parte del dibattito filosofico contemporaneo, ma è difficile trovare una definizione precisa e univoca del termine "naturalismo", poiché esso è stato applicato a tesi filosofiche radicalmente diverse le une dalle altre⁴. In termini molto generali, una posizione filosofica che si proponga di essere naturalistica dovrebbe sostenere che la realtà è esaurita dalla natura, che non esistono oggetti "soprannaturali" e che pertanto si devono accettare esclusivamente quelle entità e quei concetti che possano essere indagati dalle scienze empiriche (almeno in linea di principio). Il metodo scientifico, inoltre, dovrebbe essere l'unico metodo utilizzato per indagare qualsiasi area della realtà, inclusa ovviamente la mente umana. Si può quindi constatare come il naturalismo abbia almeno una componente ontologica (non esistono ulteriori entità oltre a quelle accettate dalle scienze empiriche) e una componente metodologica (sono i metodi accettati dalle scienze contempo-

3. In realtà, i molti problemi che – come vedremo – minacciano il monismo (specialmente in riferimento agli aspetti fenomenologici della mente) sembrano poter legittimare l'adozione di nuove forme di dualismo. Per una rassegna e una discussione delle principali teorie dualiste contemporanee a proposito della mente vedi per esempio Lavazza (2008).

4. Per un'introduzione ai problemi del naturalismo cfr. Agazzi e Vassallo (1998), De Caro e Macarthur (2005), Papineau (2007).

ranee i soli che ci possano portare a conoscere come stanno veramente le cose).

Ma se è convinzione diffusa che la mente sia parte della natura e come tale possa essere indagata e spiegata con gli strumenti e i metodi offerti dalle scienze empiriche (come ad esempio la biologia, la neurofisiologia, l'intelligenza artificiale, la psicologia cognitiva, la linguistica cognitiva e le neuroscienze), fino a che punto si può spingere una sua naturalizzazione? È possibile costruire macchine artificiali che riproducano, almeno in parte, caratteristiche essenziali dell'attività cognitiva umana? La realizzazione di macchine "intelligenti" può avere delle ripercussioni sull'immagine che abbiamo di noi stessi in quanto esseri umani? L'immagine scientifica della mente può essere compatibile con la nostra concezione ordinaria o ne rappresenta una correzione più o meno radicale? Per rispondere a queste domande bisogna innanzi tutto definire in modo più preciso quali siano le principali caratteristiche che, intuitivamente, attribuiamo al mentale, per indagare poi se sia possibile o meno integrarle all'interno di una spiegazione naturalisticamente accettabile della mente ed eventualmente riprodurle su macchine artificiali.

1) In genere, diamo per scontato che i nostri stati mentali abbiano degli effetti fisici e che, viceversa, gli stati fisici abbiano degli effetti sui nostri stati mentali. Più precisamente, riteniamo che gli stati mentali possano causare stati fisici ed essere a loro volta causati da stati fisici. Per esempio, il *desidero* di salutare un amico che si trova dall'altra parte della strada e la *credenza* che alzando e agitando la mano potrei soddisfare tale desiderio *causano* l'attivazione di certi neuroni che scaricano impulsi elettrici verso i muscoli che, a loro volta, si contraggono per produrre il movimento del braccio verso l'alto e poi quello della mano. Ma come possono degli stati mentali come credenze e desideri (che sono apparentemente estranei al mondo fisico) *causare* l'attivazione dei neuroni del mio cervello? Per il principio di chiusura causale del mondo fisico, infatti, gli stati fisici possono causare ed essere causati solo da altri stati fisici. L'*efficacia causale*, dunque, è una prima caratteristica che vorremmo poter attribuire al mentale, sebbene essa sembri mal conciliarsi con la nostra convinzione secondo cui il mentale non è riducibile al fisico.

2) Almeno alcuni stati mentali hanno la proprietà di essere diretti, di puntare o di tendere verso qualcosa (un oggetto o contenuto inten-

zionale) che potrebbe anche non essere presente o non esistere⁵. Tale proprietà è indicata col termine *intenzionalità*⁶, ma è più opportuno distinguere tra *intenzionalità del riferimento* e *intenzionalità del contenuto*. Nel primo caso, s'intende la proprietà di uno stato mentale di riferirsi a, o di riguardare qualcosa (*aboutness*): le credenze, i desideri, le intenzioni, le speranze ecc. sono sempre la credenza, il desiderio, l'intenzione, la speranza *di* qualcosa. Per esempio, credere che un cane stia abbaiano al di là della porta è uno stato mentale di credenza che *si riferisce a/riguarda* l'abbaio del cane al di là della porta. Nel secondo caso, invece, ci si riferisce al fatto che molti stati mentali hanno un contenuto intenzionale. Gli stati intenzionali – o, da Bertrand Russell in poi, atteggiamenti proposizionali – sono strutture del tipo “S verbo-di-AP che *p*”, caratterizzate da: (i) il soggetto *S*; (ii) il contenuto intenzionale *p*⁷; (iii) il verbo-di-AP, cioè l'atteggiamento o l'attitudine (*attitude*) che il soggetto ha nei confronti di *p* (lo crede vero, lo desidera, lo spera ecc.). Tra le caratteristiche più importanti degli atteggiamenti proposizionali vi è quella di essere valutabili semanticamente: una credenza può essere vera o falsa, un desiderio soddisfatto o meno e così via. Per esempio, uno stato intenzionale come credere che ci sia un cane sul divano è caratterizzato dall'essere una credenza (*tipo* di atteggiamento) che verte sul fatto che c'è un cane sul divano (*contenuto* dell'atteggiamento), ed è vera se e solo se c'è effettivamente un cane sul divano. La nozione di *intenzionalità* esprime dunque la relazione sussistente tra stato e contenuto intenzionale.

5. Alcuni stati mentali, come per esempio le sensazioni, non sembrano invece riferirsi ad alcunché, né avere un contenuto intenzionale nel senso in cui possono averlo le credenze o i desideri.

6. Il termine “intenzionalità” fu originariamente introdotto da Tommaso ed è poi stato riportato al centro dell'interesse filosofico da Franz Brentano nella seconda metà del XIX secolo. Nella filosofia analitica contemporanea il dibattito sull'intenzionalità è strettamente connesso alle questioni circa la natura della mente e degli stati mentali. Per maggiori approfondimenti cfr., per esempio, Gozzano (1997), Calabi e Voltolini (2009).

7. Per evitare possibili fraintendimenti, parleremo solo di “contenuto intenzionale” o semplicemente di “contenuto” per indicare ciò che viene anche comunemente identificato con una proposizione o, seguendo Frege (1918-19), un “pensiero”.

È opportuno notare, tuttavia, come l'intenzionalità non sia una prerogativa esclusiva degli stati mentali, poiché anche le parole e le espressioni di un linguaggio hanno la proprietà di *riferirsi a* determinate entità⁸. Nel caso del linguaggio naturale il riferimento è la relazione che connette le varie espressioni linguistiche con ciò di cui parlano i soggetti che usano tali espressioni (cioè con le entità che ne costituiscono appunto il referente). Per esempio, quando un parlante asserisce che “Lou Reed è stato il cantante dei Velvet Underground”, utilizza il nome proprio “Lou Reed” per riferirsi a un particolare individuo che ha la proprietà di essere stato il cantante dei Velvet Underground. O ancora, se si asserisce che “l'acqua è dissetante”, il termine “acqua” si riferisce alla sostanza inodore, incolore e insapore che ha come formula chimica H₂O, cioè a un particolare genere naturale. (Sulle difficoltà riguardanti la naturalizzazione del riferimento cfr. Frixione, *infra*).

3) Almeno alcuni stati mentali – detti stati fenomenici – hanno un certo aspetto qualitativo ed esperienziale. Per esempio, quando ascolto una canzone di Patti Smith, ammiro un quadro di Edvard Munch, sento un forte dolore alla gamba, ho fame, sono dispiaciuta per non essere riuscita a incontrare un amico oppure mi annoio aspettando l'autobus, mi trovo in uno stato mentale che ha un certo carattere fenomenico, un aspetto qualitativo, esperienziale, che è accessibile solo a me, in prima persona, in quanto soggetto cosciente di quello stato. Gli aspetti qualitativi, intrinsecamente non intenzionali e accessibili introspektivamente, in modo diretto e immediato, che sono responsabili del carattere fenomenico degli stati mentali, sono detti *qualia* (singolare *quale*). Per quanto detto, i *qualia* sembrano essere entità non fisiche, ineffabili, private e incomunicabili. La discussione sui *qualia* è estremamente ampia, soprattutto perché essi giocano un ruolo centrale nell'analisi della *coscienza fenomenica*, cioè la condizione per cui avere un particolare stato mentale o essere uno specifi-

8. A questo proposito, il dibattito è aperto tra chi sostiene che l'intenzionalità delle parole e delle espressioni linguistiche sarebbe solo “derivata” e “dipendente” dall'intenzionalità “originaria”, che è appunto quella mentale (cfr. Haugeland 1981; Searle 1980, 1983, 1992; Fodor 1987), e chi, al contrario, nega che vi sia una distinzione tra intenzionalità “originaria” e “derivata” (cfr. Dennett 1971, 1978, 1987).

co organismo *fa un certo effetto* al soggetto (cfr. Nagel 1974). Secondo molti è questo il vero ostacolo da superare per chi intenda offrire una teoria naturalistica della mente che sia davvero completa (a questo proposito, Chalmers, 1995, parla infatti di “problema difficile” – *hard problem of consciousness*).

2. Irriducibilità ed efficacia causale

Per offrire un resoconto della mente naturalisticamente adeguato, si tratta innanzi tutto di capire come faccia un sistema fisico, per quanto complesso e sofisticato come il nostro cervello, a dar luogo a qualcosa come la mente umana che, da una parte, sembra estranea e irriducibile al mondo fisico, mentre dall'altra sembra poter avere effetti fisici.

Tra i tentativi di offrire una teoria naturalistica della mente, il più radicale è quello offerto dall'*eliminativismo* (o materialismo eliminativo) secondo cui la mente non è altro che un costrutto teorico postulato da una teoria falsa (Feyerabend 1963; Rorty 1965; Churchland P.M. 1984; Churchland P.M. e Churchland P.S. 1999). Il mentale, pertanto, deve essere rimosso dall'indagine scientifica, che si potrà così concentrare su ciò che è davvero rilevante: il cervello e il sistema nervoso. Sebbene questa teoria risolva, o meglio, dissolva il problema dell'efficacia causale degli stati mentali, essa tuttavia rinuncia a spiegare l'immagine della mente di senso comune. Inoltre, rischia di gettare via, assieme agli stati mentali, anche le nozioni di razionalità e verità, minando così le basi della stessa scienza.

Un altro tentativo di ricondurre la mente all'interno dell'ordine naturale è offerto dal *comportamentismo*, secondo cui – in generale – la mente è assimilabile a una serie di disposizioni al comportamento (Ryle 1949), intendendo con “comportamento” qualsiasi cosa di *pubblicamente osservabile* una creatura faccia. La mente viene “dissolta” e di fatto esclusa da un'ontologia scientifica, ma – secondo

molti critici – ciò avviene nel nome di una semplificazione troppo radicale della relazione tra stati mentali e comportamenti manifesti⁹.

Per il *riduzionismo* (inteso come teoria dell'identità dei tipi – *type identity theory*¹⁰), invece, la mente non deve essere eliminata, ma ridotta – mediante un insieme di leggi ponte di tipo psicofisico – al cervello. La mente è cioè identica al cervello, nel senso che tutti gli stati mentali di un certo tipo (per esempio il dolore) sono identici a stati cerebrali di un altro tipo (per esempio alla stimolazione delle fibre-C). Il che si può anche esprimere dicendo che le proprietà mentali sono identiche a proprietà fisiche (cfr. per esempio Feigl 1958; Place 1956; Smart 1959).

Una soluzione ancora diversa è offerta dall'*emergentismo*, secondo cui esiste una sola realtà, quella fisica, ma quando i sistemi materiali diventano sufficientemente complessi, essi sviluppano delle proprietà mentali emergenti che non sono riducibili al fisico. Tali proprietà, poiché prive di efficacia causale propria, sono tuttavia considerate dei meri epifenomeni (cfr. Kim 1999). L'emergentismo ha il pregio di coniugare monismo e non-riduzionismo, ma solo al prezzo di negare l'efficacia causale degli stati mentali¹¹.

Tutte queste posizioni sono state variamente contestate. A prescindere dalle singole obiezioni, è comunque possibile rilevare come ognuna di esse, di fatto, rinunci a conciliare monismo, non-riduzionismo ed efficacia causale del mentale. Una teoria che invece si propone apertamente di farlo è il *funzionalismo* (e per questo non è un caso che il funzionalismo sia stato adottato dalla scienza cognitiva

9. In maniera più precisa, ma pur sempre schematica, secondo il comportamentismo *logico* ogni espressione che descrive uno stato mentale è traducibile o definibile in un'espressione che descrive un fenomeno comportamentale (vedi per esempio Hempel 1949), mentre secondo il comportamentismo *metodologico* una teoria psicologica genuinamente scientifica dovrebbe tenere conto solo dei dati comportamentali e non invocare mai stati interni (come rappresentazioni, stati o processi mentali) per fare previsioni o dare spiegazioni del comportamento di una creatura (vedi per esempio Skinner 1953). Vedi anche Skinner (1957, 1974).

10. Essa non va confusa con la teoria dell'identità delle occorrenze (*token identity theory*), secondo cui è la singola occorrenza di uno stato mentale a essere identica con una singola occorrenza di uno stato fisico.

11. Per avere una panoramica dell'attuale dibattito sull'emergentismo vedi, per esempio, Clayton e Davies (2006) e Kistler (2006).

per diventare poi uno dei suoi fondamenti teorici – cfr. Fodor 1968; Putnam 1963, 1967). La tesi centrale del funzionalismo è che uno stato mentale è ciò che è in virtù della funzione, o del ruolo causale che esso svolge nell'attività cognitiva di un soggetto, a prescindere da come tale ruolo sia realizzato fisicamente (*realizzabilità multipla*). In altri termini, uno stato mentale è quello che è in virtù dello schema astratto di relazioni causali che esso intrattiene con altri stati mentali, input sensoriali e output comportamentali¹². Tale schema di relazioni causali può poi essere spiegato in termini di computazioni aventi carattere astratto (funzionalismo *computazionale*): i processi mentali sono elaborazioni di informazioni o, più semplicemente, computazioni, cioè sequenze ordinate di operazioni su basi di dati governate da insiemi definiti di regole formali¹³. Ciò significa che le computazioni che costituiscono i nostri processi mentali non sono vincolate a quello che, di fatto, è il loro sostrato materiale, cioè al cervello umano. Secondo la teoria computazionale e rappresentazionale della mente (TCRM) delineata da Fodor (1975, 1987, 2000, 2008), inoltre, le computazioni avvengono su *rappresentazioni mentali*, cioè su stati simbolici che non sono solo dotati di proprietà sintattiche, ma veicolano anche informazioni sul mondo e sono pertanto valutabili semanticamente¹⁴.

La relazione mente-corpo è spiegata dalla TCRM nei termini di una *token identity theory* e della nozione di *sopravvenienza*¹⁵. Secon-

12. In generale il funzionalismo è una posizione “ontologicamente liberale” poiché – non prendendo direttamente posizione sulle proprietà intrinseche degli stati mentali, ma solo su quelle relazionali – è compatibile tanto con il monismo (sia riduzionistico, sia non-riduzionistico), quanto con il dualismo. Per i nostri scopi, tuttavia, prenderemo in considerazione solo il funzionalismo che tenta di conciliare monismo e non-riduzionismo.

13. La nozione di computazione è stata spiegata in modo preciso da Alan Turing (1937) attraverso la Macchina di Turing e la nozione di algoritmo.

14. Le rappresentazioni mentali hanno molte caratteristiche di un linguaggio. In questo senso, Fodor (1975) introduce il concetto di “linguaggio del pensiero” o “mentalese”. Non è però necessario accettare l'idea fodoriana di un linguaggio del pensiero per poter ammettere le rappresentazioni mentali.

15. La nozione di sopravvenienza è stata introdotta da Davidson (1970) proprio con l'intento di conciliare monismo ontologico e non-riduzionismo. Per una discussione critica del concetto di sopravvenienza, soprattutto in relazione alla possi-

do la prima, ogni singola occorrenza di uno stato mentale è identica a una singola occorrenza di uno stato fisico, mentre si dice che il mentale sopravviene sul fisico se e solo se una differenza al livello mentale implica una differenza al livello fisico, ma non viceversa. Sembra così possibile conciliare monismo, non-riduzionismo ed efficacia causale del mentale: (i) la TCRM è una teoria monista, poiché nega che la mente sia una sostanza autonoma e indipendente rispetto alla materia; (ii) la TCRM è un'ipotesi non-riduzionistica, dal momento che ogni stato mentale – essendo individuato in base al suo ruolo causale-funzionale – ha un'identità indipendente dalla sua particolare realizzazione fisica (una differenza al livello fisico non implica una differenza al livello mentale, come invece avviene nell'ipotesi riduzionistica); (iii) siccome gli stati mentali sono stati funzionali che sono pur sempre realizzati fisicamente, essi possono causare ed essere causati da stati fisici ed essere pertanto causalmente efficaci.

In questa prospettiva gli esseri umani hanno una mente perché hanno un sistema di elaborazione dell'informazione sufficientemente complesso, cioè un meccanismo che immagazzina rappresentazioni simboliche e, conformemente a regole formali, le manipola in virtù delle loro proprietà sintattiche, conservando altresì le proprietà semantiche veicolate da tali rappresentazioni. Il fatto che gli esseri umani abbiano una mente, invece, è del tutto indipendente dal particolare materiale biologico di cui sono composti. Con una nota metafora, possiamo affermare che la mente sta al cervello come il software di un computer sta all'hardware: come uno stesso software può “girare” su diverse piattaforme hardware senza che si perdano le sue proprietà caratterizzanti, anche il particolare sistema rappresentazionale che è la mente umana potrebbe (almeno in linea di principio) essere implementato non solo sul nostro cervello biologico, ma anche su altri supporti materiali (come ad esempio dei microcircuiti di silicio), che a quel punto potrebbero dirsi dotati di mente esattamente come lo siamo noi esseri umani. Secondo la prospettiva funzionalista, dunque, non c'è ragione perché la mente umana non possa essere considerata equivalente a un'ipotetica mente artificiale. HAL 9000, il su-

bilità di legare tale nozione a un materialismo non-riduzionistico, cfr. per esempio Kim (1993, 1998).

percomputer di *2001: Odissea nello spazio*, potrebbe essere più di una mera invenzione fantascientifica.

3. Psicologia del senso comune e intenzionalità

Nel contesto di un progetto di naturalizzazione del mentale, è legittimo domandarsi se vi sia ancora spazio per la cosiddetta “psicologia del senso comune” (*folk psychology*), che tutti noi utilizziamo abitualmente per spiegare e prevedere il comportamento degli altri esseri umani attraverso l’uso di stati intenzionali come le credenze, i desideri, le intenzioni, le speranze ecc. Per esempio, supponiamo che io veda un amico dall’altro lato della strada che guarda nella mia direzione, alza il braccio e lo agita. Se voglio offrire una spiegazione di questo suo comportamento, posso farlo facilmente attribuendogli il *desiderio* di salutarmi e la *credenza* di poter ottenere questo scopo alzando e agitando il braccio. Questo è un tipico esempio di spiegazione intenzionale offerta dalla psicologia di senso comune. Si tratta di una spiegazione mentalistica che per interpretare, spiegare e prevedere il comportamento fa uso di stati mentali intenzionali, come le credenze e i desideri, dotati di uno specifico contenuto avente carattere rappresentazionale, contenuto che gioca un ruolo esplicativo determinante (l’azione di alzare e agitare il braccio, infatti, non sarebbe stata spiegata in modo appropriato se avessi attribuito al mio amico il desiderio di ignorarmi).

La questione, ora, è se le entità teoriche postulate dalla psicologia di senso comune (atteggiamenti proposizionali e contenuti intenzionali) possano trovare posto o meno nell’ontologia di una teoria scientifica della mente e se l’intenzionalità sia quindi una proprietà genuina del mentale. Secondo il realismo intenzionale la risposta è affermativa: atteggiamenti e contenuti proposizionali sono enti genuini e la psicologia del senso comune è una teoria vera. Il realismo intenzionale si differenzia sia dalle posizioni eliminativiste sia da quelle strumentaliste o interpretativiste.

Secondo l’*eliminativismo*, gli atteggiamenti proposizionali hanno lo stesso statuto ontologico del flogisto, sono cioè entità “vuote”, postulate da una teoria della mente ancora rozza che, a una più attenta analisi scientifica, si rivela falsa (Churchland P.M. 1981, 1984, 1993;

Churchland P.S. 1986). Gli atteggiamenti proposizionali, insomma, non sono enti genuini e vanno pertanto eliminati non solo dalla nostra ontologia, ma anche da ogni nostra spiegazione della mente. Ne segue, ovviamente, che la psicologia di senso comune debba essere considerata una teoria falsa, da sostituire con una psicologia coerente con i risultati delle neuroscienze.

Anche per lo *strumentalismo* (o *interpretativismo*) credenze, desideri, intenzioni ecc. non sono enti genuini e sfuggono pertanto a un resoconto autenticamente scientifico. Ciò nonostante, la psicologia di senso comune è ancora considerata una teoria utile per fare previsioni e spiegare il comportamento di un agente, perché ci dà la possibilità di interpretare tale comportamento in termini di ragioni, adottando quello che Dennett chiama “atteggiamento intenzionale” (Dennett 1987). Gli atteggiamenti proposizionali emergono allora come il risultato di un’attività interpretativa, sono utili costrutti teorici che non fanno però parte dell’ontologia di una teoria scientifica della mente.

Siccome negano dignità ontologica a credenze, desideri, intenzioni e a tutti gli altri atteggiamenti proposizionali, eliminativisti e strumentalisti non hanno il problema di conciliare tali entità con una teoria naturalistica della mente (cfr. Bermúdez 2005). Questa è invece una questione centrale per il realismo intenzionale: com’è possibile che stati fisici del nostro cervello abbiano contenuto intenzionale, che vertano su qualcosa e lo rappresentino? Qual è l’origine di tali contenuti e come sono individuati? E com’è possibile che tale contenuto sia causalmente efficace?

Cominciamo a considerare quest’ultima domanda. A prescindere da come venga determinato il contenuto, è comunque in virtù di tale contenuto che uno stato mentale causa un altro stato mentale o un comportamento. Ma com’è possibile che un’entità astratta come il contenuto possa essere causalmente efficace? Si tratta insomma di capire come un meccanismo fisico come il nostro cervello possa realizzare catene causali di stati intenzionali che tengano conto delle proprietà semantiche di tali stati, senza dover postulare qualcosa come un “interprete interno” che abbia la capacità cognitiva di comprendere il significato degli stati intenzionali (cfr. Marruffa, *infra*).

Per vedere se e come il realismo intenzionale possa superare questa *impasse* prendiamo nuovamente in considerazione la teoria computazionale e rappresentazionale della mente avanzata da Fodor

(TCRM). Innanzi tutto, stando alla TCRM, avere un atteggiamento proposizionale significa intrattenere una certa relazione con una rappresentazione mentale, un simbolo del linguaggio del pensiero avente proprietà sintattiche e semantiche; gli atteggiamenti proposizionali ereditano così le loro proprietà sintattiche e semantiche da quelle delle rappresentazioni mentali. Uno stato mentale ne causa un altro in virtù delle proprietà sintattiche delle rappresentazioni mentali che – ricordiamo – sono causalmente efficaci in quanto realizzate fisicamente (*token identity theory*). Se le proprietà semantiche delle rappresentazioni mentali fossero riflesse dalla loro struttura sintattica, allora si potrebbe chiarire come un atteggiamento proposizionale possa avere efficacia causale in virtù del proprio contenuto. Abbiamo già visto che le regole che governano le computazioni e rendono possibili le transizioni di stato di un sistema cognitivo sono regole formali, sensibili solo alle proprietà sintattiche delle rappresentazioni mentali, ma indifferenti a quelle semantiche. Per ogni sistema di simboli formalizzabile, inoltre, è possibile definire delle regole formali che, pur essendo sensibili alle sole proprietà sintattiche dei simboli, sono semanticamente appropriate, cioè consentono di fare tutte e sole le inferenze valide sotto il profilo semantico, mettendo così al riparo da possibili inferenze arbitrarie. In tal modo la TCRM riesce a spiegare come un sistema fisico come il nostro cervello possa fare inferenze che siano efficaci causalmente e rispettino la semantica degli stati intenzionali. Possiamo sintetizzare questo risultato con il “Motto Formalista” di Haugeland: «Se ti prendi cura della sintassi di un sistema rappresentazionale, la semantica si prenderà cura di se stessa» (1985, p. 106)¹⁶.

Per la TCRM uno stato mentale deriva le sue proprietà semantiche dalla rappresentazione mentale con cui è in relazione; ma che cosa conferisce contenuto alle rappresentazioni mentali? Il dibattito a proposito di una possibile naturalizzazione dell'intenzionalità è oggi as-

16. In realtà, sebbene il funzionalismo della TCRM e il modello del calcolatore sembrino fornire un buon resoconto dei processi cognitivi umani, c'è il problema che tali processi, per essere trattabili computazionalmente, devono essere locali (cioè avere un accesso limitato alla conoscenza di sfondo) e pertanto “incapsulati”. Ma per Fodor solo alcuni processi periferici sono tali. Per una discussione approfondita del problema cfr. Marraffa (2003).

sai vivo: da una parte ci sono coloro che negano la possibilità di naturalizzare gli stati intenzionali, sottolineandone il carattere intrinsecamente normativo (cfr. Agazzi, *infra*)¹⁷; dall'altra vi sono invece coloro che sostengono diverse teorie naturalistiche a proposito della determinazione del contenuto delle rappresentazioni mentali. Tali teorie sono raggruppabili in due tipi principali: le teorie *funzionali* e le teorie *causal-informazionali*.

Secondo le teorie *funzionali* (Block 1986; Harman 1973), il contenuto di una rappresentazione mentale è determinato dalle sue relazioni (causali, funzionali, computazionali, inferenziali ecc.) con altre rappresentazioni mentali. Ci possono poi essere differenze a seconda che si tengano in considerazione tutte le relazioni con altre rappresentazioni mentali (olismo) o solo una parte (molecolarismo), oppure che si considerino irrilevanti i fattori esterni o meno.

Secondo le teorie causal-informazionali (Dretske 1981, 1988, 1995; Fodor 1987, 1990; Devitt 1996), il contenuto di una rappresentazione mentale è determinato da ciò che causa o causerebbe tale rappresentazione. Nel tentativo di risolvere il problema di spiegare perché, talvolta, una rappresentazione possa essere causata da qualcosa che non rappresenta e possa rappresentare qualcosa da cui non è stata causata sono state sviluppate la teoria della *dipendenza asimmetrica* (Fodor 1987, 1990, 1994) e quella *teleo-causale* (Dretske 1988, 1995; Millikan 1984, 1993, 2000). Secondo la prima, relazioni causali spurie sono possibili solo perché è stato già instaurato il nesso causale appropriato. Per la teoria teleo-causale, invece, il contenuto di una rappresentazione mentale è determinato dalla funzione biologico-evolutiva che assolve, per cui le relazioni causali spurie possono essere individuate e scartate facendo appello proprio a tale funzione.

Infine, per quanto riguarda l'individuazione del contenuto, si possono definire due posizioni principali: l'*internismo* e l'*esternismo*. Secondo l'*internismo*, il contenuto di uno stato mentale dipende esclusivamente da fattori interni all'individuo, nella fattispecie dagli stati cerebrali dell'individuo. L'*esternismo* sostiene invece che il contenuto dipende, almeno in parte, da fattori esterni all'individuo e appartenenti all'ambiente naturale e/o sociale in cui egli si trova. In

17. Cfr. anche Davidson (1980) e Hornsby (1997).