

Si può parlare di coscienza artificiale?

Il film A.I. di Steven Spielberg: un'occasione per parlare di coscienza artificiale

Perché il film di Steven Spielberg è stato denominato A.I., ossia intelligenza artificiale? Sicuramente il regista americano e gli esperti di marketing hanno attribuito al nome A.I. un valore evocativo che ha radici più nella scienza che non nella fantascienza. Avrebbero potuto chiamarlo CyberPinocchio o Cyberkid. Avrebbero potuto utilizzare uno dei tanti termini che la fantascienza ha messo a disposizione della letteratura: androide, robot e replicante. Invece hanno preferito appoggiarsi a un termine che ha avuto così tanto successo negli ultimi decenni: un termine che evoca la possibilità di costruire macchine (o sistemi) intelligenti.

Sarebbe bastato che Spielberg avesse chiamato il film A.C., ossia *Artificial Consciousness*, e il film stesso avrebbe potuto avere un grosso impatto sulla storia delle idee. Invece di sancire la consacrazione della fine di una fase culturale, avrebbe potuto aprirne un'altra. La storia della letteratura fantastica e della fantascienza si basa sul fatto che gli esseri artificiali abbiano coscienza, non che siano intelligenti, cioè che siano soggetti. A cinquant'anni di distanza dalla pubblicazione di *'Io, robot'* di Isaac Asimov, sarebbe stata un'ottima occasione di parlare di coscienza artificiale.

Nella letteratura del fantastico gli esseri artificiali sono sempre stati dotati di coscienza; i loro autori non si sono preoccupati dell'intelligenza delle loro creature ma della loro capacità di essere soggetti autonomi di decisioni e obiettivi. *Frankenstein* di Mary Shelley, i robot di Karel Capek (1920) e quelli di Isaac Asimov sono tutti dotati di corpi e di coscienza. Solo Stanley Kubrick riuscirà a proporre un essere artificiale, privo di corpo ma dotato di coscienza ed emozioni similumane: HAL, il computer intelligente del film *2001 Odissea nello spazio*. E forse proprio il film *2001 Odissea nello spazio* avrebbe dovuto essere denominato A.I., poiché in esso HAL personifica il vero grande interprete delle speranze riposte nella disciplina "intelligenza artificiale" che solo un grande genio come Kubrick poteva pensare di proporre, nelle sale cinematografiche, privo di corpo.

Uno dei primi personaggi artificiali dotati di coscienza presenti nella letteratura e non formati da materiale biologico è stato Pinocchio. Il personaggio di Collodi propone una struttura di legno che in un primo

momento percepisce il dolore e solo successivamente, dopo aver interagito con l'ambiente, acquista lo statuto di persona. Il suo esserci, la sua coscienza fenomenica è dunque precedente al suo sviluppo cognitivo. In altri casi, al contrario, la creatura artificiale nasce già dotata di una potente capacità elaborativa, ma raggiunge lo status di persona, reale e riconosciuta, solo dopo lunghi anni di interazione con soggetti umani. E' il caso de *l'uomo bicentenario* di Asimov e dell'androide Data in *Star Trek*. Al contrario, Frankenstein e i robot di Karel Capek sono costituiti da "protoplasma vivente"; ma ciò che conta in loro è che la loro esistenza è l'effetto della volontà prometeica dei loro creatori umani di produrre un essere a loro immagine e somiglianza. Non è importante il materiale di cui sono fatti: è un costruttore, umano e quindi fattibile, a essere la loro causa formale ed efficiente, e tanto basta.

Dall'intelligenza artificiale alla coscienza artificiale

Fin dall'anno della sua creazione, il 1956, il termine intelligenza artificiale sottolineò l'aspetto costruttivista della relativa disciplina. A volte i cultori dell'intelligenza artificiale si sono posti l'obiettivo conoscitivo di capire il funzionamento di menti e cervelli biologici, ma solo al fine di avere suggestioni e suggerimenti. Quando non l'hanno fatto sono rientrati nell'alveo di altre discipline quali la psicologia, la filosofia, la biologia, le scienze cognitive. La parola artificiale non è un semplice aggettivo, bensì implica un programma di lavoro di tipo ingegneristico nel settore della progettazione e costruzione di esseri artificiali dotati di intelligenza.

Negli anni Settanta c'è stato il tentativo di far percorrere due diverse strade alla robotica industriale e all'intelligenza artificiale. La robotica avrebbe dovuto costruire i corpi; l'intelligenza artificiale avrebbe dovuto interessarsi di menti. Poi, una volta sviluppati corpi e menti artificiali, si sarebbero potuti combinare per dare luogo a un essere similumano. L'obiettivo iniziale della cibernetica, la creazione di un soggetto artificiale a immagine e somiglianza del nostro essere soggetti coscienti, è stato così indefinitamente posposto nel tempo, e subordinato allo sviluppo propedeutico di altri moduli e sottoinsiemi (sia meccanici che logico-cognitivi) che si ritenevano necessarie fattibili.

A differenza del termine *consciousness*, il termine intelligenza ha, di fatto, un'accezione utile ed economica. A differenza di quello che succede per caratteristiche dell'essere umano quali sentimenti, emozioni e, in genere, attività legate a un pensiero vago e poco produttivo (che solo gli esseri umani possono permettersi), l'intelligenza artificiale ha

sempre avuto successo per gli aspetti applicativi: nell'area informatica, nell'analisi del linguaggio, della scienza cognitiva e dei sistemi esperti. Molto si è scritto negli ultimi anni sull'argomento della coscienza; alcuni ritengono che addirittura si sia scritto, o meglio riscritto, troppo. In realtà il problema è stato affrontato con i classici strumenti della filosofia e della psicologia. Solo ora la biologia e la medicina cominciano a non tacciare di visionarietà coloro che utilizzano nuove e più appropriate tecnologie per studiare cervelli e menti. Per la prima volta una rivista come *Nature Neuroscience* ha scritto nell'editoriale che “sebbene dieci anni fa pochi ricercatori avrebbero preso sul serio il campo della coscienza, oggi i tempi stanno cambiando. Sebbene lo studio della coscienza resta un campo elusivo, sta nascendo un nuovo consenso tra i neuroscienziati sulle possibilità di affrontare sperimentalmente il problema della natura della coscienza” (Editor, 2000). Neuroscienziati come Damasio ritengono che i prossimi dieci anni saranno decisivi per lo sviluppo di una nuova disciplina che definisca gli strumenti metodologici necessari a comprendere la coscienza (American, 1999; Damasio, 1999). Tuttavia è indubbio che affrontare il tema della coscienza da un punto di vista ingegneristico non sembra, finora, avere suscitato l'interesse della comunità scientifica internazionale. Anzi, al di fuori della fantascienza, è incerta la collocazione temporale della nascita di una *artificial consciousness*, intesa come disciplina a se stante: gli studiosi non si sono ancora riuniti in occasione di un seminario estivo (sulla falsariga di quello storico di Dartmouth che segnò la nascita ufficiale della *artificial intelligence*) per parlare di coscienza artificiale o ingegneria della coscienza. Eppure, dopo parecchi anni dal primo convegno di Tucson tenutosi nell'Aprile del 1994, si può ragionevolmente ritenere che il tema della coscienza sia stato riletto attualizzandolo alla luce dei progressi della neurobiologia e delle scienze cognitive. In quella conferenza poi ripetutasi ogni due anni, si ebbe il primo confronto internazionale sul tema della coscienza, non solo tra filosofi, ma anche con neuroscienziati, fisici, scienziati, studiosi di scienze cognitive, psicologi, studiosi di farmacologia, neuropsicologia. Da allora – grazie anche alla contemporanea simbolica apertura di una serie di studiosi di rilievo (Francis Crick, Harry Damasio, Gerard Edelman, David Chalmers) la coscienza è diventata un problema scientifico e non solo un problema filosofico o terminologico. L'importanza di questa, e altre conferenze centrate sul tema della coscienza, non è stato tanto il fatto di aver avanzato con successo nuove definizioni della coscienza, quanto di aver riaperto un problema che sembrava – forse a torto – essere stato risolto. Nuovi criteri per la valutazione della coscienza e delle teorie che

dovrebbero descriverla sono stati avanzati e non possono più essere facilmente ignorati.

Nascita della disciplina denominata “coscienza artificiale”

Nel 1982, in Giappone, si diede inizio al programma di ricerca FGCS (fifth-generation computer systems) che aveva come obiettivo quello di sviluppare le applicazioni di intelligenza artificiale. Tuttavia, già nel Convegno IJCAI'87 tenutosi a Milano, ci si interrogava, a trent'anni della nascita della disciplina, sul futuro dell'intelligenza artificiale. Nello stesso anno il ritorno prepotente delle reti neurali apriva nuove prospettive alternative al classico approccio dell'intelligenza artificiale. La rivista *Sistemi intelligenti*, che nasce nell'aprile 1989, coglieva i fermenti di quel particolare momento storico e si proponeva come una rivista di scienza cognitiva e intelligenza artificiale, che aveva come obiettivo “lo studio dell'intelligenza, delle attività mentali e del loro supporto fisico, il cervello”. Ma l'intelligenza ha bisogno di una volontà che la guidi. La razionalità deve essere mossa da fini e motivazioni che non possono essere il prodotto delle elaborazioni interne di un sistema chiuso. Le attività mentali non possono essere esclusivamente attività cognitive nel senso di elaborazione razionale e meccanica dei simboli. L'esperienza empirica ci mostra che, anzi, questo tipo di intelligenza è semmai un prodotto ontogeneticamente e filogeneticamente piuttosto tardo. E così sono diventati ineludibili gli aspetti legati alla capacità del cervello di determinare un soggetto unitario, capace di fare esperienza di se stesso e del mondo circostante, capace di genuine e primarie capacità semantiche intenzionali nei confronti di eventi esterni, passati e futuri; in grado di produrre motivazioni e fini, valori soggettivi e sensazioni. In poche parole, la coscienza.

Ma coscienza vuol dire intenzionalità come è stato indicato da numerosi autori, pur con accezioni diversissime e talora anche contrastanti (Agazzi, 1981; Brentano, 1874; Dennett, 1969, 1987; Fodor, 1987; Searle, 1983, 1992). E allora in primis il problema della realizzabilità della coscienza artificiale ha avuto all'inizio la forma della costruzione di un agente intenzionale. In questo senso si devono leggere le prime realizzazioni di Gerard Edelman (Edelman, 1987; Edelman & Mountcastle, 1978), poi sfociate in modo diretto ed inequivoco nello studio della coscienza *tout court* (Edelman & Tononi, 2000). Tuttavia i robot intenzionali di Edelman, la serie Nomad e la serie Darwin, si scontravano contro il modello oggettivista della realtà sostenuto da generazioni di studiosi. Senza una teoria della coscienza e quindi dell'intenzionalità accettata da tutti, era impossibile riconoscere in una

struttura artificiale la presenza di genuine capacità intenzionali. E infatti la domanda sui correlati neurali della coscienza resta aperta sia nel campo biologico sia nel campo, finora ipotetico, dell'artificialità.

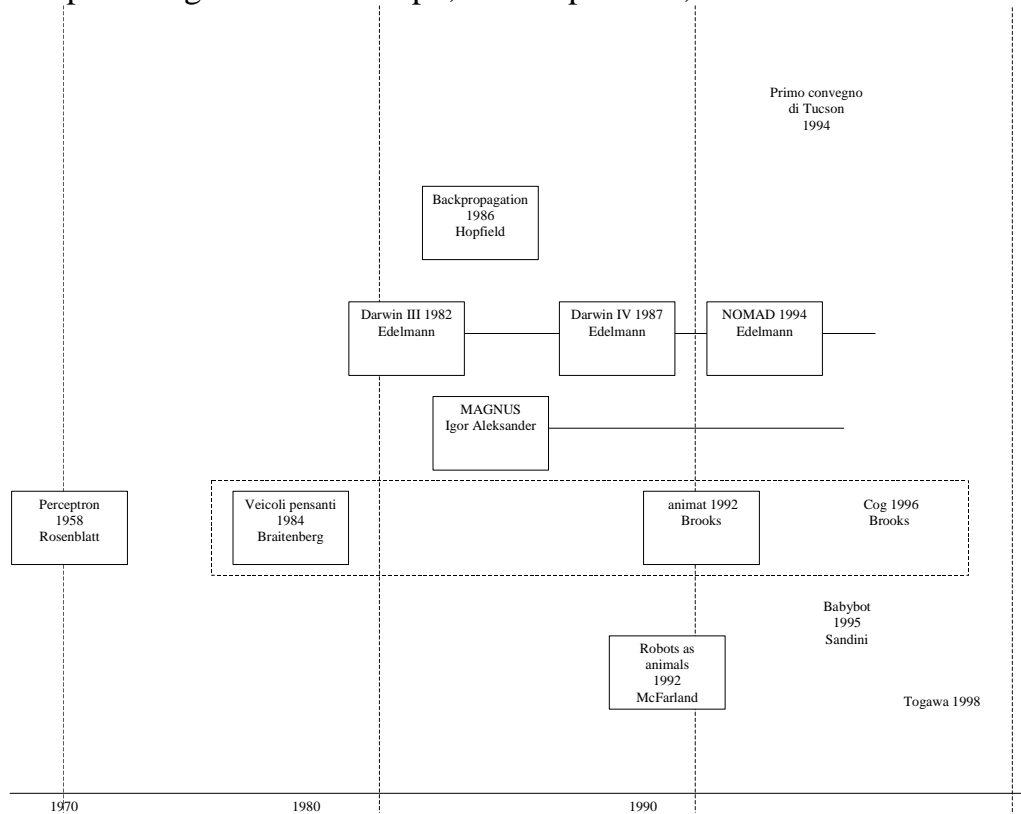


Fig.1 Sebbene non tutti gli studiosi citati appartenessero esplicitamente al filone della coscienza artificiale è tuttavia possibile, a posteriori, riconoscere il loro contributo.

Alla fine degli anni Ottanta e all'inizio degli anni Novanta il problema della mente si allarga allo studio del comportamento intelligente negli animali e nei robot. Nascono in questo contesto gli "animat" e le problematiche degli "automi spirituali", in cui convergono non solo la robotica e l'intelligenza artificiale, ma anche la scienza cognitiva, le scienze biologiche, la filosofia e la linguistica. La robotica si apre alla costruzione di "macchine inutili", quali gli esseri robotici di Rodney Brooks (Brooks, 1990, 1991; Brooks, Breazeal, Marjanovic, Scassellati, & Williamson, 1999), le quali cominciano a proporre un approccio *bottom up* basato sulle tematiche dell'apprendimento, dello sviluppo e del ruolo fondamentale del corpo, che i primi tentativi dell'intelligenza artificiale sembravano avere accantonato. E tuttavia la riscoperta della necessità dell'ambiente e dei corpi, che non sono altro che il primo elemento ambientale sottoposto al controllo dell'agente artificiale, non ha permesso di risolvere immediatamente il quesito dell'intenzionalità. Anche sviluppi ulteriori di questo tipo di approccio, sebbene abbiano

fornito informazioni circa le modalità dello sviluppo sensori motorio, ben poco hanno aggiunto sulla natura dell'essere soggettivo. Cog è sì un umanoide di forma antropomorfa, ma non condivide quasi nulla della struttura mentale e cognitiva propria degli esseri umani. Da un punto di vista strutturale e cognitivo non è altro che un insieme di riflessi elementari sapientemente orchestrati dai suoi progettisti e costruttori. L'uso delle reti neurali quali strumenti di apprendimento si limita a una, pur sofisticata, capacità di approssimazione di funzioni. Le capacità di apprendimento di Cog sono limitate ad ambiti specifici: non sono in grado di determinare uno sviluppo integrato del sistema. E, a riprova dei limiti del modello proposto, sembra che l'interesse del laboratorio del MIT per Cog stia rapidamente scemando.

Diverso e più complesso è l'approccio seguito da Giulio Sandini al LIRA-Lab che, attraverso un robot antropomorfo di piccole dimensioni (Babybot), cerca di determinare i vincoli e le potenzialità dello sviluppo integrato di un sistema (Sandini, Manzotti, Metta, & Panerai, 2000). L'idea è di individuare un percorso di fasi distinte di crescita (mutuamente sovrapposte) tali da permettere la genesi di un organismo cognitivamente complesso. Un obiettivo che potrebbe essere irraggiungibile in un solo passo e che potrebbe richiedere la maturazione separata e graduale di tutta una serie di potenzialità separate.

Solo pochi hanno indirizzato esplicitamente la ricerca delle modalità di costruzione di un soggetto artificiale genuinamente intenzionale e pertanto anche cosciente. Fra questi va sicuramente annoverato il pluriennale tentativo di Igor Aleksander, dell'Imperial College di Londra, di definire e costruire un agente artificiale cosciente, nella convinzione che uno sviluppo sufficientemente fedele delle capacità cognitive di un essere artificiale non possa che corrispondere alla nascita di un essere artificiale (Aleksander, 1994, 1996). Del resto anche Searle, storico sostenitore della natura biologica della coscienza, ha recentemente ammesso che non vi è alcun motivo a priori per negare che una creatura artificiale, capace di riprodurre i meccanismi fondamentali del cervello umano, possa corrispondere a un soggetto cosciente (Aleksander, 2000).

Quanto poi la comprensione dei meccanismi biologici sia effettivamente necessaria alla realizzazione dei meccanismi che sottostanno alla emergenza della coscienza non è affatto chiaro. Da un lato l'intelligenza artificiale classica ha sempre rifiutato di imitare espressamente i processi biologici. D'altro lato, proprio il fallimento dell'IA classica sembra spingere connessionisti e neuroscienziati a cercare nella riproduzione esatta dei meccanismi del cervello l'unica via possibile per replicare un soggetto.

E tuttavia sono molti coloro che oggi ritengono necessaria una rivoluzione delle categorie fondamentali per poter affrontare con un nuovo metodo l'integrazione della dimensione soggettiva con quella oggettiva senza che la prima debba necessariamente ridursi alla seconda (Chalmers, 1996; Di Francesco, 1996; Praetorius, 1999; Stubenberg, 1998). Si sta lentamente erodendo quell'atteggiamento di voluta sospensione dell'indagine ultima della realtà che la svolta linguistica aveva, in un certo senso, assiomatizzato (Dummett, 1998; Picardi, 1992; Ryle, 1949). Anche se con fatica, nuovi metodi sono proposti, nuove ipotesi sono messe alla prova per determinare la natura e le modalità dell'intenzionalità che sta alla base dell'esperienza cosciente (Edelman & Tononi, 2000; Manzotti, 2001; Scassellati, 2000; Togawa & Otsuka, 2000; Varela, 2000; Varela & Shear, 1999; Zlatev, 2001).

Il problema della coscienza nell'ambito della disciplina “coscienza artificiale”

Per costruire artefatti coscienti, è necessario rivisitare l'immane lavoro preparatorio che migliaia di anni di storia hanno depositato nella cultura filosofica e psicologica.

Un prerequisito per lo studio della mente, al fine della sua costruzione in strutture artificiali, è l'identificazione fra la mente e la mente cosciente. La precisazione può sembrare superflua, ma ciò dipende dal fatto che negli ultimi cinquant'anni, ossia dopo Wittgenstein, Skinner e Ryle (ognuno per la sua parte), si è cercato di spiegare la mente senza far uso della coscienza: un esito estremo delle contraddizioni interne del dualismo seicentesco. Nell'ambito delle tematiche della coscienza artificiale ci si deve interrogare su cosa s'intende per coscienza? Non la coscienza morale o etica, bensì quella che gli anglosassoni definiscono *consciousness* e che corrisponde alla capacità di un soggetto umano di fare esperienza dei propri pensieri, di se stesso e del mondo. La coscienza è stata la grande assente della ricerca scientifica nel Novecento: adesso i tempi stanno cambiando e la letteratura sull'argomento sta dilagando, impegnando studiosi e premi Nobel che provengono dalle discipline più disparate (Roger Penrose, Gerard Edelman e Francis Crick). Fino a oggi si sono applicate al problema della coscienza le categorie di Galileo, di successo nello spiegare i fenomeni fisici, ma apparentemente insufficienti nello spiegare i fenomeni mentali. Quando un essere umano fa esperienza del mondo (ossia percepisce colori e sapori, oppure prova dolore e piacere) coglie degli aspetti *qualitativi* della realtà, che la scienza galileiana non è in grado di misurare. Tuttavia le *qualità* esistono: anche se da Cartesio in poi esse sono state relegate nel chiuso di una dimensione spirituale che non può

venire aggredita scientificamente. Finora è stato possibile eludere il problema perché non vi erano le condizioni tecnologiche per poter costruire un essere cosciente artificiale. Finora la progettazione di artefatti (automobili, ponti, satelliti, computer) non ha richiesto la comprensione della coscienza. Oggi che i robot iniziano ad avvicinarsi all'essere umano - sia come capacità di calcolo sia come struttura fisica – il problema non può più essere evitato.

Una possibile soluzione del problema – soluzione che potrebbe trovare conferma nella costruzione di un soggetto artificiale – nasce da un'analisi e revisione delle categorie ontologiche fondamentali. D'altronde l'idea che la realtà sia composta da oggetti e dalle loro proprietà è stata messa in discussione in ambiti tra loro molto diversi (Davidson, 1980; Smith, 1998; Stapp, 1998; Whitehead, 1925, 1927). Sono così state proposte ontologie alternative più aderenti agli ultimi esiti della fisica che vedono, quali componenti elementari della realtà, gli eventi. In quest'ottica un soggetto cosciente corrisponde a un complesso articolato e dinamico di eventi; i cervelli umani sono gli oggetti fisici in grado di permettere a questi eventi di verificarsi. Per esempio un embrione, prima della comparsa del sistema nervoso, non può essere considerato un soggetto cosciente (come d'altra parte una persona anencefalica); dopo un sufficiente lasso di tempo, l'essere umano risultante sarà cosciente. La mente non è più una scatola vuota che riceve suoni e immagini dal mondo esterno, ma è una porzione del mondo esterno che trova in se stessa la propria unità.

E' però necessario che tali teorie siano sviluppate in modo da poter proporre modelli costruttivi. In questo modo è possibile compiere esperimenti che metteranno alla prova tali intuizioni. In quest'ottica si muove il libro *Coscienza e Realtà* (Riccardo Manzotti e Vincenzo Tagliascio, *Coscienza e realtà. Una teoria della coscienza per costruttori e studiosi di menti e cervelli*, Il Mulino, 600 pagine, 70.000 lire (Manzotti & Tagliascio, 2001)) che vorrebbe presentarsi come una proposta teorica per la costruzione di un robot dotato di coscienza. In realtà non si può fare a meno di rivisitare, attualizzandole, quelle teorie filosofiche che dovrebbero suggerire agli ingegneri i processi metodologici e tecnici per realizzare uno dei più affascinanti obiettivi proposti dalla scienza nel secolo XXI: capire chi siamo attraverso la costruzione di una macchina che potrà comunicarci quello che avviene dentro di lei allorché, analogamente a quello che avviene in un essere umano, proverà una sensazione. Sulla base della teoria presentata per costruire una macchina di questo tipo non occorre nessuna ulteriore innovazione tecnologica, né che si deve attendere la scoperta di qualche

misterioso fenomeno fisico o biologico finora completamente sconosciuto, né che si deve fare ricorso a elementi organici, quali proteine o biochip.

D'altra parte la filosofia in più di duemila anni di storia ha caricato termini quali *soggetto*, *intenzionalità*, *rappresentazione* di connotati così diversi e articolati da impedire scorciatoie o superficiali tentativi di semplificazione. Come si è detto, la scienza, allo stadio attuale, non ha la minima idea di come sia possibile che un sistema fisico (un cervello, un sistema nervoso, un insieme di neuroni) possa produrre quell'insieme strano di fenomeni che corrispondono alla nostra esperienza cosciente. Non solo. La scienza non sa neppure cosa siano questi fenomeni. Paradossalmente il problema della coscienza sta diventando sempre più urgente e importante man mano che le nostre possibilità di indagine del cervello progrediscono. Alcuni decenni fa, quando era pressoché impossibile studiare, in vivo, gli unici cervelli di cui si sappia con certezza che producano la coscienza (i cervelli umani), si poteva sempre sperare che il progredire dei mezzi di indagine avrebbe permesso di scoprire all'interno del cervello un qualche fenomeno fisico sconosciuto e straordinario che avrebbe risolto ogni incertezza. Questo non è avvenuto e il cervello, per quanto complesso, è rimasto un oggetto fisico esattamente come gli altri. D'altro lato le teorie filosofiche sulla coscienza, attuali e passate, non forniscono indicazioni chiare su come procedere nella costruzione di un soggetto e sono per questo non-scientifiche nel senso di non-falsificabili empiricamente.

Ma perché è così difficile il problema della coscienza? Un esempio chiarirà la cosa. Immaginiamo di stare guardando una penna rossa. Noi facciamo esperienza della penna e delle sue proprietà, tra le quali quella del rosso. Tuttavia, secondo la scienza, noi siamo il nostro cervello e tutte le volte che facciamo esperienza di qualcosa è perché dentro il nostro cervello succede qualcosa che provoca la nostra esperienza. E tuttavia ... il nostro cervello non diventa rosso quando noi guardiamo la penna rossa. Il nostro cervello rimane un oggetto fisico distinto dalla penna che osserva, non ne riproduce le proprietà. E' straordinaria che per tanto tempo si sia accettata l'idea che l'oggetto cervello potesse 'creare' le nostre sensazioni quasi che una 'sensazione' non fosse nulla di 'reale' o quasi che il cervello potesse compiere un'azione di 'magica creazione'. Per quanto se ne sa, nessun oggetto è in grado di creare nessun altro oggetto, al più di assemblare, trasformare, modificare. E l'oggetto cervello non può essere diverso dagli altri da questo punto di vista. Del resto tutte le discipline scientifiche mature sono rette da qualche forma del principio di conservazione. La scienza biologica conobbe uno dei suoi più grandi progressi quando rifiutò l'idea della possibilità della

generazione spontanea. La scienza fisica quando introdusse i principi di conservazione del moto e della massa. L'elenco potrebbe continuare numeroso. Non si vede perché il significato e l'esperienza dovrebbero seguire criteri diversi. Inoltre quello che il cervello dovrebbe creare non è neppure descrivibile in termini di presenza fisica. Sensazioni, pensieri, immagini, percezioni nel senso di contenuti della coscienza sono entità quantomeno evanescenti e prive di ogni traducibilità fisica oggettiva (ma non empirica). Attendarsi che, quasi per magia, i «segnali nervosi diventino immagini e sensazioni» è ancora più incredibile. L'assunto Galileiano secondo il quale tutte le entità fisiche sono entità oggettive, implica che i segnali, in qualunque sistema fisico (e quindi anche dentro i nostri cervelli) siano e restino fenomeni di tipo fisico.

Ma se gli esseri umani sono coscienti si deve spiegare come sia possibile che un sistema fisico produca l'esperienza cosciente. Pensiamo alla Terra prima della comparsa degli esseri viventi. E' piuttosto ragionevole supporre che, in assenza di organismi viventi, non vi fosse nessuna esperienza cosciente. All'altro estremo, quattro miliardi di anni dopo, abbiamo la terra popolata dagli esseri umani che, come abbiamo detto, sono evidentemente coscienti. A un certo punto l'evoluzione deve avere determinato la comparsa dell'esperienza cosciente: dove e perché?

Del resto, e l'esempio è così quotidiano che si tende a dimenticare, l'esperienza cosciente viene prodotta ogni volta che un nuovo essere umano è concepito e si sviluppa. Pensiamo a un ovulo fecondato. E' un oggetto fisico come gli altri dal punto di vista dell'esperienza cosciente. Siccome è privo di sistema nervoso pare ragionevole che sia privo anche di esperienza cosciente. Dopo qualche mese o dopo alcuni anni avrà un'esperienza cosciente paragonabile alla nostra. Un oggetto fisico privo di esperienza cosciente è diventato un soggetto cosciente che fa esperienza di sé e del mondo. Cosa è successo? Siamo veramente costretti ad accettare un'ontologia, quella metafisica, in cui il fatto quotidiano del nostro essere coscienti è tanto misterioso quanto 'l'apparizione del Genio della lampada di Aladino' come diceva Thomas Husley alla fine del secolo scorso?

Analogamente all'evoluzione e alla crescita biologica degli esseri umani possiamo pensare alla produzione di robot che siano sempre più simili, nel comportamento, agli esseri umani. Supponendo che, fra 10 o 1000 anni, si sia in grado di produrre un essere artificiale in grado di esibire un comportamento (corporeo e linguistico) paragonabile a quello degli esseri umani, dovremmo per questo concludere che, magicamente, abbia prodotto l'esperienza cosciente? Come ha scritto Jaegwon Kim, non solo non si sa come muoversi per creare un soggetto artificiale ma non si saprebbe nemmeno come valutare i successi conseguiti (Kim, 1998).

Oggi la scienza non è in grado di rispondere a nessuna di queste domande e questa incapacità, più di ogni altra considerazione, dimostra la necessità di affidarci, per cercare di affrontare il problema della coscienza, alla costruzione di artefatti dotati di “coscienza artificiale” (Edelman & Tononi, 2000).

E' necessario sottolineare che l'essere coscienti è completamente separato dall'essere vivi. Essere vivi significa essere costituiti da molecole basate sulla replicazione del DNA, essere coscienti significa essere in grado di fare esperienza del mondo. Il fatto che, per ora gli unici esseri sicuramente coscienti siano anche esseri vivi non deve precludere il tentativo di costruire esseri coscienti senza fare ricorso a strutture biologiche viventi. In base alle conoscenze odierne, niente lega la struttura del DNA o dell'atomo di Carbonio al nostro essere coscienti. E se questo fosse vero significa che è possibile supporre e immaginare il modo in cui una struttura artificiale possa produrre un soggetto dotato di esperienze coscienti. Non perché tale macchina o sistema sia in grado di compiere quel miracolo che un'insufficiente ontologia ci costringe a supporre, ma perché la sfida che ci pone la comprensione della coscienza ci costringerà a mettere in crisi luoghi comuni consolidati ma insufficienti; a rigettare ontologie autorevoli ma eccessivamente semplici; a creare nuovi metodi per descrivere eventi e fatti oggi al di fuori del dominio scientifico. In sintesi, è possibile che il secolare empasse in qui si è imbattuta la scienza, nello spiegare la natura della mente cosciente, sia causato non tanto dalla natura della problema in sé quanto dalle ipotesi che acriticamente si sono accettate sulla natura della realtà. Cambiando queste, e sottoponendole al vaglio dell'evidenza è possibile che si guadagni una migliore comprensione tanto della coscienza quanto della realtà.

Riferimenti

- Agazzi, E. (1981). Intentionality and Artificial Intelligence. *Epistemologia, IV*, 195-228.
- Aleksander, I. (1994). *Towards a Neural Model of Consciousness*. Paper presented at the ICANN 94.
- Aleksander, I. (1996). *Impossible Minds: My Neurons, My Consciousness*.: Imperial College Press.
- Aleksander, I. (2000). *How to Build a Mind*. London: Weidenfeld & Nicolson.
- American, T. E. o. S. (Ed.). (1999). *The Scientific American Book of the Brain*. New York: The Lyon Press.

- Brentano, F. (1874). *Psychologie vom empirischen Standpunkt*. Leipzig: Hahn.
- Brooks, R. A. (1990). Elephants Don't Play Chess. *Robotics and Autonomous Systems*, 6, 3-15.
- Brooks, R. A. (1991). Intelligence Without Representations. *Artificial Intelligence Journal*, 47, 139-159.
- Brooks, R. A., Breazeal, C., Marjanovic, M., Scassellati, B., & Williamson, M. (1999). The Cog Project: Building a Humanoid Robot. In N. C.L. (Ed.), *Computation for Metaphors, Analogy, and Agents* (Vol. 1562): Springer-Verlag.
- Chalmers, D. J. (1996). *The Conscious Mind: in Search of a Fundamental Theory*. New York: Oxford University Press.
- Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Clarendon Press.
- Dennett, D. C. (1969). *Content and consciousness*. London: Routledge & Kegan Paul.
- Dennett, D. C. (1987). *The intentional stance*. Cambridge (Mass.): The MIT Press.
- Di Francesco, M. (1996). *L'io e i suoi sè*. Milano: Raffaello Cortina Editore.
- Dummett, M. (1998). *Origins of Analytical Philosophy*. Cambridge (Mass.): Harvard University Press.
- Edelman, G. M. (1987). *Neural Darwinism. The Theory of Neuronal Group Selection*. New York: Basic Books.
- Edelman, G. M., & Mountcastle, V. B. (1978). *The Mindful Brain*. Cambridge (Mass.): The MIT Press.
- Edelman, G. M., & Tononi, G. (2000). *A Universe of Consciousness. How Matter Becomes Imagination*. London: Allen Lane.
- Editor. (2000). In Search of Consciousness. *Nature Neuroscience*, 3(8), 1.
- Fodor, J. A. (1987). *Psychosemantics: the problem of meaning in the philosophy of mind*. Cambridge (Mass.): The MIT Press.
- Kim, J. (1998). *Mind in a Physical World*. Cambridge (Mass.): The MIT Press.
- Manzotti, R. (2001). *Intentional robots. The design of a goal seeking, environment driven, agent*. Unpublished PhD Thesis, Genova, Genova.

- Manzotti, R., & Tagliasco, V. (2001). *Coscienza e Realtà. Una teoria della coscienza per costruttori e studiosi di menti e cervelli*. Bologna: Il Mulino.
- Picardi, E. (1992). *Linguaggio e analisi filosofica*. Bologna: Patron.
- Praetorius, N. (1999). *Principles of Cognition, Language and Action: Essays on the Foundations of a Science of Psychology*. New York: Kluwer Academic Press.
- Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.
- Sandini, G., Manzotti, R., Metta, G., & Panerai, F. (2000, July 25-27, 2000). *Development: is it the right way towards humanoid robotics?* Paper presented at the ISA-6, Venezia.
- Scassellati, B. (2000, September). *Theory of Mind for a Humanoid Robot*. Paper presented at the First IEEE/RSJ International Conference on Humanoid Robotics.
- Searle, J. R. (1983). *Intentionality, an essay in the philosophy of mind*. Cambridge (Mass.): Cambridge University Press.
- Searle, J. R. (1992). *The rediscovery of the mind*. Cambridge (Mass.): The MIT Press.
- Smith, B. C. (1998). *On the Origins of Objects*. Cambridge (Mass.): The MIT Press.
- Stapp, H. P. (1998). *Whiteheadian Process and Quantum Theory of Mind*. Paper presented at the Silver Anniversary International Conference, Claremont (Cal.).
- Stubenberg, L. (1998). *Consciousness and qualia*. Amsterdam: J. Benjamins Pub.
- Togawa, T., & Otsuka, K. (2000). A model for Cortical Neural Network Structure. *Biocybernetics and Biomedical Engineering*, 20(3), 5-20.
- Varela, F. (2000). *Neurophenomenology*. Paper presented at the Tucson 2000, Tucson.
- Varela, F. J., & Shear, J. (Eds.). (1999). *The view from within. First-person approaches to the study of consciousness.*: Imprint Academic.
- Whitehead, A. N. (1925). *Science and the modern world*. New York: Free Press.
- Whitehead, A. N. (1927). *Process and Reality*. London: The Free Press.
- Zlatev, J. (2001). The Epigenesis of Meaning in Human Beings, and Possibly in Robots. *Mind and Machines*, 11, 155-195.