

DISTINZIONE E RIFLESSIONE*

Giuseppe Trautteur

In un volume nato da un convegno sulla coscienza, i curatori (Marcel e Bisiach, 1988) sentirono la necessità di difendere la legittimità del loro oggetto di studio. Qualche anno dopo, l'incredibile aumento e il largo consenso delle pubblicazioni sull'argomento ci hanno liberato dalla necessità di una tale giustificazione. E tuttavia il venire meno di questa necessità porta ancora con sé una certa inquietudine. Per quale motivo se ci si occupa della coscienza si sente il bisogno di spiegare meglio cosa ciò significhi al nostro collega che magari sta lavorando, per dire, alla teoria dell'omotopia, mentre egli non sente la stessa esigenza? Credo che la ragione sia che lo studio della coscienza è in larga misura ricerca intorno a ciò che "coscienza" significa. Peggio: non vi è alcuna metodologia condivisa. In uno studio sulla coscienza si deve sempre ricorrere a metodologie "regionali", in accordo al proprio orientamento, metodologie che in parte determinano che tipo di coscienza si sta studiando. Alcuni di questi orientamenti e metodologie saranno brevemente ricordati di seguito.

1. Alcuni approcci alla coscienza

La coscienza, naturalmente, è stata affrontata dalle più diverse discipline, ma certamente i tentativi più antichi appartengono alla filosofia. Oggi è la filosofia della mente che in un senso in qualche modo "ufficiale" domina il campo e si confronta con l'intera varietà delle nozioni distinguibili: coscienza, consapevolezza, autocoscienza, autoconsapevolezza, contenuti di coscienza e simili. Si possono trovare – per esempio in Lycan (1987) o Flanagan (1992) – bibliografie aggiornate, così come liste ragionate di termini, posizioni e teorie.

Un secondo filone di approccio alla coscienza si è sviluppato, com'è ovvio, all'interno della psichiatria, della psicologia, della psi-

coanalisi. Qui troviamo interessi clinici, studi fortemente radicati nella psicologia sperimentale e, ovviamente, l'approccio freudiano. Il peculiare sviluppo della psicoanalisi solo recentemente ha consentito un contatto con orientamenti diversi e ciò sembra dipendere da quello che è chiamato modello o teoria funzionalista della mente, ed in particolare dai modelli procedurali o algoritmici dei processi mentali.

Sorprendentemente la coscienza dell'osservatore, totalmente irrilevante nello sviluppo delle scienze fisiche fino a tutto il diciannovesimo secolo, viene alla ribalta con le idee della teoria della relatività e, soprattutto, della meccanica quantistica. È bene sottolineare anzitutto che la pratica della fisica ha tagliato corto su questi problemi, e che la fisica è tanto oggettiva oggi quanto lo era ai tempi dell'ortodossia filosofica laplaciana. E tuttavia, in modi non ancora accertati, i meccanismi della natura (inanimata) sembrano necessitare dei processi coscienti di una parte (senziente) di essa: noi stessi. Questi studi si sono piuttosto allontanati dalla fisica tradizionale e sono perseguiti nella filosofia della fisica: in particolare come Fondamenti della meccanica quantistica. Ma alcuni autori, e fra essi soprattutto Eccles (1980; 1990) e Penrose (1989; 1994), fanno uso della fisica in senso opposto nell'affrontare la coscienza. Cioè, non nel senso che la coscienza sia necessaria alla fisica, ma nel senso che alcuni fenomeni fisici, precedentemente non indagati, possono costituire le basi per una teoria esplicativa della coscienza. Eccles usa il cosiddetto collasso del pacchetto d'onda come *locus* dell'interazione fra un'anima non fisica e il proprio cervello. Le sue idee hanno dato origine a concreti esperimenti neuropsicologici. Questi ultimi sono trattati da Dennett e Kinsbourne in (1992) i quali propongono la teoria della coscienza *multithread*, cioè a più fili. Nel far ciò essi hanno il merito di liberarci da ogni grado di omuncularismo: distruggendo il cosiddetto "teatro cartesiano" si liberano anche della tentazione di usare falsi termini soggettivi, ereditati dal vocabolario della psicologia popolare, che già implica alcuni aspetti della soggettività, mentre lo sforzo della ricerca deve essere proprio quello di spiegare cos'è che costituisce un soggetto a partire da un sostrato fisico. Condividiamo questa tentazione con i biologi che ancora trovano difficile trattarsi dal far uso della loro bestia nera: il finalismo. Penrose, d'altro canto, afferma che la comprensione della coscienza potrà seguire solamente dalla scoperta di

una teoria generale della gravità quantistica. E altri poi, fra cui forse Stapp (1993) è il più esplicito, fondono alcuni aspetti di diverse interpretazioni della meccanica quantistica al fine di dare un resoconto complessivo sia del mondo senziente che di quello inanimato.

Forse l'elemento singolo più importante nel rinnovare lo studio della coscienza è stato l'introduzione di idee provenienti dai fondamenti della matematica e dalla informatica. Queste idee si sono incontrate col problema della coscienza attraverso il tentativo più generale di rendere conto dei fenomeni mentali sulla base, o come un insieme, di procedure effettive. Quest'ultimo è il punto di contatto con la psicologia cognitivista e tale particolare approccio alla mente di solito è visto come necessariamente radicato nella sovraccitata dottrina del funzionalismo, in cui si afferma che la particolare natura del substrato fisico di un'operazione simbolica è irrilevante, purché sia fisica. Non credo che sia necessariamente così e il motivo per cui sono scettico sulla possibilità di prendere una posizione forte in proposito è che la nozione di livello (di realtà o di *software*) e la loro relativa indipendenza – uno degli ingredienti principali di questa dottrina – potrebbe non essere così chiara e legittima come spesso si ritiene. Inoltre ho sostenuto altrove (Trautteur, 1988; 1992) che l'altra nozione su cui si basa la ben nota ipotesi del sistema fisico di simboli, vale a dire la nozione di indirizzo nel *software* come una base esplicativa per il concetto di riferimento, potrebbe essere una falsa via. Inoltre, la più recente teoria le cui parole chiave sono "incarnazione" [*embodiment*], "collocazione" [*situatedness*], "radicamento dei simboli" [*symbol grounding*] sottolinea l'importanza della effettiva forma materiale che gli algoritmi prendono, siano essi parte del mondo fisico reale come indicato da "incarnazione" e "collocazione" nella robotica, siano essi parte dell'ambiente interno che costituisce l'artefatto. Si ha la sensazione che questo punto di vista giochi un ruolo importante nella scoperta di una capacità di conferire senso radicata nelle effettive componenti materiali della transazione simbolica.

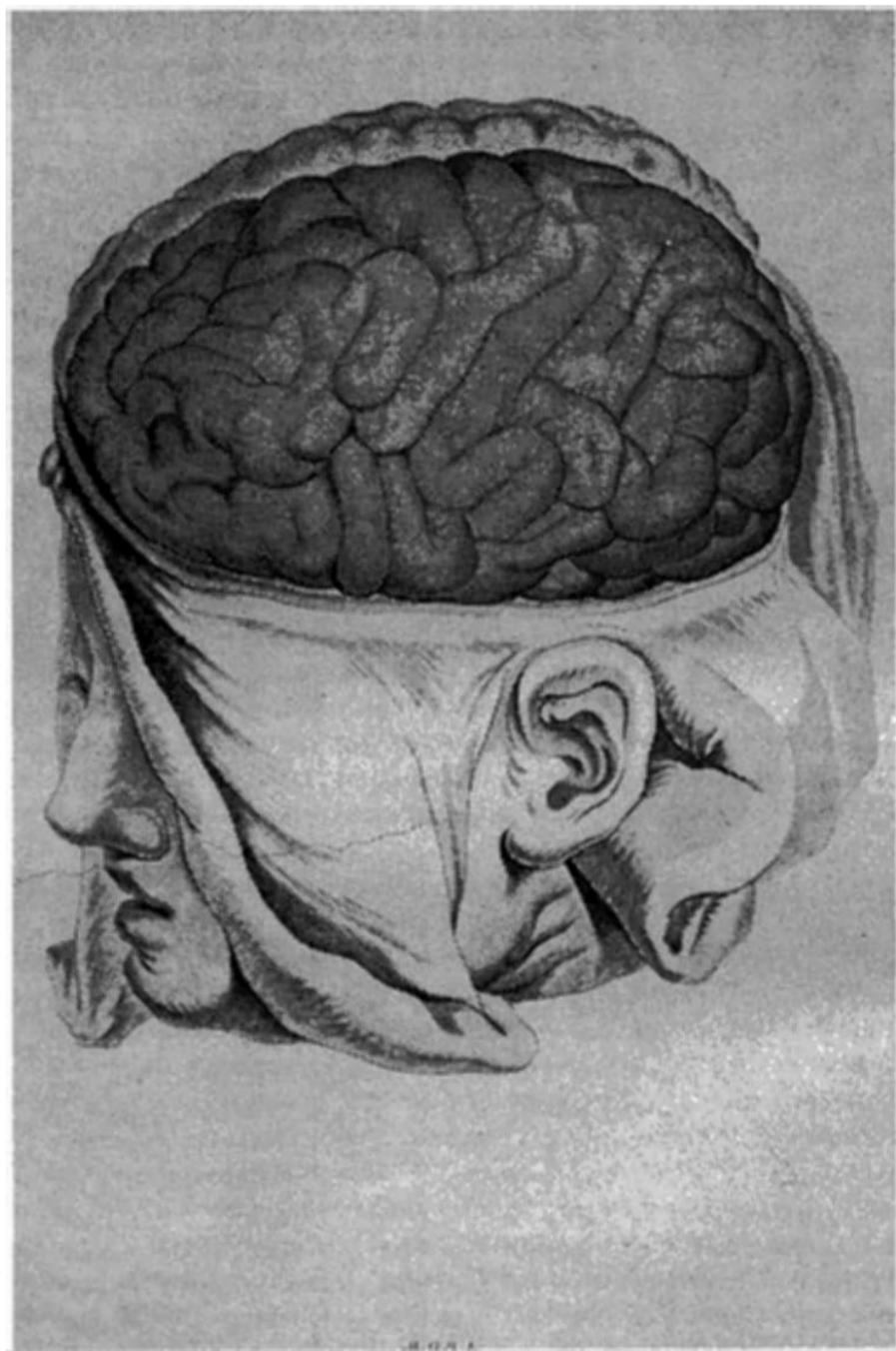
L'insieme delle attività iniziate con la cibernetica, evolutesi nell'intelligenza artificiale e oggi conosciute globalmente come scienza cognitiva o come scienze della cognizione e del cervello ha dato forma a questo approccio. Di nuovo, la demarcazione degli obiettivi e dei metodi è molto difficile da delineare. Sembrano dipartirsi due filoni da

questo approccio. Il primo è più teorico e più vicino alla filosofia della mente. Consiste nella discussione sulle capacità a priori di procedure effettive o di artefatti algoritmici contro le capacità umane o biologiche. Il secondo filone, più vicino alla psicologia, usa la metafora computazionale per i processi mentali ed ha un sapore più distintamente sperimentale. All'interno di questo secondo filone, sebbene talvolta in virulenta opposizione ideologica, ha preso forma il movimento connessionista. Recenti mediazioni quali i sistemi ibridi, non colgono soddisfacentemente il nocciolo del problema, come invece tentano di fare Weyhrauch (1980) e Hofstadter (1976) che considerano contemporaneamente sia gli aspetti fondazionali della teoria della calcolabilità sia le concrete implementazioni nell'*hardware* e nel *software*.

Altri ambiti di ricerca in cui la coscienza gioca un ruolo importante sono la linguistica, la psicolinguistica e gli aspetti neurologici del linguaggio, sia normali che patologici. Solo un esiguo numero di linguisti presta attenzione esplicitamente alla coscienza, per esempio Jackendoff (1987, 1994) e Fauconnier (1984), perché essere consapevoli o coscienti è un presupposto della competenza linguistica. Recentemente tuttavia gli studi sul linguaggio naturale dal punto di vista della (psico)linguistica e da quello dell'intelligenza artificiale mostrano una tendenza ad un più stretto rapporto, anche se ciò che si indaga qui, ed in generale in tutti gli studi nell'area cognitiva, è più l'atto mentale che il problema dell'essenza del soggetto cosciente.

Infine non va dimenticata l'attenzione che la coscienza ha ricevuto negli *studi giuridici* sia come base per la definizione della persona, nel contesto dei diritti civili, sia come sostrato della responsabilità etica – in cui troviamo la coscienza come consapevolezza strettamente connessa alla coscienza morale e dove importanti conseguenze legali e sociali sono correntemente basate su credenze psicologiche piuttosto fragili -, fino ai problemi della coscienza degli animali e dei loro possibili diritti. Di fatto persino la possibilità di conferire diritti civili, sulla base di un preteso possesso di coscienza, ad artefatti come *robot* e programmi era già stata esaminata, per esempio in Putnam (1964).

È anche un po' snervante che questo campo sia disseminato di controversie. Ce ne sono state due principali, dalle quali ne derivano altre due. Il primo dibattito fu originato dalla affermazione di Lucas (1961), argomentata a partire dai teoremi limitativi della logica, che



nessun artefatto potrà mai eguagliare la mente umana. La mente umana deve possedere una particolare proprietà di intuizione e, in ultima analisi, una differenza di natura rispetto alle sue imitazioni artificiali. Sebbene indipendente da quello di Lucas, sia nell'argomentazione che nel contenuto, il cosiddetto argomento della stanza cinese di Searle (1980), che ha originato il secondo dibattito, è tuttavia ad esso correlato, perché ha l'intento di mostrare che alla mente umana, e perciò biologica, è inerente una proprietà, in sostanza l'intenzionalità, inaccessibile alle altre combinazioni fisiche. Entrambe le tesi – quella di Lucas e quella di Searle – sono state fortemente osteggiate, ma non è mai stata avanzata alcuna risposta chiara e definitiva. Più tardi, il recupero delle reti neurali da parte dei fisici, che individuano analogie matematiche tra certi sistemi fisici e il vecchio modello del cervello originariamente proposto da McCulloch e Pitts (1943), e dagli psicologi (Rumelhart *et al*, 1986a; 1986b), ha originato il ben noto dibattito tra il connessionismo e la cosiddetta intelligenza artificiale classica o simbolica, già menzionata sopra. Un interessante *outsider* in questo secondo dibattito è Edelman (1989).

L'ultimo dibattito ha minore risonanza se paragonato ai primi tre, e ha origine nella robotica (Brooks, 1986; 1991a). Nessun tentativo di ricerca sulla coscienza vi è perseguito; l'accento è posto su quelli che potrebbero essere chiamati gli antecedenti evoluzionistici della coscienza, che si ritiene siano la collocazione [*situatedness*], cioè l'agire in un ambiente reale, sensorialmente ricco, e la robustezza, cioè la flessibilità e la capacità di sopravvivenza nell'affrontare l'ambiente reale, che sono qualità peculiari degli esseri biologici reali. Il retroterra teorico sottolinea l'effettiva immersione nella realtà, in contrasto con la simulazione di ambienti artificiali e inventati – in diretta opposizione al funzionalismo – e parzialmente converge con il connessionismo, per il fatto che l'importanza e forse il bisogno di rappresentazioni esplicite, ancora essenziali nei sistemi basati sul funzionalismo, è negata.

2. Temi di ulteriore ricerca

Nella maggioranza degli approcci qui brevemente indicati, si ha l'impressione che il problema venga liquidato prima ancora che la reale ricerca inizi. Si consideri per esempio il soggetto che pronuncia una fra-

se in prima persona. La trattazione di Chisholm (1980) a riguardo, è forse uno degli approcci filosofici più chiari al problema. Tuttavia la sua nozione di "attribuzione diretta" – per intendere ciò che nel nostro contesto chiameremmo "la credenza cosciente su se stessi" – e la conseguente discussione esplicativa addotta, non raggiungono mai ciò che era il tema specifico dell'indagine; la esplicita e dettagliata connessione fra i processi mentali e le loro manifestazioni materiali. Ciò avrebbe dovuto essere tentato dalla cibernetica e dall'intelligenza artificiale – uno degli approcci alla coscienza sopra menzionati – ma i risultati ottenuti non sono quelli che ci si poteva aspettare quaranta anni fa. È accaduto che l'unità dell'agente cosciente e le sue esperienze interne fossero rapidamente perse di vista, se mai seriamente considerate, e che la ricerca fosse smembrata in molte direzioni diverse, originando ciò che Dennett ha chiamato lo studio del mentale senza il soggetto. Questo stato di cose è forse un ulteriore elemento che spinge lo psicologo Jaynes a esprimere una certa diffidenza nei confronti dell'approccio algoritmico alla coscienza: «Solo se potessimo mostrare che la mente funziona come un calcolatore, e con una certa precisione, avremmo un terreno solido per simulare la coscienza usando programmi di calcolatore: si noti che l'ordine qui è importante. La questione non è se il calcolatore sia simile alla mente, perché sicuramente lo è. I calcolatori già simulano numerosi processi mentali. Ma ciò è come dire che una mappa di Venezia è simile a Venezia. Certamente lo è. Ma ... l'analogia nell'altra direzione, che Venezia sia simile alla mappa, ha poco significato, ammesso che ne abbia. Ed è lo stesso, penso, per il paragone di un insieme di programmi di calcolatore in qualsivoglia struttura o combinazione, con tutte le elaborazioni parallele che si desiderino aggiungere, rispetto al funzionamento mentale. Quei programmi di calcolatore possono essere simili ad un insieme di funzioni mentali per il fatto che sono stati progettati per esserlo. Ma ciò non rende valida l'affermazione che la mente è simile a un calcolatore. Le somiglianze nella vita reale non sono transitive» (Jaynes, 1995).

Ulteriori temi di ricerca potrebbero allora essere forse riassunti in due punti: uno sostanziale ed uno metodologico.

Il primo: assumere una posizione di dichiarata adesione alla materialità e alla fisicità, anche se il nostro scopo è quello di rendere conto delle sensazioni interiori, unitarie, private che il soggetto prova quan-

do parla (o percepisce, o agisce). Ciò viene a volte indicato come la trasparenza del sé a se stesso. Questo rendere conto, tuttavia, deve essere soggetto ai criteri di adeguatezza comuni a ogni spiegazione scientifica intersoggettiva. Sembra che ciò sia quanto Crick (1994) voglia esprimere con la sua *astonishing hypothesis*.

Il secondo: non ricercare una teoria completa della coscienza, né aspirare in particolar modo ad una precisa denotazione o definizione del termine "cosciente". Sforzarsi piuttosto di sperimentare procedure effettive (per esempio algoritmi, programmi, reti, ecc.) allo scopo di valutare ed afferrare almeno alcuni degli elementi, strutture e principi architettonici che sembrano essere importanti condizioni necessarie – anche se non sufficienti – per il darsi della coscienza nel dominio biologico e, possibilmente, in quello artificiale. Un traguardo intermedio, ma straordinariamente proficuo, potrebbe essere la spiegazione del motivo esatto per cui certi aggregati materiali sono coscienti – di fatto sono persone – mentre la maggioranza non lo sono.

Essendo desiderosi di comprendere la coscienza come essa esiste biologicamente, molto più che una sua mera duplicazione *in corpore vili*, dovremmo chiederci quanto utile possa essere l'approccio della scienza cognitiva e della IA. Dopo tutto cosa ci garantisce che una duplicazione anche soddisfacente (ma sulla base di quali criteri?) della coscienza – tralasciandone, naturalmente, il tremendo impatto sociale ed etico – potrebbe illuminarci sulla sua natura? Sebbene questa domanda metodologica rimanga irrisolta, ne accettiamo il rischio implicito che sottostà all'intero approccio della scienza cognitiva e della IA, incluso il tentativo di Edelman, e facciamo lo stesso con le difficoltà che esso genera – efficacemente espresse dalle osservazioni di Jaynes su Venezia e le sue mappe. Benché i due punti precedenti non costituiscono ancora un programma di ricerca, sono tuttavia in sintonia con una ipotesi di lavoro che ritengo stia prendendo forma. Questa ipotesi non è mai stata formulata chiaramente, ma è individuabile come retroterra unificante nel lavoro di molti ricercatori del campo. Si basa su due ingredienti: una distinzione del "sé" in due parti e una riflessione di queste parti l'una nell'altra tale da costituire un'unità. L'agente su cui fa perno l'atto riflessivo è inizialmente una "terza" parte rispetto alle parti precedentemente separate dalla distinzione. Da qui il nome: "schema a terza persona".

3. Lo schema a "terza persona"

Il concetto di terza persona che prendiamo in esame qui è quello di uno spettatore o testimone. Come tale emerge di volta in volta nella letteratura filosofica, psicologica e linguistica. Ci può essere una certa somiglianza con la cosiddetta obiezione del "Terzo Uomo" alla teoria platonica delle idee nel Parmenide [si veda (Schweizer, 1994) per una recente discussione dell'argomento]. Una entità di questo genere è stata denominata "osservatore nascosto" nel contesto dell'ipnosi (Hilgard, 1986; Watkins e Watkins, 1979-80) e compare come una parte o un aspetto della persona – conscia e inconscia, normale e patologica – in svariate situazioni neurologiche e psicologiche. Si veda per esempio Gabel (1988) per la sua rilevanza nella ricerca sul sogno, e Marcel (1993) per la sua importanza nella anosognosia. Una terza persona compare anche negli studi semantici o lessicografici come generico parlante della lingua che, come testimone, fissa o fonda il significato di una parola che viene definita. Le prime due persone sono, ovviamente, l'autore del dizionario e il suo lettore generico.

Comunque, molte di queste terze persone assomigliano del tutto ad individui reali almeno nel senso che l'interazione a tre parti fra il parlante/attore, l'ascoltatore/paziente, e lo spettatore indipendente non è altro che un'analogia con una semplice interazione sociale. C'è qui il pericolo della fallacia omuncolare, dal momento che il peso della spiegazione del soggetto cosciente è differito ai suoi agenti interni. Minsky (1989), nel suo *The Society of Mind*, usa in modo assai pertinente la metafora della "società" come uno schema architettonico, senza cadere, naturalmente, nella trappola omuncolare. Ma la valutazione di Minsky del problema della persona non pone in rilievo il sé: la coscienza è vista, essenzialmente, come un resoconto successivo. Tuttavia il principio architettonico coinvolto nel suo monitoraggio del cervello *B* sul cervello attivo è certamente sulla linea della "terza persona".

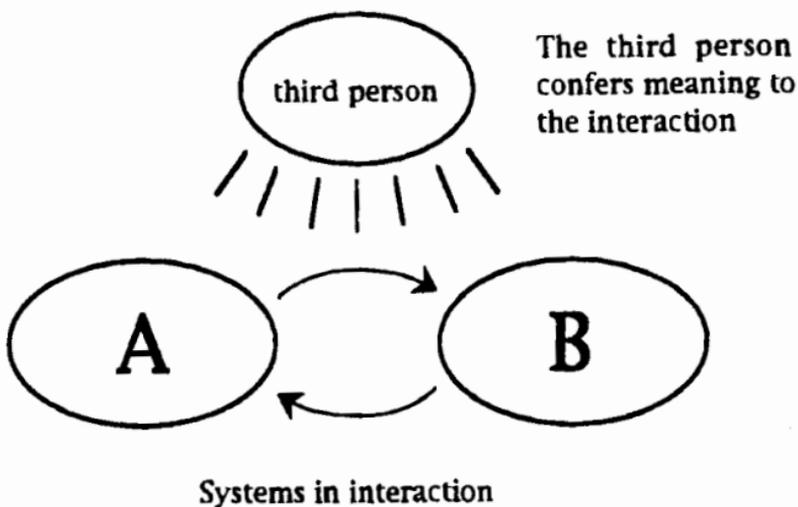
Il passo successivo compiuto nello schema a terza persona, che è proposto qui come un sistema implementabile, sta nel fatto che la terza persona deve coincidere con l'intero soggetto cosciente. Assumiamo che due sistemi, *A* e *B*, siano impegnati in una conversazione (si veda il riquadro *a* della figura a pag. 139) e che entrambi sostengano il ruolo del parlante o dell'ascoltatore a seconda delle circo-

stanze. È comunemente accettato che la conversazione ha un significato reale, intenzionale solo nel caso che i sistemi siano esseri umani. Tuttavia, nel caso in cui né *A* né *B* siano esseri umani potrebbe comunque darsi che la loro conversazione sia dotata di significato per una terza entità umana spettatrice. Questo spettatore è la persona che conferisce significato e può entrare nel disegno sostituendosi ad uno dei sistemi *A* o *B*. Questo è ciò che comunemente accade quando usiamo un calcolatore in modo interattivo. Il punto è che un essere umano può allo stesso tempo sia prendere parte all'interazione sia – inconsciamente o preconsiamente, per così dire – esserne testimone. Questo “esser testimoni” può o meno essere usato poi in modo introspettivo, ma sembra costituire un ingrediente necessario di quel nucleo della coscienza che stiamo inseguendo: *la sensazione di essere presenti a se stessi*. Weyhrauch implicitamente esprime il nocciolo di questo schema con il costrutto META presente nel suo FOL (Weyhrauch, 1980) in cui il numero dei livelli di riflessione è due.

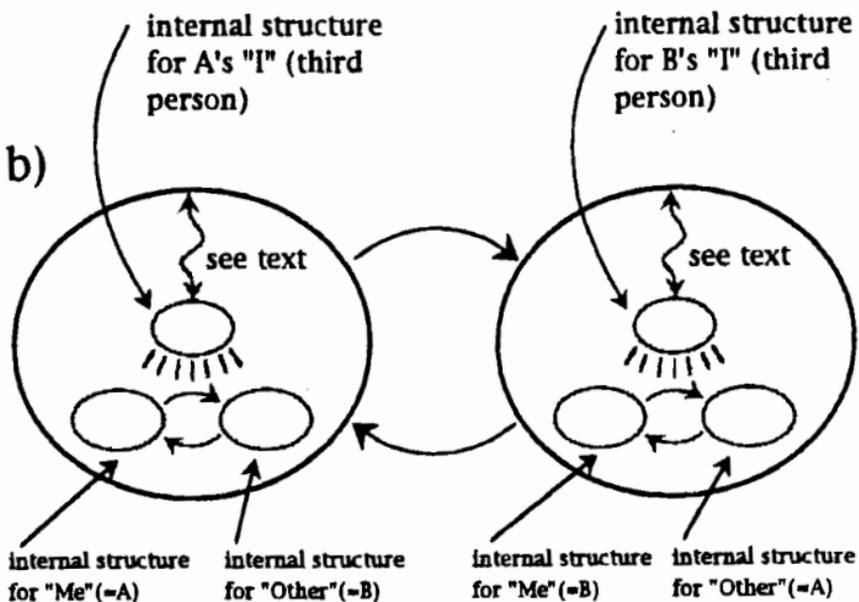
Per poter essere testimone di ciò che accade, il soggetto deve farsi un'idea – avere un modello interiore – dell'interazione che ha luogo, inclusi uno o più *simulacra* di se stesso. Suppongo che l'interiorizzazione della situazione generi – di fatto richieda – negli agenti *A* e *B* una struttura *extra* (si veda il riquadro *b* della figura) che *a priori* non sembrerebbe essere necessaria per l'esecuzione di compiti cognitivi, ma non coscienti. Questa struttura consiste di un insieme di modelli interni del sistema stesso, di entità esterne – l'altro partecipante al dialogo o, in generale, l'ambiente – e della “terza persona” (si veda la figura a pagina seguente).

Ma chi compie l'atto del “vedere”? Chi compie l'introspezione? Un tale schema sembra essere soggetto ad una ricorsione senza fine. E certamente, modelli mentali ricorsivi, nei quali il soggetto pensa (o modella) il soggetto, che pensa (o modella) il soggetto, e così via, hanno avuto un ruolo importante nelle teorie che si sono candidate alla spiegazione della coscienza. Johnson-Laird nel suo contributo a Marcel-Bisiach (1988): Johnson-Laird (1988a) e in Johnson-Laird (1988b) e Brian Cantwell Smith (1984, 1985, 1986) hanno trattato estesamente questo punto. Ma la ricorsione non è sufficiente. Certamente la ricorsione non è molto allettante come principio architettonico perché non ci sono entità fisiche genuinamente ricorsive. In

a)



b)



definitiva tutti gli artefatti ricorsivi (programmati) riposano sull'iterazione e quindi su un calcolo che implica un estremo superiore fissato per il numero dei distinti processi ricorsivi. Per porre fine alla *mise en abîme* (Dällenbach, 1977) della trappola ricorsiva, una parte della struttura interna, potenzialmente in illimitato sviluppo, deve diventare "identica" all'intero sistema. (È interessante notare che, leggendo "identico" con il significato di "in corrispondenza biunivoca", si avrebbe allora ciò che caratterizza anche gli insiemi di cardinalità infinita.) Nel nostro schema l'identità non può essere coincidenza materiale, altrimenti non sarebbero entità distinte. Quindi queste strutture algoritmiche devono essere distinte e tuttavia le stesse. Ma cosa significa esattamente che due procedure effettive sono identiche? Cos'è esattamente una singola procedura? È qui che l'analisi del fondarsi nel sostrato materiale e dell'identità fisica degli *alias* di una struttura all'interno del medesimo sistema *software* diventano essenziali. Per quanto ne so, solo Pitrat (1989, 1990) ha svolto una ricerca in questa direzione, analizzando da un punto di vista metalinguistico il fenomeno dell'auto-avvio [*bootstrap*] di una procedura e la natura delle differenti entità intermedie che occorrono nel processo.

Il punto cruciale – ripetiamolo: non per una spiegazione completa della coscienza, ma per raggiungere almeno il suo necessario elemento essenziale: l'identificazione, attraverso la riflessione, di una entità con una parte di se stessa – è indicato dalle frecce tratteggiate nel riquadro *b* della figura e dipende da problemi ancora irrisolti, ma accessibili, dell'architettura del *software*. È qui che interviene la nozione di parallelismo. Il parallelismo non deve essere inteso come l'esecuzione simultanea di processi indipendenti, ma come lo sviluppo spaziale di un algoritmo complesso in esecuzione. Damasio, nel suo *Descartes' Error* (Damasio, 1994), in cui tratta principalmente le connessioni tra emozione e cognizione, sostiene anche – in un modo che ricorda il concetto junghiano di sincronicità – che la coscienza deve dipendere dal fatto che processi in qualche modo correlati accadono allo stesso tempo, ma in luoghi diversi del cervello.

Mentre un algoritmo, così come una struttura o un testo, è situato fisicamente dove lo sono i suoi simboli, è molto difficile stabilire dove è situato un algoritmo attivo in esecuzione. Non è in nessun luogo e al tempo stesso è in molti luoghi diversi. Può darsi che la massima di

Hume: «un oggetto può esistere, e tuttavia non essere in alcun luogo» (D. Hume, *A Treatise of Human Nature*, Book I, Part IV, Sect. V, [Penguin Books, 1969, p. 284]) – che egli scrisse nella sua infruttuosa ricerca sull'anima – sarà confermata dalla spiegazione algoritmica della coscienza. Come suggerito prima, un "differente" principio di identità potrebbe valere per queste entità attive e diffuse, che potrebbero dunque far parte della *res extensa*, mantenendo però gli attributi della *res cogitans*. Suppongo, in sostanziale accordo con la trattazione di Chisholm delle espressioni in prima persona, ma con un fondamento fisico molto più esplicito, che l'autoriferimento meccanico è antecedente e fondante per una capacità referenziale veramente consapevole e intenzionale. Da dove si origina in noi la comprensione autentica? C'è la speranza che un lavoro di sperimentazione algoritmica, basato sulle linee appena illustrate, possa costituire l'inizio di una considerazione innovativa del problema della coscienza.

* Traduzione e adattamento di "Distinction and reflection" in: TRAUTTER, G. (Ed.), *Consciousness: Distinction and Reflection*, Bibliopolis, 1995, con licenza dell'Editore.

RIFERIMENTI BIBLIOGRAFICI

- BROOKS, R. (1986), *A Robust Layered Control System for a Mobile Robot*, in «IEEE J. Of Robotics and Automation», RA2-1 (1986), pp. 14-23.
- BROOKS, R. (1991a), *Intelligence without representation*, «Artificial Intelligence» 47 (1991), pp. 139-59.
- CHISHOLM, R. (1980), *The First Person: An Essay on Reference and Intentionality*, Harvester Press, 1980.
- CRICK, F. (1994), *The Astonishing Hypothesis*, Scribner, 1994.
- DÄLLENBACH, L. (1977), *Le récit spéculaire. Essai sur la mise an abîme*, Seuil, Paris, 1977.
- DAMASIO, A.R. (1994), *Descartes' Error*, Grosset/Putnam, 1994, (trad. it., *L'errore di Cartesio*, Adelphi, Milano, 1995).
- DENNETT, D.C. and KINSBOURNE, M. (1992), *Time and the observer: The where and when of consciousness in the brain*, «Behav. and Brain Sci.», 15 (1992), pp. 183-247.
- ECCLES, J.C. (1980), *The Human Psyche*, Springer Verlag, 1980.
- ECCLES, J.C. (1990), *A unitary hypothesis of mind-brain interaction in the cerebral cortex*, «Proc. R. Soc. London Biol. Sci.», 240 (1990), pp. 433-51.
- EDELMAN, G.M. (1989), *The Remembered Present: A Biological Theory of Consciousness*, Basic Books, 1989, (trad. it., *Il presente ricordato*, Rizzoli, Milano, 1991).

- FAUCONNIER, G. (1984), *Espaces mentaux*, Les Éditions de Minuit, 1984.
- FLANAGAN, O. (1992), *Consciousness Reconsidered*, Bradford Books, 1992.
- GABEL, L. (1988), *Dissociative phenomena and monitoring of self: Experimental, clinical and theoretical considerations*, «Integr. Psychiatr.», 6 (1988), pp. 53-60.
- GRAY, J.A. (1995), *The Contents of Consciousness: a Neuropsychological Conjecture.*, «Behav. and Brain Sci.», in stampa.
- HILGARD, E.R (1986), *Divided Consciousness. Multiple Controls in Human Thought and Action*, Wiley, 1986.
- HOFSTADTER, D. R. (1979), *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic Books, 1979, (trad. it., *Gödel, Escher, Bach*, Adelphi, Milano, 1984).
- HUME, D., *A Treatise of Human Nature*, 1739, (trad. it., *Trattato sulla natura umana*, a cura di E. LECALDANO, 3 voll., Laterza, Bari, 1975).
- JACKENDOFF, R. (1987), *Consciousness and the Computational Mind*, MIT Press, 1987, (trad. it., *Coscienza e mente computazionale*, Il Mulino, Bologna, 1990).
- JACKENDOFF, R. (1994), *Patterns in the Mind*, Basic Books, 1994.
- JAYNES, J. (1995), *The Diachronicity of Consciousness*, in G. TRAUTTEUR, (Ed.), *Consciousness: Distinction and Reflection*, Bibliopolis, 1995.
- JOHNSON-LAIRD, P.N. (1988a), *A Computational analysis of consciousness.*, in MARCEL, A.J., BISIACH, E. (Eds.), *Consciousness in Contemporary Science*, Clarendon Press, 1988.
- JOHNSON-LAIRD, P.N. (1988b), *The Computer and the Mind*, William Collins & Co, 1988 (trad. it., *La mente e il computer. Introduzione alla scienza cognitiva*, Il Mulino, Bologna).
- LEIDLMAIR, K. (1995), *The schizofrenic computer*, in G. TRAUTTEUR, (Ed.), *Consciousness: Distinction and Reflection*, cit.
- LUCAS, J.R. (1961), *Minds, Machines and Gödel.*, in «Philosophy», XXXVI (1961), pp. 112-27.
- LYCAN, W. (1987), *Consciousness*, Bradford Books, 1987.
- MARCEL, A.J. (1993), *Slippage in the unity of Consciousness*, in CIBA Foundation Symposium 174, *Experimental and theoretical studies in consciousness*, Wiley, 1993.
- MATURANA, H. (1995), *Biology of Consciousness*, in G. TRAUTTEUR, (Ed.), *Consciousness: Distinction and Reflection*, cit.
- MARCEL, A.J., BISIACH, E. (Eds.), *Consciousness in Contemporary Science*, Clarendon Press, 1988.
- MCCULLOCH, W.S., PITTS, W. (1943), *A logical calculus of the ideas immanent in the nervous activity*, in «Bull. Math. Biophysics», 5 (1943), pp. 11-43.
- MINSKY, M. (1986), *The Society of Mind*, Simon and Schuster, 1986.
- PENROSE, R. (1989), *The Emperor's New Mind*, Oxford Univ. Press, 1989 (trad. it., *La mente nuova dell'imperatore*, Rizzoli, Milano, 1992).
- PENROSE, R (1994), *Shadows of the Mind*, Oxford Univ. Press., 1994 (trad. it., *Ombre della mente*, Rizzoli, Milano, 1996).
- PITRAT, J. (1989), *Qu'est-ce qu'un individu au royaume de la connaissance?* in «Cahiers du Laforia», 77 (1989), pp. 37-58.
- PITRAT, J. (1990), *Metaconnaissance*, Hermes, 1990.
- RUMELHART, D.E., MCCLELLAND, J.L., PDP RESEARCH GROUP (1986ab), *Parallel Distributed Processing. Vol.1: Foundations*, MIT Press, 1986.
- RUMELHART, D.E., MCCLELLAND, J.L., PDP RESEARCH GROUP (1986b), *Parallel*

- Distributed Processing, Vol. 2: Psychological and Biological Models*, MIT Press, 1986.
- SCHWEIZER, P. (1994), *Self-predication and the Third Man*, in «Erkenntnis», 40 (1994), pp. 21-42.
- SEARLE, J. (1980), *Minds, Brains, and Programs*, in «Behav. and Brain Sc.», 3, (1980), pp. 41.
- SMITH, B.C. (1984), *Reflection and Semantics in LISP*, in *Conf. Record of the 11th POPL of the Ass. for Computational Linguistics*, Jan., 1984.
- SMITH, B.C. (1985), *Prologue to Reflection and Semantics in a Procedural Language*, in R.J. BRACHMAN, H.J. LEVESQUE (Eds.), *Readings in Knowledge Representation*, Morgan Kaufman, 1983, pp. 31-9.
- SMITH, B.C. (1986), *Varieties of Self-reference*, in HALPERN, J.Y. (Ed.), *The Theoretical Aspects of Reasoning about Knowledge: Proc. of the 1986 Conference*, Morgan Kaufmann, 1986, pp. 19-43.
- STAPP, H.P., *Mind, Matter and Quantum Mechanics*, Springer Verlag, 1993.
- TRAUTTEUR, G. (1988), *On Addresses*, in *Proc. of the 1988 IEEE Inter'l Conf. on Systems, Man and Cybernetics*, Beijing, 1988, Vol. II, pp. 1267-9.
- TRAUTTEUR, G., *Problems with Symbols. A Commentary to Herbert Simon "Scientific Discovery as Problem Solving"*, in «Int. Studies in the Phil. of Sci.», 6 (1992), pp. 61-4.
- WATKINS, J.G., WATKINS, H.H. (1979-80), *Ego states and hidden observers*, in «J. of Altered states of Consciousness», 5 (1979-80), pp. 3-18.
- WEYHRAUCH, R. (1980), *Prolegomena to a Theory of Mechanized Formal Reasoning*, in «Artificial Intelligence», 13 (1980), pp. 133-70.