

FABIO DE SIO

"MACHINAE DOCILES".**MODELLI, MACCHINE, CERVELLI**

1. Premessa 2. Le neuroscienze come scienza pratica 3. Modelli nelle neuroscienze
4. Modelli, macchine, cervelli 5. Il Criterio, l'Icona e la Virtù

... abbiamo cominciato le nostre discussioni e sessioni nello spirito del "come se". [...] Abbiamo esplorato ogni sorta di "se". Poi, piuttosto bruscamente, mi sembra, abbiamo cominciato a parlare la lingua dell'"è". Dicevamo all'incirca le stesse cose, ma ora come se fossero davvero così. Questo mi fa ricordare una definizione di gravidanza: "il risultato di aver preso sul serio qualcosa che era stato porto per gioco", e mi domando se non siamo diventati gravidi e in pericolo di parto prematuro.

(Ralph Gerard, 1950)

1. Premessa

Il modello (lat. *Modulus*, da *modus*) è una misura. La misura un'unità convenzionale che si confronta con gli oggetti, per conoscerne il rapporto. Nel linguaggio comune, con il termine "modello" s'indica il prototipo, o la rappresentazione in scala di qualcosa da realizzare, o ancora un esemplare o un esempio. Il riferimento alla misura è sempre presente, seppur in modo non univoco.



cyop&kaf misentropia

Nelle scienze, l'identità fra modello e misura è essenziale, e in questo il modello si distingue dalla metafora e dall'analogia come uno specifico modo di rappresentazione¹. La scelta di un modello nella ricerca scientifica è un atto assoluto, arbitrario, ancorché, ovviamente, non strettamente soggettivo. Questo perché la funzione del modello, potremmo dire la sua ragione, è economica.

¹ H. J. Rheinberger, *Toward a History of Epistemic Things. Synthesizing Proteins in a Test Tube*, Stanford University Press, Stanford (CA) 1997, p. 111.

In relazione al risparmio o valorizzazione di una certa risorsa (tempo, energia mentale) un modello può essere inteso come una scorciatoia, una rappresentazione intuitivamente accessibile di un fenomeno complesso o addirittura di un referente in tutto o in parte ignoto (si pensi all'atomo di Bohr). Ancora, si può interpretarlo come l'istanza estrema di una legge di natura (cioè il caso ideale in cui la legge è completamente rispettata in ogni suo aspetto) o, epistemicamente, come una «macchina nomologica», un generatore induttivo di generalizzazioni².

2. Le neuroscienze come scienza pratica

L'elemento comune a tutte queste accezioni è il riferimento all'utilità, a una dimensione relazionale e contingente. Un modello, dunque, non è valido perché vero, ma in quanto utile. Si tratta di una differenza profonda, ben espressa da Duhem all'inizio del secolo: nel criticare il vizio (per lui tipicamente inglese) della modellizzazione, egli ha espresso questo scarto nella contrapposizione tra «*esprit profond*» ed «*esprit ample*», fra una scienza cartesiana, intesa come conseguimento della verità per via razionale, e una scienza baconiana, il cui «oggetto è del tutto pratico, [...] del tutto industriale»³.

Duhem, tuttavia, sembra aver sottovalutato la differenza fra “pratico” e “applicativo” (leggi: “industriale”). Il concetto di “applicazione”, infatti, «ha senso soltanto nel contesto di una distinzione già presupposta tra teoria e pratica; laddove una simile disgiunzione sia rifiutata, l'applicazione diventa per definizione parte costitutiva della pratica»⁴. In una prospettiva pratica, laddove la possibilità di predizione e manipolazione è in sé un indice di successo, la questione della verità passa in secondo piano, il vero criterio della scienza diviene la produttività. Non è un caso che nella filosofia della scienza, soprattutto di lingua anglosassone, la crescente attenzione per i modelli sia coincisa con uno spostamento dell'ottica dalla dimensione teorica alla pratica, dalla rappresentazione (e alla questione dell'adeguatezza) al rappresentare (e alla questione dell'efficacia)⁵.

² Cfr. N. Cartwright, *Models: The blueprints for laws*, in «Philosophy of Science», 64, 1997, pp. 292-303.

³ P. M. M. Duhem, *La théorie physique: son objet, et sa structure*, Chevalier & Rivière, Paris 1906, p. 108.

⁴ E. F. Keller, *Models of and Models for: Theory and Practice in Contemporary Biology*, in «Philosophy of Science», XLVII (supplement), 2000, pp. S72-S86, qui p. S75.

⁵ Cfr. I. Hacking, *Representing and Intervening. Introductory Topics in the Philosophy of Natural Sciences*, Cambridge University Press, Cambridge 1983.

Nell'interrogare il significato di questa sostituzione, i suoi limiti e le sue conseguenze, farò qui riferimento a due tendenze fondamentali delle neuroscienze, storicamente incarnate in due momenti della definizione progressiva di questo campo di studi: l'elaborazione di modelli meccanici dei processi mentali e l'emergere delle nuove tecnologie di visualizzazione dell'attività cerebrale. Tale selezione si basa sull'assunto che un'indagine sui modelli nelle neuroscienze al contempo permetta una comprensione più chiara della loro natura strumentale e mediata, e apra una diversa prospettiva sul confine fra il naturale e l'artificiale, il mondo biologico e quello culturale. I modelli meccanici del cervello incarnano in quest'ottica tanto una posizione metafisica quanto un'opzione epistemologica forte. Cioè a dire, essi offrono una possibile risposta tanto sulla natura dei processi adattativi e l'adeguatezza della loro descrizione in quanto sistemi fisici (e, viceversa, l'adeguatezza di una descrizione funzionale e teleologica delle macchine), quanto sulla possibilità di stabilire una relazione significativa fra fenomeni, livelli di spiegazione e programmi di ricerca separati e indipendenti. In entrambi i casi, si può dire che i modelli mettano in relazione livelli separati di realtà, cioè ambiti fenomenici in linea di principio autonomi, non necessariamente riconducibili l'uno all'altro. Vedremo infine in che senso si possa affermare che, nell'attuare queste connessioni non necessarie, i modelli possano assurgere alla funzione di produttori di realtà.

L'ipotesi di partenza è che il significato euristico ed epistemologico della modellizzazione nelle neuroscienze sia legato a due nuclei problematici principali: (1) la questione della teleologia, che considero consustanziale allo studio del comportamento; (2) il particolare tipo di riduzione esplicativa di queste discipline, in cui confluiscono tre dimensioni in principio autonome (strutturale o anatomica; funzionale o fisiologica; fenomenica o psicologica) in assenza di una gerarchia ontologica ed esplicativa prestabilita. V'è poi una terza ipotesi, che riguarda il carattere pratico delle neuroscienze in senso lato: per il loro oggetto e per la tipica riflessività del loro procedere (in cui ipotesi sul e modelli dell'oggetto di studio riverberano sull'ente osservatore – cervelli che studiano se stessi), le neuroscienze in generale possono essere intese come scienze pratiche in sommo grado. Il concetto di "scienze pratiche", come accennato, va inteso nel senso che il momento teorico della generalizzazione e quello pratico della manipolazione non sono separabili. Il modello-misura, come cercherò di mostrare, agisce da mediatore fra questi

diversi piani, da luogo di sintesi di ambiti fenomenici diversi e, infine, da produttore di realtà nella forma di nuovi nessi fra l'essere, il rappresentare e il dover essere.

3. Modelli nelle neuroscienze

L'orizzonte teorico delle neuroscienze è organizzato intorno al problema dell'adattamento individuale e, quindi, del funzionamento teleologico degli organismi. Ovviamente, il problema della teleologia non è esclusivo delle neuroscienze. La biologia nel suo insieme, e per tutta la sua breve storia, si è dovuta misurare con questo peccato originale, né si può dire che una soluzione degna di cotanto impegno sia stata raggiunta⁶. Come ha osservato lo zoologo John Z. Young: «I biologi hanno tentato in ogni modo di eliminare la teleologia dalla loro scienza, come era avvenuto nella fisica e nella chimica. [...] È lecito inferire che il tipo di spiegazioni in uso nella scienza fisica siano incompatibili con spiegazioni in termini di "scopo" [*purpose*], e questo ci costringe ad affrontare un paradosso che non può al momento essere completamente risolto»⁷. Vi sono, tuttavia, diversi modi per aggirare questo formidabile ostacolo. Il più diffuso, forse perché più semplice, consiste nella rimozione del problema o nel suo spostamento a un livello superiore. È quanto accade, per esempio, con le spiegazioni evolutive, cioè con la giustificazione di uno stato di cose in base alla sua storia e a un principio economico come è la selezione naturale. Una delle virtù dell'evoluzionismo è proprio la messa fra parentesi del problema della causa agente⁸.

Un simile procedimento non può appartenere all'orizzonte generale delle neuroscienze, meno che mai allo studio del cervello. Ponendosi come ultimo oggetto il comportamento adattativo e il suo sostrato fisiologico, questa costellazione di discipline non può evitare il confronto diretto con il problema ricorrente della relazione fra causalità e finalità, meccanismo e intenzionalità, leggi della fisica e diavoleto di Maxwell⁹. È significativo che i primi compiuti modelli meccanici del sistema nervoso siano stati sviluppati a cavallo della Seconda Guerra Mondiale, nel tentativo di affrontare direttamente, piuttosto che

⁶ Cfr A. Volpone, *Teleonomia e altre forme di naturalizzazione della "scienza dei fini" in biologia*, in «Quaderni della Scuola di Nettuno», 2005, pp. 11-135.

⁷ J. Z. Young, *Philosophy and the Brain*, Oxford University Press, Oxford 1987, p. 4.

⁸ D. Tarizzo, *Come Darwin ha cambiato la filosofia?*, in «S&F_scienzaefilosofia.it», www.scienzaefilosofia.it, 2, 2009, pp. 152-162.

⁹ Cfr. E. F. Keller, *Refiguring life: Metaphors of twentieth-century biology*, Columbia University Press, New York 1995, cap. 2.

respingere come non-scientifico o irrilevante, il problema della teleologia e della sua riduzione a meccanismo¹⁰. Di più: questa ristrutturazione di campo non è derivata semplicemente da un approfondimento delle conoscenze sul sistema nervoso, o da singole grandi scoperte, quanto dall'esperienza di macchine di un tipo nuovo, assai più complicate e dotate di capacità assimilabili a quelle di sistemi nervosi complessi: computazione, adattamento e predizione¹¹. La risonanza tra l'approfondimento delle proprietà elettriche del sistema nervoso e lo sviluppo di macchine elettroniche sembrerebbe aver aperto un nuovo spazio di comparazione. Un'apertura dovuta alla scoperta che «c'è talmente tanto in comune fra i meccanismi viventi e quelli artificiali che un paragone tra loro non appare più così ingenuo come quando la parola "macchina" richiama l'immagine di una locomotiva o di un telefono»¹². Il nuovo approccio al cervello come organo di manipolazione di segnali, quindi, sembra essere penetrato nella ricerca sperimentale dal mondo delle macchine.

Storicamente, lo sviluppo di questi nuovi modelli ha contribuito a portare a sintesi prospettive diverse e sino ad allora programmaticamente distinte: la strutturale o molare (anatomia); la funzionale o molecolare (fisiologia); la fenomenologica (psicologia)¹³, facilitando il collegamento fra osservazioni in principio indipendenti come il carattere adattativo del comportamento, la presenza di connessioni ridondanti nel cervello e quella di veri e propri circuiti nervosi. Il modello del meccanismo con *feedback* consentiva di combinare in modo sensato diversi livelli d'analisi e, cosa più importante, di saggiare questa combinazione sperimentalmente.

La possibilità di concepire plausibili modelli meccanici del sistema nervoso e del comportamento rappresenta ben più che una semplice licenza epistemica: si tratta di un'autentica istanza filosofica. *Behavior, purpose and teleology* di Bigelow, Rosenblueth

¹⁰ Cfr. R. Cordeschi, *The Discovery of the Artificial: Behavior, Mind, and Machines Before and Beyond Cybernetics*, Kluwer Academic Publishers, Dordrecht 2002.

¹¹ Cfr. A. Rosenblueth, N. Wiener, J. Bigelow, *Behavior, purpose and teleology*, in «Philosophy of science», X, 1, 1943, pp. 18-24; J. Z. Young, *Doubt and certainty in science: a biologist's reflections on the brain*, Clarendon Press, Oxford 1951.

¹² W. Grey Walter, *Features in the electro-physiology of mental mechanisms*, in *Perspectives in Neuropsychiatry*, a cura di D. Richter, H. K. Lewis, London 1950, pp. 67-78, qui p. 69.

¹³ Cfr. J. B. Watson, *Psychology as the behaviorist views it*, in «Psychological review», XX, 2, 1913, pp. 158-177; F. B. Skinner, *The Behavior of Organisms*, Appleton-Century, New York 1938; K. S. Lashley, *Brain Mechanisms and Intelligence*, Hafner Publishing Co., Inc., New York 1963 [1929]; F. De Sio, *La scienza occidentale e il polpo. Riflessioni sulla relazione ambigua fra teoria, pratica e animali nella neurobiologia sperimentale*, in *Anatomia del Corpo, Anatomia dell'Anima. Meccanismo, senso, linguaggio*, a cura di A. Trucchio, Quodlibet, Macerata 2008, pp. 11-84, part. pp. 16-46.

e Wiener, per esempio, appare su una rivista filosofica, come altri importanti contributi degli stessi autori sulla cibernetica e la modellizzazione¹⁴. *The Nature of Explanation* di Kenneth Craik¹⁵ prende le mosse proprio da problemi filosofici (la critica dell'apriorismo e il tentativo di delineare una nuova filosofia empirista). Quasi contemporaneamente, Warren S. McCulloch concepisce una vera e propria epistemologia sperimentale, mirata ad affrontare problemi filosofici tradizionali (come la questione degli universali o il rapporto mente/corpo) secondo una strategia induttiva coerente¹⁶.

L'approccio "sintetico" (basato sull'analisi di strutture complesse, in contrapposizione a quello analitico, che parte dai costituenti elementari¹⁷) è condiviso da scienziati di orientamento e provenienza diversi come uno stile di ragionamento (assimilabile all'*esprit ample* di Duhem), la possibile base di un nuovo metodo sperimentale. Craik si spinge a dichiarare: «Proprio come nella fisica, così in psicologia, fisiologia e filosofia è necessario compiere esperimenti per comprendere chiaramente quali siano i fatti che richiedono spiegazioni, e poi ancora esperimenti per saggiare le spiegazioni ipotetiche proposte»¹⁸. All'interno di questa filosofia sperimentale, i modelli costituiscono il principale strumento epistemico, aprono spazi di comparazione, in base non tanto all'assunto che il cervello e il sistema nervoso siano macchine, ma che svolgano determinate funzioni direttamente comparabili a quelle di alcune macchine.

Questo ci porta alla terza caratteristica suggestiva dei modelli neuroscientifici: la riflessività. La costruzione di modelli meccanici del cervello, della mente o del processo di pensiero inevitabilmente ricomprende l'osservatore e, entro certi termini, lo modifica, in un processo costante di retroazione. In questo senso abbiamo definito le neuroscienze come scienze pratiche in sommo grado. Per il suo oggetto di studio e la sua procedura, la conoscenza neuroscientifica ha un impatto diretto sul concetto di specificità umana, ancor più rilevante se si considerano le possibilità di intervento che questo sapere

¹⁴ Cfr. A. Rosenblueth, N. Wiener and J. Bigelow, *Behavior, purpose and teleology*, in «Philosophy of science», X, 1, 1943, pp. 18-24; A. Rosenblueth and N. Wiener, *The Role of Models in Science*, in «Philosophy of Science», XII, 1945, p. 316.

¹⁵ Cfr. K. J. W. Craik, *The nature of explanation*, Cambridge University Press, Cambridge 1967², cap. 1.

¹⁶ Cfr. L. E. Kay, *From logical neurons to poetic embodiments of mind: Warren S. McCulloch's project in neuroscience*, in «Science in Context», XIV, 4, 2002, pp. 591-614; L. E. Kay, *Who wrote the Book of Life? A History of the Genetic Code*, Stanford University Press, Stanford (CA) 2000, cap. 3; T. H. Abraham, *(Physio) logical circuits: The intellectual origins of the McCulloch-Pitts neural networks*, in «Journal of the History of the Behavioral Sciences», XXXVIII, 1, 2002, pp. 3-25.

¹⁷ K. J. W. Craik, *The nature of explanation*, cit., p. 109

¹⁸ *Ibid.*, p. 100.

promette. È nelle neuroscienze, più che in qualsiasi altro ambito di ricerca, che l'interazione e la confusione tra il livello dei fatti e quello dei valori è non solo sempre presente, ma spesso essenziale.

4. Modelli, macchine, cervelli

Roberto Cordeschi ha osservato come la definizione cibernetica di "modello" rappresenti una dipartita sostanziale dalla tradizionale connotazione del termine¹⁹. Non è nella rappresentazione di un fenomeno o di una struttura che il modello trova la propria giustificazione, ma nell'incorporazione di una serie di regole ch'esso condivide con il fenomeno di riferimento. Questa differenza è fondamentale, in quanto riconosce alla pratica della modellizzazione un reale valore euristico: osservando il modello (teorico o reale) in azione, è possibile trarre conclusioni sul fenomeno corrispondente. Questo perché entrambi gli elementi del binomio svolgono la medesima funzione e fanno riferimento al medesimo insieme di assiomi e vincoli. Una società o un'impresa possono essere definite come organismi, previa precisazione dei termini significativi del confronto. Il modello assume così una propria identità, oltre la semplice analogia. Se, infatti, l'analogia tra il cuore e una pompa idraulica ci appare perfettamente legittima e produttiva, in quanto fa emergere una caratteristica essenziale del cuore, né Harvey né alcun altro in seguito avrebbe mai concepito un'analogia inversa. Il pensiero di affrontare le macchie idrauliche come fossero dei cuori ci sembra giustamente assurdo, oltre che piuttosto inutile. I modelli della cibernetica, in special modo i modelli meccanici del cervello, offrono invece proprio questa opportunità. «Il problema dell'interpretazione della natura e varietà della memoria nell'animale», ci ricorda Norbert Wiener, «ha un parallelo nel problema di costruire memorie artificiali per la macchina»²⁰. La biunivocità del parallelismo è la chiave d'interpretazione del modello, che apre possibilità nuove di rappresentazione e sperimentazione di entrambi i termini del binomio. I cervelli sono una classe speciale di macchine, e alcune macchine possono essere accostate a cervelli. La ragione del parallelismo risiede nella selezione precisa e giustificata delle

¹⁹ Cfr. R. Cordeschi, *The Discovery of the Artificial: Behavior, Mind, and Machines Before and Beyond Cybernetics*, cit.

²⁰ N. Wiener, *Cybernetics: Control and communication in the animal and the machine*, MIT Press, Cambridge (MA) 1948, p. 14.

caratteristiche salienti (in questo caso il comportamento predittivo, come caso speciale dell'adattamento) e dei vincoli comuni ai due sistemi, quindi nel rimando a un livello di realtà non ulteriore ma condiviso, che è quello del meccanismo. Non c'è un *tertium*, ma solo una relazione di disvelamento reciproco fra i due termini del binomio.

Nell'orizzonte di una scienza pratica, la monodimensionalità del rapporto teoria-esperienza, della relazione di verità e di verifica fra due livelli di realtà gerarchicamente organizzati, è sostituita da un livello unico ma multi-dimensionale. La ragion d'essere d'un modello, il suo criterio di adeguatezza, è che funziona. Che un modello funzioni significa:

1. Che l'approccio meccanicistico è adeguato alla spiegazione di tutti gli elementi della relazione (non vi sono salti ontologici fra naturale e artificiale, solo differenze di percezione – questa è la sua dimensione critica);
2. Che il modello stesso non si limita a riprodurre fedelmente fenomeni ma ne produce (dimensione euristica);
3. Che il modello meccanico può essere convertito a qualche impiego, è in grado di fare qualcosa di comprensibile e sensato (dimensione performativa).

La giustificazione in base al funzionamento è cruciale, poiché lo stesso processo (empirico) di definizione del criterio, cioè la stessa decisione se un modello funzioni o no, ha un effetto retroattivo sull'oggetto o la funzione riprodotti (il cervello, il "pensiero").

Un esempio di questa riverberazione fra livelli di realtà è l'esperimento mentale sulla capacità di pensiero delle macchine, il "test di Turing"²¹. La domanda fondamentale, «le macchine possono pensare?» è qui immediatamente rimpiazzata da una riformulazione apparentemente analoga e più accessibile all'esperienza: «è concepibile che una macchina riesca a ingannare un soggetto umano in un gioco d'imitazione?». Il criterio di successo, in questo caso, è lo sviamento di un osservatore non onnisciente. La possibilità di rinvenire processi di pensiero in una macchina è identificata con la sua capacità d'ingannare. La messa a punto di questo esperimento implica, oltre alla negoziazione delle caratteristiche funzionali di una macchina "pensante", la riconsiderazione di presupposti e pregiudizi sulla distinzione fra umano e artificiale. Nell'esperimento di Turing, l'unico modo per giungere alla inequivocabile certezza che la macchina possiede dei veri e propri stati di coscienza sarebbe essere la macchina stessa, il che non è

²¹ Cfr. A.M. Turing, *Computing machinery and intelligence*, in «Mind», LIX, 236, 1950, pp. 433-460.

possibile, né comunque sarebbe comunicabile. L'unica possibilità è l'inferenza in base all'osservazione del flusso degli stati nella macchina. Qui emerge chiara la dimensione critica del modello meccanico: in uno spazio relazionale non v'è luogo per la presupposizione di un'essenza. Il dicibile è limitato da ciò che si può esperire. È ciò che la macchina fa che la qualifica come essere pensante entro precisi limiti, e il suo essere pensante non è una qualità intrinseca a priori, ma un risultato dell'osservazione. Nell'interazione con il modello meccanico, il solipsismo filosofico, l'astrazione dell'osservatore dal fenomeno osservato non è un'opzione. Siamo nel regno della contingenza, in cui le convinzioni dello sperimentatore hanno conseguenze ontologiche. La macchina sta qui per un'alterità che può o meno essere accettata, presupposta nell'interazione, ma mai com-presa dall'attore umano. Al contrario, il confronto con la macchina "pensante", in quanto situazione limite, fa emergere le incongruenze di un approccio essenzialistico o biologistico alla coscienza come, per esempio, il tentativo di localizzarla in qualche zona del cervello o rete di cellule²².

Il gioco dell'imitazione si fonda, in sensi diversi, sull'esplorazione del limite: il limite della simulabilità di caratteristiche "propriamente umane" come quello di sostenibilità di una serie di assunti considerati fondamentali, tanto a favore della comparazione cervello-macchina quanto contro. Né l'osservatore né l'oggetto osservato (non la macchina, non la sua *performance*) possono essere considerati assoluti. La dimensione esplorativa del modello e la sua relazione con il limite traspare dalla definizione stessa delle macchine di Turing in quanto meccanismi *possibili*, cioè in teoria realizzabili e, ancor più, dagli svariati esempi di traduzione della macchina di Turing universale in macchine particolari, contingenti, reali. Trasposto nella realtà della costruzione di artefatti, il gioco dell'imitazione mostra tutta la sua portata critica e, al contempo, permette di apprezzare le altre due dimensioni su considerate, l'epistemica e la performativa. La costruzione, o anche la semplice progettazione (realizzazione grafica) dell'algoritmo che sta per una macchina possibile richiede la definitiva esplicitazione e chiarificazione del rapporto tra soggetto osservante e fenomeno osservato: l'algoritmo esce dallo spazio immaginativo e si fa oggetto d'esperienza dei sensi, oltre che del puro spirito. Questo passaggio implica il rispetto di alcune ulteriori regole e vincoli. La regola fondamentale della modellizzazione è che il fenomeno da imitare (p.e. il comportamento) sia descritto «in termini di reazioni

²² *Ibid.*

uniche e precisamente definibili a situazioni uniche e precisamente definibili»²³, cioè in modo non equivoco. Al vincolo dell'univocità si aggiunge ovviamente quello della realizzabilità, anche solo in linea di principio (in questo senso il modello può anche trascendere la propria natura economica in senso stretto: si possono concepire modelli semplicemente troppo costosi per essere costruiti) e, infine, quello della contingenza. Il modello meccanico nasce come esternalizzazione di caratteristiche funzionali tipiche del cervello o dell'organismo in relazione con l'ambiente, una operazione al contempo «selettiva e proiettiva»²⁴ che ha come implicazione immediata l'accettazione dell'impossibilità di dedurre il comportamento imitativo a partire dalla struttura. La specifica costituzione di un modello meccanico non rappresenta che uno fra i tanti modi possibili di replicare una funzione. Appare qui la seconda dimensione del modello, quella epistemica. Da un lato, se interpretata in rapporto alla strategia sperimentale tipica della modellizzazione, l'impossibilità di dedurre la funzione dalla struttura è indice di una precisa opzione per lo studio dei sistemi, delle organizzazioni rispetto a quello delle unità costitutive, dei componenti. Dall'altro, se riportata al gioco contingente della relazione fra osservatore e osservato, essa ha come implicazione immediata l'impossibilità per l'osservatore di dedurre tutte le proprietà dell'osservato, sebbene l'osservato sia una sua "creazione". Una via sperimentale per conciliare teismo ed evoluzione, se vogliamo. Ma ciò che qui più c'interessa è la sostanziale interconnessione fra le tre dimensioni del modello meccanico. Esso, abbiamo visto, è una soluzione contingente ed economica a un problema di rappresentazione. In quanto tale, deve essere semplice, plausibile, realizzabile. Ma è anche uno strumento di esplorazione o, meglio, una ipotesi esplorativa. Nella meccanizzazione di funzioni e fenomeni apertamente negoziati o «precisamente definibili», ma pur sempre complessi e non deducibili a priori, si apre infatti lo spazio dell'emergenza, della sorpresa. Già McCulloch e Pitts, nel modellare una rete di neuroni ideali come una macchina di Turing, avevano osservato nel 1943 la possibilità di comportamenti patologici non esplicitamente ricercati, di creazione di circuiti perversi che non può essere imputata a errori di progettazione. Fra le possibilità di azione e sviluppo delle reti di neuroni logici è semplicemente presente la loro

²³ D. M. MacKay, *Mindlike Behaviour in Artefacts*, in «The British Journal for the Philosophy of Science», II, 6, 1951, pp. 105-117, qui p. 109.

²⁴ *Ibid.*

degenerazione patologica, fenomeno che apre un nuovo spazio di comparazione e rappresentazione. Il modello, semplicemente, può incorporare la disfunzione allo stesso modo in cui incorpora la funzione. In questo senso affermo che un modello può produrre realtà, in un senso già spiegato, cioè nello spazio della relazione sperimentale.

Ancora più evidenti sono gli esempi di macchine concretamente realizzate. Si pensi, per esempio, alle "tartarughe" costruite alla fine degli anni Quaranta dal neurologo inglese William Grey Walter, o al celebre "omeostata", il meccanismo autostabilizzante di William Ross Ashby (psichiatra)²⁵. Nella loro visibile differenza (le tartarughe sono macchine semoventi, dotate di sensori ed effettori che le fanno interagire con il mondo circostante; l'omeostata può essere interpretato come un modello della relazione di sistema nervoso e ambiente) queste macchine sono comparabili al livello funzionale (nella loro costante ricerca di equilibrio fra i vari elementi in relazione) e fenomenico. Costruite secondo il criterio della massima economia (minima quantità di elementi) esse avevano la funzione critica di mostrare come il comportamento complesso e apparentemente irriducibile degli organismi adattativi fosse in realtà riproducibile da sistemi molto semplici. Entrambi i modelli hanno avuto successo ben oltre le aspettative, producendo comportamenti inattesi (le "relazioni affettive" fra tartarughe, la capacità di recupero dell'omeostata dalla modificazione casuale dei suoi circuiti) e, con l'aggiunta di piccole modifiche o componenti (principalmente di un'unità di memoria), suggerendo addirittura una possibile rappresentazione dei diversi stadi dell'evoluzione comportamentale. Fra il serio e il faceto, Grey Walter usava per i diversi modelli una tassonomia linneana (*Machina speculatrix*, *Machina docilis*, o *Machina sopora*, il nome da lui inventato per l'omeostata di Ashby). Su questa falsariga, in un testo fondamentale dell'Intelligenza Artificiale, Valentino Braitenberg ha costruito una vera e propria storia naturale delle macchine, fondata sulla progressiva duplicazione e moltiplicazione di pochi elementi strutturali²⁶.

Dagli esempi citati appare chiara la relazione strettissima fra le dimensioni euristica e critica dei modelli meccanici: il fatto che macchine determinate, finite, possano mostrare caratteristiche emergenti sembra corroborare la validità dell'approccio meccanico e

²⁵ Cfr. W. G. Walter, *The living brain*, Duckworth, London 1953; W. R. Ashby, *Design for a Brain*, Chapman and Hall, London 1960. Cfr. anche A. Pickering, *The Cybernetic Brain: Sketches of Another Future*, University Of Chicago Press, Chicago 2010.

²⁶ Cfr. V. Braitenberg, *Vehicles: Experiments in synthetic psychology*, The MIT Press, Cambridge (MA) 1986.

riduzionistico a funzioni complesse e apparentemente qualitative. Come abbiamo anticipato, è la stessa distinzione fra naturale e artificiale che è revocata in dubbio da questi esperimenti sintetici, e ciò deriva proprio dalla capacità degli automi di ingannare i nostri sensi e sfidare assunti pregiudiziali radicati. Braitenberg ha tradotto questa capacità in una strategia argomentativa: dopo aver descritto la struttura e la performance delle macchine, alla fine di ogni capitolo pone la medesima domanda: se non sapeste che questo è un artefatto, non lo scambiereste per un animale aggressivo, o amichevole, o spaventato?²⁷ Nel gioco della relazione contingente, questa domanda ha perfettamente senso. Tocca a noi giustificare la nostra distinzione pregiudiziale fra naturale e artificiale o, come ha suggerito fra gli altri Herbert Simon, liberarcene completamente²⁸. In questa luce, l'affermazione di partenza, che il modello generi realtà, nella forma di nuove relazioni dinamiche e contingenti fra essere, rappresentazione e dover essere acquista plausibilità. L'ultima dimensione, la performativa, non fa che confermare il nostro assunto. Le macchine su descritte, come altre analoghe o più complesse (si pensi al *Perceptron* di Rosenblatt o alla *Pattern Recognition Machine* di Taylor²⁹) sono modelli validi in quanto realizzabili, ed effettivamente realizzati. Non solo, in tutti i casi citati la finalità sperimentale e/o ludica è sempre affiancata da una idea più o meno chiara delle possibili applicazioni pratiche dei modelli. La dimensione applicativa, coerentemente con il concetto pratico di scienza introdotto all'inizio, non è conseguenza delle altre due. Al contrario, non è possibile distinguerla completamente, tanto a priori, quanto a posteriori. La trasfigurazione dei modelli da cose epistemiche a oggetti tecnici presenta anch'essa una tipica funzione riflessiva: l'applicazione sostiene la plausibilità e, a un diverso livello, agisce o promette di agire sul rappresentato, sia nello spazio della rappresentazione che in quello della sostituzione o potenziamento. Ancora, siamo qui in una dimensione completamente contingente, in cui la "danza degli agenti" risulta in relazioni epistemiche e ontologiche mai definite.

Questo c'introduce alle considerazioni finali, in primo luogo alle virtù che un simile coinvolgimento dell'osservatore nel fenomeno richiede.

²⁷ *Ibid.*

²⁸ Cfr. H. Simon, *The Sciences of the Artificial.*, MIT Press, Cambridge (MA) 1969, pp. 6-9.

²⁹ Cfr. F. Rosenblatt, *The perceptron: A probabilistic model for information storage and organization in the brain*, in «Psychological review», LXV, 6, 1958, pp. 386-408; D. Gabor, *Models in Cybernetics*, ne «I Modelli nella Tecnica», Accademia Nazionale dei Lincei, Roma 1956.

5. Il Criterio, l'Icona e la Virtù

In questo saggio, ho tentato di esplorare e rendere esplicita la dimensione contingente della modellizzazione in quanto rapporto non gerarchico fra oggetto, soggetto osservante e relazione sperimentale. Il coinvolgimento esplicito dell'osservatore nello spazio della rappresentazione, pur inevitabile e potenzialmente produttivo, sottolinea il lato rischioso del ragionar per modelli e, credo, suggerisce la possibilità, sempre presente, che questo gioco possa degenerare. L'assenza di un criterio trascendente e la parallela necessità di esplicitare per quanto possibile tutti gli elementi e criteri della relazione evidenzia il ruolo dell'arbitrarietà delle scelte iniziali. Lo sviluppo di un modello del comportamento, per quanto autonomo, dipende criticamente dalla scelta delle variabili iniziali e, nella sua natura relazionale, anche dalla precisa coscienza della sua funzione economica e "utilitaristica". Dipende, in sostanza, da una certa disciplina che l'osservatore deve darsi nel gioco, o, in sintesi, dalla virtù epistemica della consapevolezza del limite³⁰.

Laddove questa consapevolezza venga meno, si rischia di passare dal gioco del "come se" a quello dell'"è", dal livello dell'utilità a quello della verità.

Per esempio, come è accaduto nelle scienze cognitive, il meccanismo comune può trasformarsi, da modello, in criterio: dalla relazione biunivoca tra cervelli e macchine si può passare alla considerazione di uno dei due termini come misura assoluta³¹. Le operazioni del cervello si intendono come casi speciali di computazione. La mente da spiegare e la mente dell'osservatore acquistano lo stesso livello di realtà, e il modello meccanico diviene principio, e non luogo, di spiegazione³².

Un'altra possibile degenerazione s'intravede nelle conseguenze delle tecniche di visualizzazione dell'attività cerebrale, come la PET o l'fMRI. Qui, la potenza dell'immagine rischia di giocare brutti tiri.

³⁰ Sulle virtù epistemiche cfr. l'introduzione a P. Galison, L. Daston, *Objectivity*, MIT Press, Boston 2007.

³¹ G. A. Miller, E. Galanter and K. H. Pribram, *Plans and the structure of behavior*, Holt, Rinehart and Winston, New York 1960 è un esempio di questa confusione. Cfr. Anche J. Dumit. *Circuits in the Brain and How they Got There*, relazione al convegno «Our Brains-Our Selves?», Harvard University, Boston (MA), 1-3 maggio 2008.

³² Cfr. M. Polanyi, *The Hypothesis of Cybernetics*, in «The British Journal for the Philosophy of Science», II, 8, 1952, pp. 312-318.

Contrariamente al modello, l'immagine del cervello in azione è una scatola nera (nel senso di Bruno Latour³³): ci sembra di osservare in essa un fenomeno quale esso è, un oggetto naturale (il cervello) colto nell'esercizio delle sue funzioni. La giustificazione della relazione, la necessità di chiare definizioni e chiare regole sembra passare in secondo piano. È vero, la mediazione di una macchina non è negabile, ma essa ci mostra il cervello in azione. Ha una funzione ancillare, ci aiuta a vedere, non partecipa alla costruzione.

Da oggetti mondani che stanno per il cervello si passa al cervello come modello di un mondo, di un sistema di relazioni biologiche, epistemiche e tecnologiche che sono al contempo espresse e oscurate nell'immagine dinamica e colorata di un atto di percezione, di memoria. Il cervello in se stesso, colto in tempo reale nella propria attività specifica, sembra aver perso ogni riferimento all'intrinseca mediatezza dei modelli, al loro carattere economico e al loro stato di equilibrio fra istanze teoretiche e pratiche diverse. Eppure, questo nuovo e più alto livello di sintesi, che non ha mancato di suscitare la preoccupata attenzione di filosofi, storici e neuroscienziati "critici"³⁴, a ben guardare non si discosta poi tanto nei suoi caratteri costitutivi dai modelli "mediati" cui s'è accennato sopra. Anche in questo caso, la costruzione dell'immagine emerge da un sostrato tecnologico non neutro, da un sistema di ipotesi di funzionamento e dati di struttura e da una serie di processi di mediazione. Il risultato, tuttavia, appare trasparente, immediato. La virtù epistemica è incarnata nelle virtù fisiologiche del meccanismo che è pur sempre dentro l'osservatore, ma che al contempo lo trascende e lo giustifica in quanto tale.

FABIO DE SIO svolge attività di ricerca di storia della scienza alla Stazione zoologica Anthon Dohrn di Napoli. La sua ricerca è sostenuta da una borsa del Wellcome Trust of London

³³ Cfr. B. Latour, *Science in Action*, Harvard University Press, Boston 1987.

³⁴ Cfr. F. Ortega, F. Vidal, *The Cerebral Subject. Some Ideas for a Book Project*, presentazione all'MPIWG Departmental Colloquium, Berlino, 09.01.07; Id., *Mapping the cerebral subject in contemporary culture*, in «RECIIS», I, 2, Jul-Dec 2007, pp. 255-259; C. Borck, *Toys are Us. Models and Metaphors in Brain Research*, in *Critical Neuroscience* a cura di S. Choudhury, J. Slaby, Blackwell, Chichester (in corso di pubblicazione).