

EDOARDO DATTERI

THE EPISTEMIC ROLES OF AUTOMATA FROM CYBERNETICS TO CONTEMPORARY ROBOTICS

I RUOLI EPISTEMICI DEGLI AUTOMI,
DALLA CIBERNETICA ALLA ROBOTICA CONTEMPORANEA

Automata, or robots, have often played a crucial role in the reflection on, and discovery of, the mechanisms underlying the behaviour of living systems. During the first decades of the XX century, automata were invoked to support mechanistic theories about animal behaviour. More recently, robotic models of living systems have been used for predictive and explanatory purposes. This article explores these epistemic uses of robots with historical and contemporary examples.

I. INTRODUZIONE

La costruzione di automi che riproducono le sembianze e il comportamento di esseri umani o animali non umani – automi che qui chiameremo ‘biomorfi’ – ha spesso svolto un ruolo di rilievo nella teorizzazione sui meccanismi che regolano il comportamento dei sistemi viventi e nella riflessione su ciò che accomuna e distingue animali e macchine. All’analisi del ruolo conoscitivo (o epistemico) degli automi biomorfi è dedicato questo contributo. Metteremo in evidenza, anche attraverso esempi, come la costruzione di automi abbia spesso fornito sostegno a posizioni meccanicistiche sulla natura del comportamento animale, e abbia contribuito alla previsione e spiegazione del comportamento di particolari specie viventi.

Questo è naturalmente solo uno dei ruoli che gli automi, o robot, biomorfi sono stati finora chiamati a svolgere nella scienza e nella società contemporanea. All’inizio degli anni ’90 del secolo scorso il robotico Rodney Brooks ha riportato all’attenzione di robotici e studiosi di Intelligenza Artificiale il fatto che, in alcuni casi, lo studio dei meccanismi neurali o psicologici che

stanno alla base del comportamento animale può fornire utili spunti per progettare robot efficienti: è la cosiddetta *biologically inspired robotics*.¹ Un esempio contemporaneo di *biologically inspired robotics* riguarda lo sviluppo di robot in grado di afferrare oggetti di varia forma e dimensione, capacità naturalmente cruciale per una gran varietà di applicazioni pratiche. L'idea esplorata da alcuni importanti gruppi di ricerca è quella di prendere spunto dall'agilità e dalla destrezza con cui i polpi, usando i loro tentacoli flessibili e sinuosi, riescono ad afferrare vari tipi di oggetti adattandosi alla loro forma e ottenendo dunque una notevole stabilità di presa. Sulla base di questa idea sono stati realizzati robot dotati di bracci flessibili, tentacolari, che funzionano sia come 'zampe' per la locomozione sui fondali marini sia come organi di presa, proprio come i polpi.² Anche la scelta di conferire ai robot un aspetto umanoide è spesso imposta da motivazioni di carattere pratico: può per esempio essere dettata dall'obiettivo di generare forme di empatia nei confronti della macchina, a sua volta importante per assicurare una migliore interazione uomo-robot in certi campi di applicazione.³

Risalendo il corso del tempo fino a epoche più remote troviamo molti esempi di automi biomorfi la cui costruzione, documentata o narrata in fonti più o meno veritiere, era dettata da motivazioni di carattere non conoscitivo. Fonti medievali come le *chansons de geste*, per esempio, sono colme di riferimenti ad automi che assumevano sembianze umane e animali. Pur intrattenendo uno stretto rapporto con il mondo vivente, emulandone la forma e il comportamento, essi non erano pensati per finalità conoscitive: il loro ruolo nella narrazione era per esempio quello di enfatizzare la superiorità tecnologica di popolazioni lontane e pagane (saraceni, slavi, turchi, arabi, persiani) spesso attribuita a conoscenze esoteriche e demoniache, marcando la distinzione tra l'occidente cristiano e il resto del mondo non cristiano.⁴ Nel corso della storia si conferisce aspetto umano o animale alle macchine per intrattenere, per stupire, per narrare.

In altri casi – quelli che stiamo per discutere – l'emulazione meccanica del vivente è invece motivata da un deliberato intento di provocare una riflessione, o addirittura di stimolare lo sviluppo di ipotesi scientifiche, sul comportamento animale. Più precisamente, nei casi che discuteremo la costruzione e l'osservazione di automi fornisce basi empiriche per giungere a conclusioni di interesse scientifico sul comportamento di classi più o meno genericamente delimitate di sistemi viventi. Naturalmente, come del resto accade in qualunque area di ricerca scientifica, la validità di tali studi riposa su un numero molto ampio di assunzioni metodologiche talvolta

¹ R. PFEIFER, M. LUNGARELLA, F. IIDA, *Self-Organization, Embodiment, and Biologically Inspired Robotics*, in «Science», CCCXVIII, 2009, pp. 1088-1093.

² M. CIANCHETTI, M. CALISTI, L. MARGHERI, M. KUBA, C. LASCHI, *Bioinspired Locomotion and Grasping in Water: The Soft Eight-Arm OCTOPUS Robot*, in «Bioinspiration and Biomimetics», X, 2015, 3. DOI:10.1088/1748-3190/10/3/035003.

³ P. DUMOUCHEL, L. DAMIANO, *Living with Robots*, Harvard University Press, Cambridge, MA 2017.

⁴ E.R. TRUITT, *Medieval Robots. Mechanism, Magic, Nature, and Art*, University of Pennsylvania Press, Philadelphia, PA 2016.

controverse, talvolta implicite, quasi sempre di non facile analisi. In queste pagine ci limiteremo a fornire una panoramica degli usi epistemici degli automi cercando un bilanciamento tra completezza, sistematicità e sintesi, rimandando ad altri testi pubblicati in letteratura⁵ per una più approfondita discussione dei relativi problemi di metodo scientifico.

II. AUTOMI E MECCANICISMO

Come ha estesamente discusso il filosofo della scienza italiano Roberto Cordeschi nel volume *The discovery of the artificial*,⁶ nelle prime decadi del XX secolo la costruzione di macchine biomorfe ha alimentato, nella comunità di ricerca scientifica e filosofica, l'ipotesi che alcune forme di comportamento fino ad allora ritenute esclusiva prerogativa dei sistemi viventi potessero essere spiegate in termini meccanicistici, in direzione dunque contraria alle posizioni vitaliste allora diffuse.⁷ Più in particolare – ancor prima dello sviluppo della cibernetica, dei contributi di Alan Turing alla riflessione sui limiti e le potenzialità delle macchine calcolatrici, e della effettiva costruzione di calcolatori digitali – si rafforza l'idea secondo cui le macchine possono generare comportamenti imprevedibili, apprendere, manifestare comportamenti orientati a uno scopo, elaborare informazioni contenute nella propria memoria. Al rafforzarsi di tali posizioni contribuisce la costruzione di automi che non erano destinati, come prima modalità d'uso, ad alcun impiego di carattere 'pratico': si trattava di macchine appositamente costruite come strumenti per favorire una riflessione sulla distinzione tra sistemi viventi e sistemi meccanici.

Iniziamo dall'opposizione, spesso tracciata nella storia del pensiero, tra sistemi meccanici e sistemi imprevedibili. È ben nota la posizione cartesiana secondo cui è in linea di principio possibile costruire un automa che replichi fedelmente il comportamento degli animali e alcuni aspetti del comportamento dei sistemi viventi: tale automa potrebbe addirittura proferire parole, intese come reazioni stereotipate a stimoli di un certo tipo (toccandola in un punto potrà per esempio gridare che sente dolore). Cartesio tuttavia afferma l'impossibilità di costruire una macchina che coordini diversamente le parole per rispondere al senso di ciò che si dirà in sua presenza. L'automa, inoltre, disporrà al più di un insieme limitato di disposizioni all'azione, po-

⁵ E. DATTERI, *Filosofia delle scienze cognitive. Spiegazione, previsione, simulazione*, Roma, Carocci, 2012.

⁶ R. CORDESCI, *The Discovery of the Artificial. Behavior, Mind and Machines Before and Beyond Cybernetics*, Springer Netherlands, Dordrecht, 2002; E. DATTERI, *Intelligenza Artificiale e Scienza Della Natura*, in P. PECERE (ed.), *Il Libro Della Natura. II. Scienze e Filosofia Da Einstein Alle Neuroscienze Contemporanee*, Carocci, Roma 2015, pp. 180-208.

⁷ G. ALLEN, *Mechanism, Vitalism and Organicism in Late Nineteenth and Twentieth-Century Biology: The Importance of Historical Context*, in «Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences», XXXVI, 2005, 2, pp. 261-283.

tendo dunque agire (o meglio reagire) in un numero limitato di circostanze. Tali limitazioni, afferma Cartesio, non affliggono gli esseri umani, in grado di reagire opportunamente in tutte le occorrenze della vita.⁸

La distinzione tra esseri umani e animali non umani tracciata da Cartesio attribuisce rigidità e invariabilità al comportamento delle macchine (e degli animali), una posizione condivisa alla fine del diciannovesimo secolo, tra gli altri, da Charles Lloyd Morgan⁹ nelle sue riflessioni sull'automatismo nel mondo animale. Questa posizione viene messa in discussione, anche attraverso la costruzione di automi, nei primi decenni del XX secolo.

Immaginiamo il comportamento di un automa la cui velocità di movimento sia direttamente proporzionale all'intensità della luce che colpisce un sensore fotoelettrico (automi di questo genere sono stati descritti dal neuroscienziato Valentino Braitenberg in un testo di riferimento per le scienze cognitive contemporanee).¹⁰ In ambienti come quelli in cui ordinariamente viviamo sarebbe per noi impossibile prevedere l'esatta quantità di luce che il suo sensore percepirà istante per istante: non potremmo quindi prevedere la velocità del suo movimento, nonostante la relazione tra intensità luminosa e velocità rimanga immutata nel corso del tempo. Inserendo un numero maggiore di sensori e complicando anche di poco il modo in cui essi determinano le reazioni dell'automata, senza violare in alcun modo l'automatismo del meccanismo interno, si ottiene facilmente un automa il cui comportamento è del tutto imprevedibile per l'occhio umano.¹¹

Questa intuizione è stata discussa dal pioniere dell'elettroencefalografia William Grey Walter in un articolo pubblicato su *Scientific American* nel 1950.¹² Grey Walter è noto per avere costruito alcuni automi o 'tartarughe' biomorfe della specie da lui chiamata *Machina speculatrix*. Le

⁸ R. CARTESIO, *Discorso sul Metodo*, in R. CARTESIO, *Opere filosofiche 1*, a cura di E. GARIN, Laterza, Bari 1996, pp. 289-350; R. CARTESIO, *L'Uomo*, ivi, pp. 203-288.

⁹ C.L. MORGAN, *Animal Automatism and Consciousness*, in «The Monist», VII, 1896, 1, pp. 1-18.

¹⁰ V. BRAITENBERG, *Vehicles. Experiments in Synthetic Psychology*, The MIT Press, Cambridge, MA 1986.

¹¹ Sottolineiamo la precisazione "per l'occhio umano". La descrizione qui fornita non permette di inferire che il comportamento dell'automata è in linea di principio imprevedibile. Conoscendo infatti l'esatta quantità di luce che il suo sensore percepisce istante per istante sarebbe sicuramente possibile, a meno di malfunzionamenti interni, prevedere accuratamente la velocità del sistema. La considerazione che stiamo sviluppando, e che emerge chiaramente dal testo di Braitenberg, riposa sul fatto che, nello scenario che stiamo ipotizzando, date le normotipiche capacità cognitive e sensoriali degli esseri umani, è praticamente impossibile risalire a occhio nudo (dunque senza l'aiuto di particolari strumenti di misura) all'intensità della luce che colpisce il sensore dell'automata istante per istante. In breve, agli occhi di una persona che osserva l'automata muoversi in un ambiente ordinario, il comportamento dell'automata appare imprevedibile.

¹² W. GREY WALTER, *An Imitation of Life*, in «Scientific American», CLXXXII, 1950, 5, pp. 42-45.

tartarughe sensorizzate di Grey Walter erano in grado di muoversi nella casa del suo costruttore evitando gli ostacoli, dirigendosi verso fonti di luce (ma allontanandosene se si avvicinavano troppo) e dirigendosi autonomamente verso le fonti di ricarica. Nell'articolo citato, Grey Walter sottolinea come basti collegare pochi sensori e motori tra di loro per ottenere una 'strana ricchezza' di comportamento che egli caratterizza in termini di incertezza, casualità e addirittura libero arbitrio: «Solo pochi elementi» – commenta Grey Walter – «ma densamente connessi tra di loro possono generare praticamente infiniti modi di esistenza».¹³

Un altro motivo di potenziale imprevedibilità degli automi, di cui compaiono prove di esistenza nei primi decenni del XX secolo, riguarda i fenomeni detti di apprendimento meccanico. Nell'automa fotoelettrico che abbiamo immaginato poco sopra, la relazione tra stimoli sensoriali e risposte motorie rimane immutata nel corso del tempo. Nel periodo che stiamo discutendo vengono costruiti circuiti elettrici e dispositivi meccanici in grado di variare il proprio comportamento a parità di stimolo, introducendo dunque un ulteriore livello di flessibilità rispetto a quello dell'automa fotoelettrico o addirittura creando nuove connessioni stimolo-risposta (iniziando cioè a reagire a stimoli che precedentemente non davano luogo ad alcuna reazione). Uno di essi, costruito dallo psicologo Clark Hull e dall'ingegnere Robert Krueger, fu pubblicato nel 1931 nel *Journal of General Psychology*.¹⁴ Era un semplice circuito elettrico che replicava sotto certi aspetti il fenomeno dell'apprendimento condizionato studiato da Pavlov. Inizialmente il circuito manifestava una sola relazione stimolo-risposta. Tuttavia, dopo aver somministrato per un certo tempo quello stimolo assieme a uno stimolo di diversa natura, che inizialmente non evocava risposta da parte della macchina, la sola presentazione del secondo stimolo generava la risposta (la somministrazione contemporanea dei due stimoli immagazzinava in un componente elettronico chiamato 'condensatore' una quantità di carica in grado di attivare la seconda connessione stimolo-risposta). In breve la macchina era in grado, a seguito di una certa esperienza di stimoli, di sviluppare una reazione a un nuovo stimolo. Non vi è nulla di non automatico nella macchina di Hull e Krueger e negli innumerevoli sistemi di apprendimento meccanico sviluppati da allora (che oggi vengono compresi nell'area chiamata *machine learning*). Perché una macchina sia in grado di modificare nel tempo l'entità della propria reazione a un certo stimolo sensoriale è sufficiente che l'entità di tale reazione dipenda da una certa caratteristica fisica del sistema (per esempio, dal livello di carica di un condensatore elettrico), e che un'altra parte del circuito stesso sia in grado di variare il valore di tale caratteristica fisica sulla base, per esempio, della presentazione di un ulteriore stimolo. In breve, è sufficiente che tale sistema possieda almeno due meccanismi stimolo-risposta e che la reazione automaticamente prodotta da uno di essi modifichi la caratteristica fisica che media la relazione automatica tra l'altro stimolo e l'altra risposta.¹⁵

¹³ Ivi, p. 44.

¹⁴ R.G. KRUEGER, C.L. HULL, *An Electro-Chemical Parallel to the Conditioned Reflex*, in «*Journal of General Psychology*», V, 1931, pp. 262-69.

¹⁵ La connessione tra apprendimento e imprevedibilità negli automi meriterebbe una discussione approfondita che esula dagli obiettivi di questo lavoro; vale però la pena fornire alcune veloci precisazioni per suggerire la complessità del problema. In primo luogo, analogamente a quanto

Riassumendo, nei decenni che stiamo considerando la costruzione di macchine e automi fornisce prove di esistenza molto convincenti in merito alla possibilità che certe forme di variabilità e apprendimento, precedentemente ritenute peculiari dei sistemi viventi, non fossero affatto al di sopra del livello dell'automatismo. La ridefinizione del confine tra meccanismi e sistemi viventi ispirata in questi decenni dalla costruzione di automi, o dalla riconsiderazione di meccanismi già ben noti (come il regolatore di Watt, spesso citato nella letteratura d'epoca), riguarda anche la possibilità che le macchine manifestino comportamenti genuinamente orientati verso uno scopo. In quello che oggi è considerato uno dei manifesti della Cibernetica,¹⁶ Rosenblueth, Wiener e Bigelow sostennero che alcuni dispositivi meccanici detti 'a retroazione negativa' possono sensatamente essere considerati come orientati a uno scopo.¹⁷ I dispositivi a retroazione negativa reagiscono automaticamente alla differenza tra uno stato di cose esterno, segnalato da un dispositivo sensoriale (le cui proprietà dipendono cioè da certe caratteristiche ambientali), e uno stato di cose rappresentato internamente dal valore di una certa proprietà fisica del sistema. Un esempio familiare di tali dispositivi è il termostato, che reagisce alla differenza tra la temperatura esterna, segnalata da un sensore, e un valore di riferimento rappresentato

precisato in nota rispetto all'automa fotoelettrico discusso in precedenza, un automa che apprende si qualifica come imprevedibile non in linea di principio ma relativamente alle caratteristiche sensoriali e cognitive di chi lo osserva. In breve, se già 'a occhio nudo' è praticamente impossibile prevedere il comportamento di un automa regolato da meccanismi di coordinazione percezione-azione senza apprendimento e in contesti non particolarmente poveri di stimoli, lo sarà a maggior ragione se la relazione tra percezione e azione cambia nel tempo sulla base, magari, di altre proprietà esterne o interne a loro volta difficili da tracciare nel tempo 'a occhio nudo'. In secondo luogo, l'imprevedibilità è relativa al livello di analisi a cui ci collochiamo nel descrivere il comportamento del sistema. Un automa che apprende a giocare sempre meglio a scacchi in virtù di un meccanismo interno di *machine learning* modificherà nel tempo e in modi praticamente imprevedibili (stante le precisazioni appena formulate) le caratteristiche fini del proprio stile di giocata, ma nondimeno continuerà a giocare a scacchi: modificherà dunque alcuni aspetti, ma non altri, del proprio comportamento.

¹⁶ A. ROSENBLUETH, N. WIENER, J. BIGELOW, *Behavior, Purpose and Teleology*, in «Philosophy of Science», X, 1943, 1, pp. 18-24.

¹⁷ Gli autori riflettono sul fatto che una pistola, in quanto sistema non a retroazione negativa, può dirsi orientato a uno scopo soltanto in senso derivato: contribuisce infatti a perseguire lo scopo dell'essere umano che la usa. I sistemi a retroazione negativa, secondo gli autori, possono dirsi invece intrinsecamente orientati a uno scopo perché, istante per istante, modificano certe loro caratteristiche in funzione della differenza percepita tra la condizione corrente e una condizione di riferimento, come nell'esempio del termostato che stiamo per illustrare. Per una riflessione critica, si veda H. JONAS, *A Critique of Cybernetics*, in «Social Research», XX, 1953, 2, pp. 172-192.

internamente da una certa caratteristica fisica del circuito (per esempio, dalla rotazione di un potenziometro controllato da una manopola). Accendendo il riscaldamento quando tale differenza supera una certa soglia, e spegnendolo quando non la supera, il termostato in un certo senso persegue uno scopo – mantenere costante la temperatura – rimanendo tuttavia un dispositivo totalmente automatico.

Un esempio di automa orientato a uno scopo fu costruito attorno al 1915 dagli ingegneri John Hammond Jr. e Benjamin Miessner.¹⁸ Era un piccolo veicolo fornito di ruote e di due sensori fotoelettrici montati frontalmente: quando la luce colpiva uno dei due sensori, un meccanismo elettromeccanico sterzava i motori in modo da dirigere il veicolo dal lato colpito dalla luce. Vi è un senso piuttosto forte in cui tale automa, pur essendo appunto totalmente meccanico, perseguiva un obiettivo – raggiungere le fonti di luce. La fonte di ispirazione per la costruzione dell'automata derivava dal lavoro del neurofisiologo Jacques Loeb che, nel 1900, aveva proposto teorie meccanicistiche a proposito del comportamento di certi insetti, tra cui le falene.¹⁹ Lo stesso Loeb affermò che la costruzione dell'automata forniva supporto all'idea che alcuni aspetti del comportamento animale – come certe forme di orientamento verso fonti di stimoli sensoriali – potessero essere spiegate in termini puramente meccanicistici.

Nel suo orientarsi verso fonti di luce, l'automata di Hammond e Miessner metteva in atto una pronta reazione al presentarsi di certi stimoli sensoriali. Non si può dire che avvenisse al suo interno alcuna elaborazione di tale stimolo che andasse oltre l'immediata produzione di tale risposta, a meno di non voler estendere molto il significato che ordinariamente diamo a tale termine. L'idea che le macchine possano memorizzare informazioni sull'ambiente e manipolare il contenuto della propria memoria per reagire in modo opportuno allo stato di cose esterne – capacità che oggi diamo per scontata – si sviluppa negli stessi anni anche a seguito della costruzione di macchine e automi in grado di manifestare forme rudimentali di memoria meccanica e di elaborazione delle informazioni.

Da un certo punto di vista già il circuito di Hull e Krueger descritto sopra era in grado di memorizzare una nuova connessione stimolo-risposta: il 'sito' della sua memoria consisteva come abbiamo sottolineato in un condensatore, ovvero un componente elettronico in grado di mantenere una certa carica elettrica entro certe condizioni. In un articolo pubblicato nel 1933 su *Scientific American* Thomas Ross²⁰ discusse una memoria elettrica che «dimostra, entro certi limiti, gli elementi di base dell'attività mentale»: alla somministrazione di uno stimolo elettrico un semplice circuito attivava un magnete che spostava una piccola sfera metallica da un punto all'altro del circuito. Una volta giunta nella nuova posizione la sfera chiudeva un altro circuito elettrico che prima era aperto, permettendo da quel momento in poi il passaggio di elettricità attraverso di esso. La somministrazione del primo stimolo aveva dunque generato una modifica permanente delle proprietà elettriche del circuito, che da quel momento in poi conservava la

¹⁸ Cfr. CORDESCHI, *The Discovery of the Artificial*, cit.

¹⁹ J. LOEB, *Comparative Physiology of the Brain and Comparative Psychology*, G.P. Putman's Sons, New York, 1900.

²⁰ T. ROSS, *Machines That Think*, in «Scientific American», LXXXIX, 1933, pp. 206-208.

traccia mnestica dello stimolo ricevuto. Lo stesso Ross costruì due automi in grado di memorizzare la struttura di semplici percorsi. Una delle due veniva collocata su un percorso ricco di bivi che conducevano a vicoli ciechi; durante la prima navigazione l'automa 'sbagliava' entrando nei vicoli ciechi, ma tali 'errori' venivano registrati in una memoria meccanica che permettevano all'automa, al secondo giro, di compiere l'intero percorso evitando i bivi.

III. DAL MECCANICISMO ALLE TEORIE MECCANICISTICHE: AUTOMI COME MODELLI DI SISTEMA VIVENTE

Nei casi discussi finora, la costruzione di automi ha fornito prove costruttive di esistenza a sostegno dell'idea che certi aspetti del comportamento animale, fino a quel momento ritenuti prerogativa dei sistemi viventi, potessero essere spiegati in termini meccanicistici. Gli automi possono però contribuire alla comprensione del comportamento vivente in modi più specifici: alcuni automi costruiti nell'epoca cibernetica, e molti automi costruiti in tempi più recenti, intrattengono un rapporto molto stretto con teorie e modelli proposti per spiegare e prevedere particolari forme di comportamento animale. Il ruolo a essi assegnato non è quello di dare sostegno a generiche intuizioni meccanicistiche, ma di fornire evidenze per corroborare specifiche teorie scientifiche o per prevedere il comportamento di particolari sistemi viventi in determinate condizioni sperimentali.

Le discipline scientifiche che si occupano di prevedere e spiegare il comportamento dei sistemi viventi – ci riferiamo in particolare alle scienze cognitive, alle neuroscienze, alla psicologia – producono teorie che spesso descrivono i meccanismi alla base del comportamento umano e animale. In alcuni casi tali teorie meccanicistiche vengono utilizzate per prevedere come si comporterà un particolare sistema vivente in particolari circostanze. In altri casi la teoria viene proposta per spiegare tali comportamenti. Supponiamo, per esempio, di disporre di una teoria sui meccanismi neurali che permettono agli astici di muoversi verso le loro fonti di nutrimento nei fondali marini (è un esempio che riprenderemo più avanti). Nel primo caso, quello predittivo, la utilizzeremo per prevedere le prossime mosse di un astice che stiamo osservando sul fondo di una piscina. Nel secondo caso la invocheremo per spiegare perché l'astice ha appena fatto certe mosse, oppure perché si comporta sempre in un certo modo al presentarsi di certi stimoli.

Le teorie meccanicistiche sono spesso oggetti complicati, e non sempre è possibile elaborarle 'a mente' per formulare previsioni basate su di esse. In alcuni casi è possibile incorporare, o simulare, la teoria in una macchina, facendo sì che il meccanismo che essa attribuisce al sistema vivente in esame (per esempio, all'astice) governi il funzionamento della macchina stessa la quale diventa, per così dire, generatrice automatica di previsioni basate sulla teoria. Supponiamo allora di simulare la teoria meccanicistica sull'astice in un robot in grado di muoversi sott'acqua: se la teoria descrive fedelmente i meccanismi neurali che controllano l'astice, e se è stata simulata fedelmente nell'automa, vi saranno buone ragioni per ritenere che il comportamento dell'automa riproduca quello che l'astice avrebbe prodotto nelle stesse circostanze. Avremo dunque utilizzato una macchina che simula una buona teoria su un sistema vivente per prevedere il

comportamento di quest'ultimo. Questa strategia predittiva è intuitivamente allettante e problematica allo stesso tempo. Chi ci assicura che la teoria incorporata nella macchina riproduca davvero i meccanismi neurali dell'astice? Chi ci assicura che sia stata fedelmente simulata nella macchina? Chi ci assicura che il comportamento dell'astice, in quelle condizioni, non venga influenzato da meccanismi neurali che il robot non simula, oppure da variabili ambientali che non è in grado di percepire? Senza queste assicurazioni non abbiamo forti garanzie per fidarci del fatto che il robot-astice riproduca il comportamento che avremmo effettivamente osservato nell'animale; e spesso non è facile ottenere assicurazioni di questo tipo.

Un esempio dell'uso di automi per scopi predittivi è descritto in alcuni recenti lavori del robotico John Long,²¹ che assieme al suo gruppo di ricerca è impegnato nella costruzione di modelli robotici di animali estinti. Per ovvie ragioni non abbiamo alcun modo per risalire al comportamento di animali estinti se non quello di inferirlo a partire da ciò che sappiamo sulla loro struttura osseo-muscolare e sui loro meccanismi di controllo del comportamento. Non possiamo cioè far altro che affidarci alla miglior teoria che abbiamo a nostra disposizione sul loro funzionamento e formulare previsioni comportamentali – certo non previsioni di comportamenti che avverranno in futuro (sono animali estinti), ma affermazioni condizionali della forma “se quell'animale si trovasse ora in certe condizioni si comporterebbe in questo modo”, che hanno strutturalmente molto in comune con ciò che ordinariamente chiamiamo “previsione”, pur riferendosi al passato. Tra i vari animali studiati da John Long c'è il *Plesiosaurus*, un gigante animale marino il cui scheletro fu ritrovato nel XIX secolo in Inghilterra. L'analisi del fossile suggerisce che il *Plesiosaurus* fosse in grado di navigare utilizzando tutte e quattro le zampe, contrariamente a molti tetrapodi acquatici odierni (come le tartarughe marine e gli elefanti di mare) che utilizzano soltanto le due zampe frontali. Quale vantaggio in termini di adattamento al proprio ambiente assicurava al *Plesiosaurus* l'uso di quattro anziché due zampe? E perché le specie tetrapode si sono evolute verso la navigazione a due zampe? Per rispondere a queste domande Long e i suoi collaboratori hanno costruito un automa, chiamato Madeleine, che presenta alcune caratteristiche morfologiche e funzionali simili al *Plesiosaurus* ed è in grado di navigare utilizzando tutte e quattro le zampe oppure soltanto le due frontali. L'analisi del comportamento dell'automa ha condotto gli autori dello studio a concludere che l'uso di quattro zampe plausibilmente permetteva al *Plesiosaurus* di muoversi più o meno con la stessa velocità che avrebbe potuto raggiungere con due zampe, al prezzo però di un maggior consumo di energia. Qual è il vantaggio allora? In una maggiore accelerazione: questa ‘previsione al passato’ del comportamento dell'animale estinto, aiutata dalla costruzione di un automa, solleva dunque altre domande sul ruolo adattativo che la capacità di ‘scattare’ verso la preda potesse presentare nell'ambiente in cui il gigante marino viveva.

Come anticipato all'inizio di questa sezione, i robot sono chiamati a svolgere anche ruoli speculari rispetto a quello appena discusso: ruoli connessi cioè non con la previsione o ricostruzione di un comportamento non osservato – come nello studio sul *Plesiosaurus* – ma con la spiegazione di comportamenti già osservati. Il contributo dell'automa, in questo caso, è quello di

²¹ J. LONG, *Darwin's Devices. What Evolving Robots Can Teach Us About the History of Life and the Future of Technology*, Basic Books, New York, 2012.

fornire ragioni per accettare o rifiutare una particolare teoria meccanicistica proposta per la spiegazione del comportamento di particolari specie viventi. Illustriamo questa possibilità richiamando l'esempio dell'astice che abbiamo discusso sopra ma riferendoci a uno specifico studio condotto da Frank Grasso e collaboratori.²² L'obiettivo dello studio era quello di valutare l'adeguatezza esplicativa di una semplice ipotesi sul meccanismo che guida gli astici alla ricerca di fonti di nutrimento sui fondali acquatici. Secondo tale ipotesi, l'astice è in grado di risalire la scia di sostanze chimiche rilasciata dalle sue prede nell'acqua sfruttando i sensori chimici (in gergo, chemorecettori) presenti sulle sue antenne. Il meccanismo è intuitivamente semplice: l'astice si muove sempre nella direzione dell'antenna che segnala una maggior concentrazione chimica. Se per esempio a un certo punto l'antenna destra segnala la presenza di sostanze chimiche non segnalate dall'antenna sinistra (situazione che plausibilmente si verifica quando alla destra dell'animale si trova una "nuvola" di sostanza chimica disciolta nell'acqua), l'astice sterza verso destra, e viceversa per l'altro lato. La strategia sperimentale adottata da Grasso e dal suo gruppo è stata quella di incorporare questo semplice meccanismo in un piccolo robot acquatico, chiamato *RoboLobster*, e di valutare la sua capacità di riprodurre il comportamento della sua controparte vivente, ovvero la capacità di raggiungere fonti di scie chimiche disciolte nell'acqua. Le prestazioni del robot non sono state soddisfacenti: in alcune sessioni sperimentali, soprattutto quando la posizione iniziale era troppo lontana dalla destinazione, *RoboLobster* si 'perdeva' nell'acqua. Questo fallimento prestazionale è però stato interpretato come un risultato teorico degno di nota: dopo alcuni controlli tesi a eliminare spiegazioni alternative gli autori hanno concluso che i fallimenti di *RoboLobster* segnalano l'eccessiva semplicità del meccanismo proposto che, in certi casi, non è in grado di guidare l'animale verso la sorgente. In altre parole, la costruzione di *RoboLobster* ha condotto gli autori a rifiutare un'ipotesi esplicativa sui meccanismi che orientano gli astici verso le loro fonti di nutrimento.

IV. CONCLUSIONI

In questo contributo abbiamo illustrato, attraverso esempi, alcuni ruoli che la costruzione di un automa può svolgere nella riflessione teorica sui processi dell'agire umano e animale. Abbiamo messo in luce come gli automi abbiano fornito sostegno ad analisi puramente meccanicistiche di certi aspetti del comportamento fino ad allora ritenuti di esclusiva pertinenza dei sistemi viventi. In altri casi, i robot hanno svolto un ruolo cruciale nella previsione o nella spiegazione del comportamento di particolari specie viventi. Queste destinazioni d'uso conoscitivo degli automi sollevano numerosi problemi epistemologici e di metodo scientifico che esulano dagli scopi di questo contributo. Problemi destinati però a diventare sempre più urgenti, dato che

²² F. GRASSO, T. CONSI, D. MOUNTAIN, J. ATEMA, *Biomimetic Robot Lobster Performs Chemo-Orientation in Turbulence Using a Pair of Spatially Separated Sensors: Progress and Challenges*, in «Robotics and Autonomous Systems», XXX, 2000, 1-2, pp. 115-131.

l'avanzamento delle tecnologie robotiche, e la conseguente possibilità di replicare in misura sempre più fine la morfologia, il comportamento, i meccanismi neurali e cognitivi dei sistemi viventi, si preparano a generare sempre più provocatori specchi meccanici dell'agire umano e animale.

edoardo.datteri@unimib.it

(Università degli Studi di Milano-Bicocca)

BIBLIOGRAFIA

- G. ALLEN, *Mechanism, Vitalism and Organicism in Late Nineteenth and Twentieth-Century Biology: The Importance of Historical Context*, in «Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences», XXXVI, 2005, 2, pp. 261-283.
- V. BRAITENBERG, *Vehicles. Experiments in Synthetic Psychology*, The MIT Press, Cambridge, MA 1986.
- R. CARTESIO, *Discorso sul Metodo*, in *Opere filosofiche 1*, a cura di E. GARIN, Laterza, Bari, 1996, pp. 289-350.
- R. CARTESIO, *L'Uomo*, in *Opere filosofiche 1*, a cura di E. GARIN, Laterza, Bari, 1996, pp. 203-288.
- M. CIANCHETTI, M. CALISTI, L. MARGHERI, M. KUBA, C. LASCHI, *Bioinspired Locomotion and Grasping in Water: The Soft Eight-Arm OCTOPUS Robot*, in «Bioinspiration and Biomimetics», X, 2015, 3. DOI: 10.1088/1748-3190/10/3/035003.
- R. CORDESCHI, *The Discovery of the Artificial. Behavior, Mind and Machines Before and Beyond Cybernetics*, Springer Netherlands, Dordrecht, 2002.
- E. DATTERI, *Filosofia delle scienze cognitive. Spiegazione, previsione, simulazione*, Roma, Carocci, 2012.
- E. DATTERI, *Intelligenza Artificiale e Scienza Della Natura*, in P. PECERE (ed.), *Il Libro Della Natura. II. Scienze e Filosofia Da Einstein Alle Neuroscienze Contemporanee*, Carocci, Roma 2015, pp. 180-208.
- P. DUMOUCHEL, L. DAMIANO, *Living with Robots*, Harvard University Press, Cambridge, MA 2017.
- F. GRASSO, T. CONSI, D. MOUNTAIN, J. ATEMA, *Biomimetic Robot Lobster Performs Chemo-Oriented Navigation in Turbulence Using a Pair of Spatially Separated Sensors: Progress and Challenges*, in «Robotics and Autonomous Systems», XXX, 2000, 1-2, pp. 115-131.
- W. GREY WALTER, *An Imitation of Life*, in «Scientific American», CLXXXII, 1950, 5, pp. 42-45.
- H. JONAS, *A Critique of Cybernetics*, in «Social Research», XX, 1953, 2, pp. 172-192.
- R.G. KRUEGER, C.L. HULL, *An Electro-Chemical Parallel to the Conditioned Reflex*, in «Journal of General Psychology», V, 1931, pp. 262-69.
- C.L. MORGAN, *Animal Automatism and Consciousness*, in «The Monist», VII, 1896, 1, pp. 1-18.
- J. LOEB, *Comparative Physiology of the Brain and Comparative Psychology*, G.P. Putman's Sons, New York, 1900.
- J. LONG, *Darwin's Devices. What Evolving Robots Can Teach Us About the History of Life and the Future of Technology*, Basic Books, New York, 2012.
- R. PFEIFER, M. LUNGARELLA, F. IIDA, *Self-Organization, Embodiment, and Biologically Inspired Robotics*, in «Science», CCCXVIII, 2009, pp. 1088-1093.
- A. ROSENBLUETH, N. WIENER, J. BIGELOW, *Behavior, Purpose and Teleology*, in «Philosophy of Science», X, 1943, 1, pp. 18-24.
- T. ROSS, *Machines That Think*, in «Scientific American», LXXXIX, 1933, pp. 206-208.
- E.R. TRUITT, *Medieval Robots. Mechanism, Magic, Nature, and Art*, University of Pennsylvania Press, Philadelphia, PA 2016.