

Rivista
trimestrale
di analisi e
critica

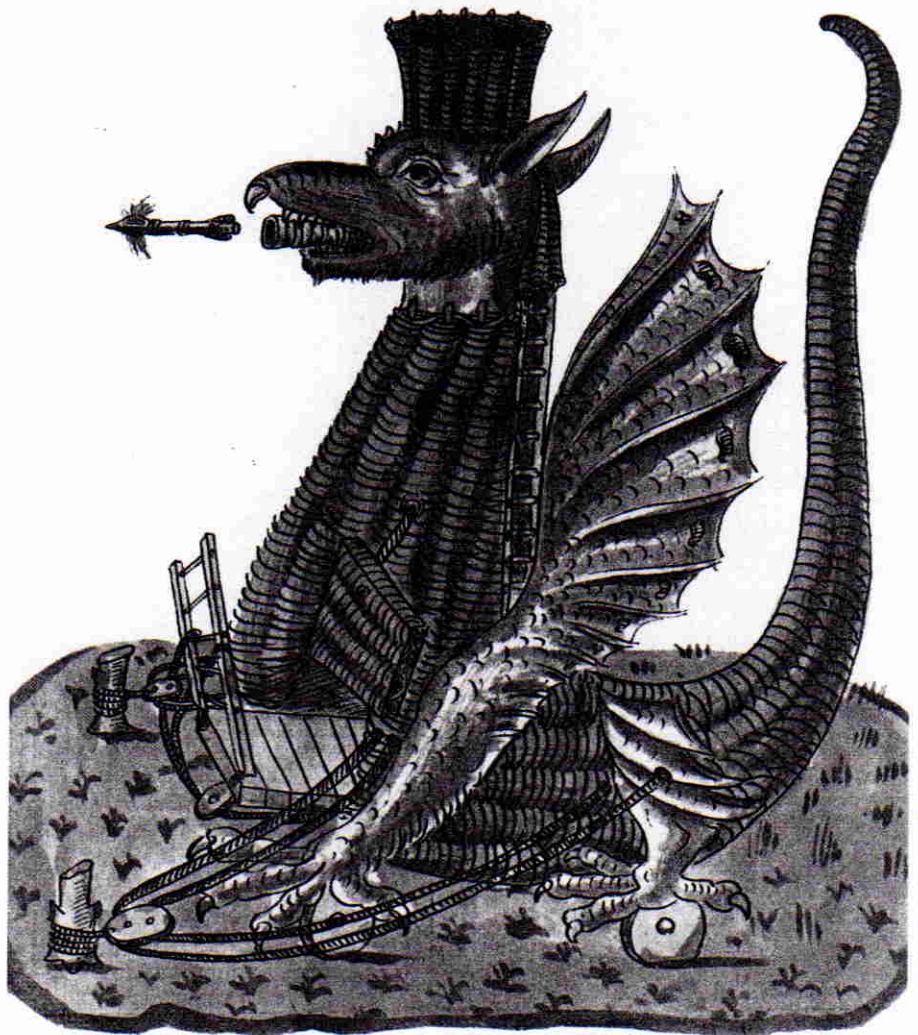
ANNO VII
N. 4 (28) 1989
L. 10.000

Nuova Civiltà delle Macchine

LO SPAZIO DELL'INGEGNERIA NELLA SCIENZA MODERNA

Arabia machina ad expugnationem civitatum magna et ingenii viri ponentis in manifestis instrumentis bellis secula

Arnaldo Maria Angelini
Antonio Ruberti
Giuseppe Biorci
Gianni Vattimo
Umberto Colombo
Vincenzo Tagliasco
Vittorio Marchis
Nicola Cabibbo
Luigi Donato
Ugo Montanari
Gian Maria Gros Pietro
George Bugliarello
Francesco Barone
Carlo Andreoni
Luciano Bazzocchi
Guido Fiorentino
Gian Franco Querciagrossa



Nuova Eri

Una prospettiva cognitivista della mente e dell'intelligenza artificiale

72

1.

Ogni rappresentazione di un fenomeno naturale deve vedersela col paradosso di riprodurre in qualche modo il proprio oggetto senza esserne, ovviamente, una vera e propria duplicazione¹. Una teoria generale della mente si imbatte in una difficoltà di grado superiore, in quanto il suo stesso oggetto riflette in sé l'universo degli oggetti possibili. Una siffatta teoria si muove costantemente sull'orlo di un abisso senza fondo: la tentazione di abbandonarsi al vortice dei reciproci rispecchiamenti e divenire così una teoria metafisica della conoscenza e del mondo in generale. Il proposito appare ancora più arduo qualora vi si giunga da vie traverse del sapere umano: non cavalcando specifiche problematiche di neurobiologia o di psicologia sperimentale, ma perseguendo scopi e esigenze originate in un settore affatto differente. Eppure, tale è la situazione odierna dell'intelligenza artificiale (IA), una ramificazione dell'informatica e dell'ingegneria elettronica tutta protesa nella nuova sfida di Prometeo: riprodurre le funzioni mentali dell'uomo con macchine e sistemi artificiali.

Fino a tutti gli anni '60, l'IA si è accontentata di simulare le funzioni logico-deduttive della mente, utilizzando teorie matematiche e sistemi formali; ma presto ci si è accorti che le attività più ovvie del buon senso quotidiano, presupposto di ogni ulteriore astrazione intellettuale, richiedono capacità e attitudini ben diverse. D'altra parte, né la filosofia né la psicologia, a cui con ingenua aspettative i ricercatori d'intelligenza artificiale si sono rivolti, sembrano offrire indicazioni a tutta prima utilizzabili. La psicologia non si è mai spinta fino ad ipotesi funzionali tanto

dettagliate da suggerire concrete soluzioni e meccanismi efficaci, il che fa temere che la riproduzione artificiale delle attività mentali sia in pratica impossibile. La filosofia ha messo in luce le aporie di fondo e gli ostacoli di più alto livello, lasciando supporre che tale riproduzione sia impossibile in linea di principio.

Ciò malgrado, la pervicacia e l'ingegnosità di alcuni pionieri hanno portato a due risultati con cui, piaccia o non piaccia, occorre fare i conti. Sul piano di fatto, alcune funzioni che siamo abituati a considerare intelligenti sono state effettivamente simulate, a costo di qualche semplificazione, da strumenti artificiali; i prototipi più riusciti esibiscono capacità non trascurabili, e tutto fa pensare che possano progredire in varie direzioni. È possibile, insomma, realizzare sistemi più complessi di quanto non si riesca a teorizzare e pianificare. La seconda evidenza è che si incominciano ad intravedere tecnologie capaci, in linea teorica, di proiettare e dare corpo a qualunque sogno, il quale riesca a dettagliarsi in strutture coerenti e in processi via via più elementari. Nel caso che il compito di simulare l'intelligenza non sia per principio impossibile, questa è l'occasione per tentare; gli stessi distinguo posti dalla riflessione filosofica potrebbero incontrare qui, se non proprio una controprova fattuale, almeno un'incarnazione esemplificativa finora impensabile. Se dunque l'IA vuole procedere per la sua strada, non può sperare in puntelli esterni e deve ridisegnare una mappa dei possibili percorsi, avanzando ipotesi completamente innovative; l'assenza di un impianto teorico di fondo, per quanto ipotetico possa essere, impedisce infatti la lettura e la valorizzazione di un quadro ancora piuttosto caotico.

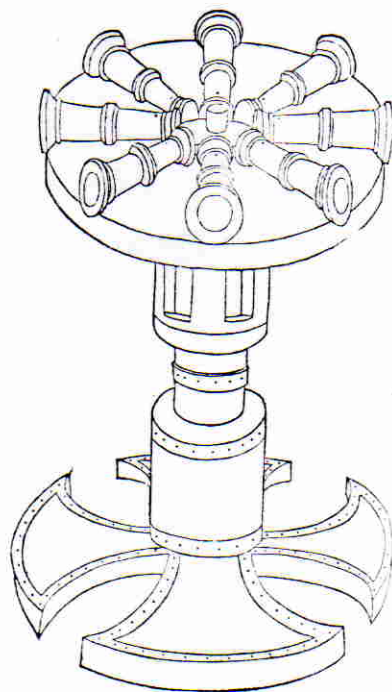
2.

Il primo tentativo organico in questa direzione giunge dal lavoro di Marvin Minsky, personaggio chiave dell'IA statunitense e promotore di quasi tutte le iniziative più brillanti. Il corposo saggio pubblicato nel 1985-86, ora uscito in traduzione italiana col titolo *La società della mente* (Adelphi, Milano 1989), raccoglie le riflessioni germinate intorno alla sua attività di ricerca e le riorganizza in una struttura logico-espositiva audace ed originale. Una novità significativa è data dal superamento di ogni separazione tra lo studio della mente umana e le mete un po' astratte dell'IA. *La società della mente* può essere letto come un testo di psicologia denso di osservazioni sperimentali e riferimenti all'evoluzione storica delle capacità intellettive (tanto lungo la scala filogenetica che nello sviluppo dall'infanzia all'età adulta), il quale però costantemente illustra termini classicamente cognitivi (sensazione, pensiero, ricordo, idea, emozione) con modelli computabili, schemi neuronali, circuiti logici e programmi di simulazione. Oppure, lo si può considerare un grandioso progetto di IA, un piano globale per la riproduzione artificiale di una mente intelligente (variamente articolato e dettagliato in piani di lavoro ed esempi realizzativi), il quale prende a modello la mente umana e le sue caratteristiche più qualificanti e imprescindibili.

L'assunto di fondo è che la mente (a questo punto, non la mente dell'uomo o l'intelligenza artificiale, ma una mente «trascendentale» in genere, in quanto funzionalmente analoga a quanto noi sperimentiamo oggettivamente e introspektivamente come tale) non è qualcosa di individuale e di semplice, ma è una stratificazione di più «menti» tra loro in competizione, è una società dinamicamente organizzata in una infinità di livelli operativi, unità ricettive, centri decisionali, elementi di controllo, di memorizzazione, di inibizione. Le componenti della mente operano alla stregua di «agenzie» gerarchicamente interconnesse, le quali si giovano di «agenti» tra loro cooperanti. Il punto nevralgico è la «burocrazia» delle agenzie: per ciascun compito specifico si viene a determinare una certa struttura funzionale che rafforza determinati collegamenti e altri ne inibisce. Ad esempio, la capacità di reazione alla forma cubica è utilizzata dall'agenzia «vedere» in funzione dell'agenzia «trovare», primo passo di «aggiungere» cui fa appello il «costruttore di torri» quando l'attività «giocare coi blocchetti di legno» conquista sufficiente priorità rispetto agli interessi alternativi del bambino. Ciascuna linea

gerarchica è sempre ipotetica e transitoria: ipotetica in quanto è un tentativo di organizzazione per livelli che nel corso dell'esperienza si è rivelata efficace, ma può necessitare di revisioni o capovolgimenti; transitoria, perché può essere sempre sopraffatta da altre esigenze, che si riversino dalle agenzie di più alto livello o riemergano dai sensori dei bisogni più elementari.

Per adeguarsi ad una materia tanto articolata, la narrazione di Minsky (come ammette l'autore, il libro «dovrebbe essere letto più come una storia avventurosa dell'immaginazione che come un testo di dottrina scientifica», p. 638) si ramifica in alcune centinaia di singoli capitoletti, raggruppati per temi più generali, che alla fine rivelano una struttura significativamente più ambiziosa di quanto il tono di ogni parte lasciasse prevedere. Il carattere non apodittico della costruzione concettuale e la sobrietà della linea espositiva, che senza pretesa di sicurezza sceglie una tra le varie direzioni aperte dalle considerazioni precedenti, finiscono per rafforzare il discorso complessivo assai meglio di quanto non accadrebbe in un impianto votato a un costante rigore deduttivo. La ricchezza di osservazioni locali evita così il pericolo della frammentarietà e confluisce in un quadro globalmente convincente.



Da R. Valturio. De re militari. 1472

Tra le idee piú feconde, emerge la caduta di ogni barriera pregiudiziale tra pensiero astratto e pensiero ordinario, e perfino tra pensiero razionale e impulso emotivo. Le emozioni hanno infatti un ruolo insostituibile nel contrassegnare valori e mete, a loro volta necessari per spingere a linee ragionate e comportamentali prive di motivazioni contingenti. Poiché, inoltre, la versatilità dell'intelletto umano risiede nella capacità di definire scopi nuovi e nel trasmetterli culturalmente alle generazioni successive, relazioni emotive quali «la paura e l'affetto, l'attaccamento e la dipendenza, oppure l'amore e l'odio» (p. 317) sono anche il mezzo per adottare o modellare i sistemi assiologici altrui. Come nel mondo dei blocchetti lo scopo particolare di costruire una torre deve avvalersi delle capacità generiche di vedere, afferrare, spostare, ecc., così in generale le sottomete sono in realtà piú astratte delle mete; il pensiero teorico non è necessariamente frutto di una ascesi induttiva sempre piú comprensiva, bensí il semplice affinamento di capacità elementari coinvolte nei compiti posti al buon senso. In definitiva, «le nostre investigazioni piú astratte hanno origine dalla ricerca dei mezzi per conseguire fini ordinari», mezzi tanto piú efficaci quanto piú conosciamo e comprendiamo come funziona il mondo in cui siamo (p. 343). Tale riavvicinamento tra pensiero informale e pensiero logico suggerisce contributi originali, come quando si osserva che la spiegazione freudiana dei motti di spirito, i quali avrebbero a che fare con tabù, censure e ferite emotive, si applica anche al gusto dell'assurdo e al ridicolo del paradosso, in quanto anche il pensiero deduttivo si nutre di divieti, di costrizioni, di tabù. L'umorismo ha una funzione pratica, «ha a che fare col modo di apprendere dei nostri censori, e soprattutto col pensiero "negativo"» (pag. 546), con quelle strutture mentali, cioè, volte ad inibire i ragionamenti rivelatisi «pericolosi»; una macchina artificiale che voglia «imparare ad apprendere» deve conoscere anche l'ironia.

3.

La posizione di Minsky, come si intuisce, è sostanzialmente riduzionista, nel senso che accetta «l'ipotesi che ogni cervello, macchina o altra cosa che posseda una mente debba essere composto di cose piú piccole, affatto incapaci di pensiero» (p. 637); ma anche nel senso che molte strutture mentali di alto livello, nell'ambito della concatenazione concettuale o della padronanza

linguistica, adottano e riutilizzano fisicamente, per così dire, schemi logici e procedure messe a punto nella prima infanzia a corredo della coordinazione motoria e della rappresentazione spazio-temporale. La stessa sintassi del linguaggio, almeno per le strutture comuni a piú ceppi glottologici, rimanda sia «agli interessi dei nostri antenati» e «ai tipi di domande e di problemi che in quel momento apparivano importanti» (p. 525), sia al modo in cui via via si articola nel bambino la capacità di pensare e di organizzare rappresentazioni e ricordi.

Da *La società della mente* è completamente bandito, in modo perfino irritante, qualsiasi riferimento esplicito alla cultura filosofica ed anche, con l'eccezione di qualche accenno a Freud, Piaget e Papert, alla tradizione psicologica. Questo indubbiamente riflette le attitudini del suo autore, ma è anche una scelta di campo, empirista e antiaccademica. Non è un caso se tra le piú varie epigrafi che contrassegnano diversi capitoli trovano posto David Hume e Thomas Hobbes, con due citazioni entrambe di una descrittività quasi naturalistica. Facendo di necessità virtù, e a costo di passare per ingenuo, Minsky aspira così a riacquistare uno sguardo sgombro, una sensibilità descrittiva, che un'eccessiva riverenza per la tradizione finirebbe per inibire. Tale commistione di coraggio e impudenza, non priva di un suo fascino, permette se non altro prese di posizione molto franche, come a proposito della discussione sul rapporto fra la mente e il corpo: «In queste idee non vedo alcun interesse poiché, per quanto mi riguarda, il cosiddetto problema mente-corpo non racchiude alcun mistero. *La mente è semplicemente ciò che fa il cervello*» (p. 563).

4.

Riconoscendo dunque al libro di Minsky una consistente portata culturale, soprattutto per le assidue sollecitazioni pratiche sul versante dell'intelligenza artificiale, che qui non è possibile esemplificare, resta però il rammarico per l'assenza di un confronto esplicito col principale approccio antagonista ai temi trattati. A fronte della concezione cognitivista dei processi cerebrali, per la quale il cervello elabora rappresentazioni simboliche ed opera, in modo piú o meno algoritmico, su elementi semanticamente rilevanti, negli anni '80 ha ripreso piede l'impostazione connessionista e reticolare. In questa seconda prospettiva, le capacità cerebrali sarebbero piuttosto il risultato dell'addestramento di reti neuro-

nali in origine connesse in modo generico e quasi casuale, mediante meccanismi di retroazione a carattere stocastico e probabilistico; l'informazione in esse codificata non è simbolica e localizzata, ma sub-simbolica e distribuita². Le simulazioni recentemente realizzate a *computer* su tali principi mostrano curve di apprendimento (e di «dimenticanza»), reazioni a traumi distruttivi e capacità di adattamento all'imprevisto, le quali appaiono sorprendentemente analoghe a determinate prestazioni del cervello biologico. Perché allora Minsky, per altro assai tollerante in linea generale e convinto assertore che il cervello è una commistione di centinaia di modalità elaborative differenti, compresa quella reticolare e connessionista, insiste poi a utilizzare il modello cognitivista anche là dove parrebbe meno plausibile, a livello di circuiti neuronali e connessioni sinattiche? È davvero realistico immaginare che specifici «agenti» capaci di rispondere «rosso», oppure «piccolo», oppure «rotondo», forniscano disciplinatamente le rispettive conclusioni (sì/no) all'«agenzia» che sta controllando l'ipotesi «mela»?

Una prima risposta sta nel fatto che il libro, uscito nel 1985, raccoglie il lavoro di almeno un decennio precedente, durante il quale la teoria reticolare non aveva ancora rafforzato i propri paradigmi. Dare di essa una valutazione equilibrata avrebbe significato spostare il baricentro dell'intera opera.

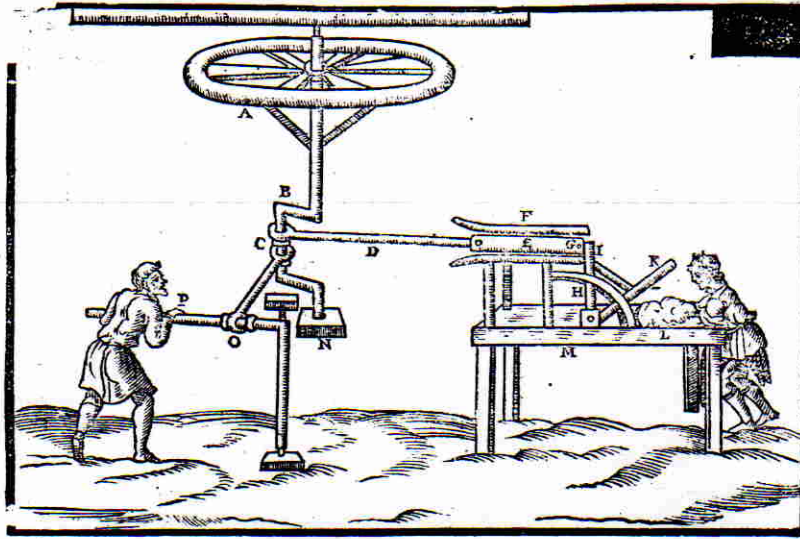
In quanto al merito, bisogna riconoscere che in effetti un'ipotesi di lavoro è tanto più efficace e produttiva quanto più dipana coerentemente i propri assunti fino alle estreme conseguenze. Appunto in funzione della competizione interteorica il cognitivismo dell'intelligenza artificiale deve saper delineare tutte le proprie capacità esplicative, anche e soprattutto nelle aree per esso più difficili e problematiche. Se le teorie non hanno sufficiente «vischiosità» e tenacità, di là dal personale incaponirsi dei loro sostenitori, anche il processo di «falsificazione» perde di valore e diventa una dissennata opera di distruzione. Non è affatto scontato che, almeno in alcuni casi, il micro-processo cerebrale non abbia aspetti simbolici, ed è fondamentale sapere fin dove l'ipotesi cognitivista riesce a spingersi.

La terza ragione coinvolge considerazioni epistemologiche più generali. Nella tradizione moderna, ci si aspetta che una teoria scientifica non solo dia conto dei fenomeni, ma li interpreti e li inquadri in un sistema di significati. L'interpretazione di Einstein non ha modificato le equazioni di Maxwell o le trasformazioni di Lorentz, eppu-

re ha conferito loro una portata del tutto nuova. Viceversa, l'insoddisfazione per le caratteristiche del fotone o per il principio di Heisenberg nasceva dalla difficoltà nell'attribuire loro un qualsivoglia significato, una volta scardinati i binomi classici onda-particella o posizione-velocità.

Ora, se c'è una teoria intrinsecamente affezionato all'aspetto interpretativo della teorizzazione scientifica, essa è appunto l'approccio cognitivo allo studio del mentale. E se c'è una teoria dichiaratamente costruttivista e sospettosa di ogni attribuzione semantica, questa è il connessionismo neuronale. Ciò non significa che il cognitivismo non disponga di consistenti strumenti matematici e di una solida impostazione sperimentale e costruttiva, né che il connessionismo non sia un'ipotesi dotata, a un certo livello, di potere esplicativo: ma la loro contrapposizione tende ad esaltare caratteristiche diverse. Da più parti si auspica una fusione delle due prospettive: mentre le funzioni mentali ad alto livello possono essere meglio descritte e simulate su un impianto cognitivista, il connessionismo pare più adatto a spiegare l'attività neuronale e a riprodurre le prestazioni meno astratte. Ma il punto di incontro dei due approcci («dall'alto» e «dal basso») resta problematico e puramente convenzionale. Non c'è ragione di principio perché il connessionismo arresti i propri costrutti reticolari per convertire l'informazione in termini simbolici, né perché il cognitivismo sospenda le proprie propaggini logico-formali per affidarsi a differenti strategie: il compromesso avrebbe solo un significato pratico e contingente.

Per sua natura, il modello connessionista opera, rispetto alle funzioni acquisite, a «scatola nera»: è in grado di «apprendere» a rispondere in modo pertinente alle sollecitazioni esterne, ma non di giustificare la propria risposta. È, cioè, un artefatto che si presenta, nel suo stato finale, come un oggetto sconosciuto e spesso sorprendente, da analizzare *ex novo* con strumenti probabilistici, stocastici, o addirittura estetici. Un sistema reticolare in grado di distinguere i caratteri di stampa non ingloba nessuna teoria topologica, né suggerisce criteri operativi; è la prova che semplici neuroni, anche scarsamente prestrutturati, possono apprendere il compito, ma appare altrettanto muto quanto al modo. Potremo abituarci a ritenere questa una «spiegazione» del fenomeno, così come abbiamo finito per accettare certe formulazioni matematiche della fisica sub-atomica o delle dinamiche dei sistemi complessi? Se anche il cervello fosse una scatola nera analizzabile solo con metodi di campiona-



Macchina per battere e assodare la pasta
(da G. Branca, *Le machine*, 1629)

mento e di correlazione, come soddisfare l'esigenza di ricostruire in termini intuitivi le linee di «ragionamento», la concatenazione dei «pensieri» o il flusso delle «idee»? La tenacia con cui Minsky tiene fede al metodo cognitivo esprime la volontà di trarre dall'indagine una spiegazione teorica e psicologica dei nostri processi mentali, e non solo una dimostrazione della loro possibilità a partire da un meccanismo fisico sufficientemente complesso (cosa che si può ben ammettere fin dal principio). Se una teoria scientifica è necessariamente una costruzione astratta, un modello concettuale, non deve stupire che alcuni degli schemi proposti per spiegare le singole operazioni del cervello non assomiglino, in modo plausibile, ad un fascio di neuroni.

5.

Vi è infine un ulteriore contesto nel quale l'alternativa tra connessionismo e cognitivism comporta conseguenze ancor meno settoriali. In quanto modelli del cervello umano, le simulazioni artificiali sono volte, in ultima istanza, a trattare e comprendere il mondo fenomenico. Il modo in cui una macchina connessionista potrebbe esperire e teorizzare sul mondo esterno (facciamo un po' di fantascienza!) porrebbe gli stessi problemi e conserverebbe i misteri del nostro modo di apprendere. Una macchina cognitivista, al contrario, deve i suoi successi (come pure gli insuccessi) alle teorie, alle regole e agli schemi in

base a cui opera; è essa stessa, in qualche maniera, una costruzione teorico-ipotetica sottoposta a controllo (e falsificazione) ad ogni passo¹. Questo ne fa un obiettivo decisivo per tutta la ricerca scientifica, perché costituirebbe il banco di prova per qualsiasi teoria sufficientemente formalizzata e completa da riuscire a vivere con le sole proprie forze, e comunque ne determinerebbe la necessaria formalizzazione. Una macchina cognitivista potrebbe anzi suggerire e controllare nuove ipotesi teoriche, in quanto di esse avrebbe bisogno (letteralmente!) come il pane. Oggettivato in leggi e rapporti indipendenti dalla capacità interpretativa dello scienziato, l'impianto teorico della scienza moderna si troverebbe direttamente esposto al vaglio della situazione di fatto e potrebbe dimostrare punto per punto il livello della propria adeguatezza.

Il fallimento di determinate istanziazioni di macchine cognitive costituisce la falsificazione, ovviamente, della composizione teorica in esse rappresentata e non, come alcuni frettolosamente hanno concluso⁴, dello stesso programma cognitivista: come se l'esito negativo di un esperimento di laboratorio potesse segnare il fallimento della scienza fisica, o dell'intera ricerca sperimentale. Al contrario, la consapevole fallibilità delle ipotesi e delle realizzazioni della prospettiva cognitivista la promuove a strategia di ricerca attiva e feconda, in sintonia con le investigazioni scientifiche più accreditate. La ricchezza di idee ma anche la consequenzialità operativa di personalità come Marvin Minsky fanno pensare che essa avrà molto da dire nel prossimo futuro.

¹ «In immagine e raffigurato qualcosa dev'essere identico, affinché quella possa essere un'immagine di questo» (Wittgenstein, *Tractatus logico-philosophicus*, 2.161).

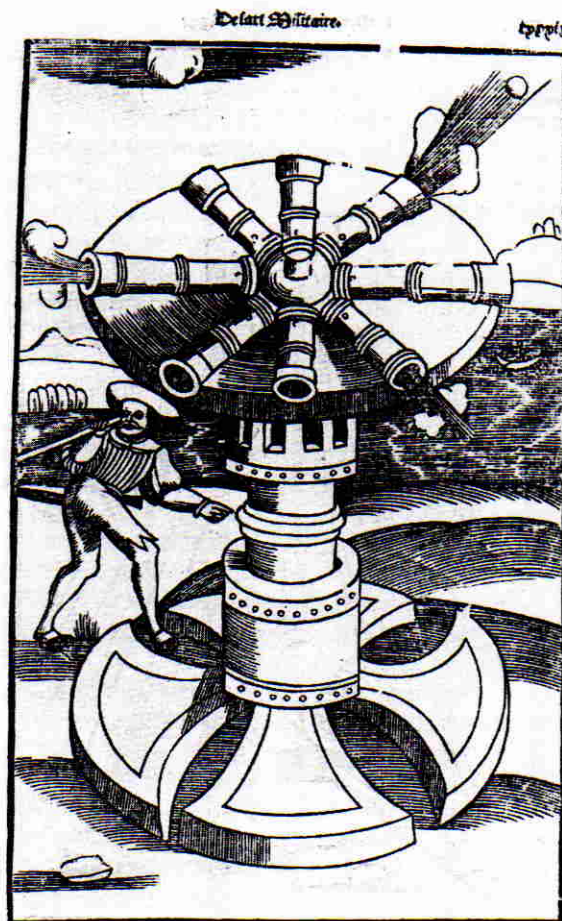
² «Una rappresentazione è distribuita quando una determinata entità è rappresentata da un insieme di unità della rete. Tuttavia, questo non è sufficiente. [...] Quello che invece è essenziale per poter parlare di rappresentazioni distribuite è che gli elementi componenti attraverso i quali rappresentiamo una certa entità non abbiano significato per noi, ma siano elementi caratterizzati soltanto da proprietà fisico-matematiche» (Domenico Parisi, *La mente come cervello*, in «Sistemi intelligenti», agosto 1989, p. 226).

³ Quando si dice che un sistema connessionistico apprende «per tentativi ed errori», l'epistemologo popperiano drizza tutte le sue antenne. Ma invano si cercherebbe nel connessionismo un processo di falsificazione, perché un tale sistema non rappresen-

ta né presuppone alcuna teoria, non raffigura nulla che si possa dire vero o falso. Il modello connessionista è piuttosto un perfetto meccanismo «tolemaico», con innumerevoli epicicli correttivi, tutto orientato ad aggiustamenti *ad hoc* per dar conto dei fenomeni, o almeno dei più recenti. Un sistema connessionistico evolve complessivamente secondo principi termodinamici, ma dal punto di vista epistemologico il suo prototipo è il termostato dello scaldabagno.

⁴ Si veda ad esempio Pietro Morasso, *Reti neurali e intelligenza* in «Sistemi Intelligenti», agosto 1989, pp. 243-4: «Gli insuccessi rendono onore all'approccio costruttivista [del cognitivismo e dell'IA], che per definizione assume l'onere della prova, esibendo non mere teorie ma artefatti il cui funzionamento può essere valutato, confrontato, confutato. Questo ha garantito un metabolismo estremamente elevato della ricerca poiché gli insuccessi [...] non possono non costringere ad un mutamento di indirizzo».

77



Da F. Vegezio, *De re militari*, 1511-1512