

INTELLIGENZA ARTIFICIALE E INTELLIGENZA NATURALE

Problemi e prospettive

S am

INTRODUZIONE

Nell'estate del 1956, al Dartmouth College nel New Hampshire, si tenne uno storico convegno chiamato "Dartmouth Summer Research Project on Artificial Intelligence" la cui più evidente e meno significativa influenza fu l'introduzione nell'uso della locuzione "Artificial Intelligence" o "Intelligenza Artificiale". Organizzato da Marvin Minsky e John McCarthy (l'ideatore del LISP); questo raduno doveva avere lo scopo di radunare i pionieri della nuova disciplina per un fervido scambio di idee sulle possibili e future linee di sviluppo. Nonostante i buoni propositi fu piuttosto una conferenza sullo "stato dell'arte", ma pose ugualmente una base importante per i successivi sviluppi, se non altro come punto di partenza ufficiale.

Questo interesse per l'IA non sgorgava dal nulla, ma era strettamente legato ai rapidi successi nella progettazione e realizzazione di elaboratori, cominciati con l'ENIAC di Mauchly e Eckert nel 1943 e giunti allo stesso anno della conferenza con il JOHNNIAC di J. Von Neumann. Le prime conquiste in questo campo, che già potremmo azzardarci a chiamare informatica, sembravano un buon indizio da cui dedurre che presto il computer avrebbe potuto svolgere tutte le mansioni intellettuali e sgravare l'uomo "dalla fatica del concetto". Questo perché i primi tentativi furono incentrati su capacità intellettive strettamente legate alla logica e quindi alla matematica booleana con cui funzionano i computer.

Altra forte spinta fu quella della Cibernetica, altra disciplina nata in quegli anni, che, dopo l'articolo "Un calcolo logico delle idee immanenti nell'attività nervosa" di McCulloch e Pitts del 1943, alimentò non solo speranze, ma vere e proprie scuole di pensiero sulla natura della attività mentale. Il nome stesso rivela gli originali intenti della disciplina (in greco significa "arte del governare la nave") che erano quelli di studiare le attività di controllo e di comunicazione negli esseri viventi; col passare del tempo questa disciplina si è venuta sempre più specializzando nello studio dell'attività del cervello umano ed è stata sempre più a sproposito citata in ambiti ben separati dal suo campo di studi, quali la teoria della conoscenza, che certo è collegata alle attività cerebrali di controllo, ma di certo non si identifica con esse.

Infine, nel 1958 Rosenblatt progettò e realizzò la prima rete neurale, il *Perceptron*, mentre poco tempo addietro Newell e Simon avevano ottenuto i primi risultati nel campo della manipolazione di simboli da parte di una macchina.

In seguito i progressi tecnici sono stati notevoli, anche se nel campo della sola IA essi sono stati poco vistosi o appariscenti e legati più che altro a continui cambi di impostazione e di prospettive.

Questa è la genesi dell'intelligenza artificiale, segnata dunque dallo svilupparsi dell'Informatica e della Cibernetica.

Come spesso è accaduto nella storia delle scienze sperimentali (e l'Informatica può rientrarvi a pieno titolo, come diremo più avanti), la giustificazione pratica (ovvero sia l'utilità) prevale sulla giustificazione teorica e quindi si iniziò a parlare di intelligenza artificiale e a lavorarvi alacremente prima ancora di averne una definizione rigorosa e di inquadrarla teoricamente. O forse dovremmo dire che vi fu una base teorica, ma solo da un punto di vista matematico-logico: il lavoro in cui Turing presenta la sua celebre Macchina universale è del 1936, come anche la proposta di un test per verificare se una macchina sia o meno intelligente.

Quello che manca ancora è la definizione di intelligenza, un concetto talmente comune che la sua esplicitazione è assai complessa e tuttora non è totalmente delineato.

A questo proposito intervennero allora e intervengono tuttora la filosofia, la psicologia e, più tardi, la neurofisiologia. Queste discipline hanno tra i loro obiettivi lo studio della mente e della conoscenza, seppur da punti di vista differenti.

Abbiamo già accennato al fatto che anche la matematica ha preso la parola nella discussione sull'IA, in forza dello stretto legame tra la logica e il pensiero razionale.

Ora, si potrebbe essere indotti a credere che l'IA sia una sorta di argomento multidisciplinare in cui ogni settore ha una sua parte ben precisa; ebbene, non è così: l'IA è piuttosto transdi-

sciplinare, ovvero sia attinge a svariati campi della conoscenza unendone i contributi per raggiungere i propri obiettivi, o quanto meno per proseguire la propria ricerca.

Questa natura ibrida ha fatto sì che l'IA sia ricca di questioni irrisolte e affascinanti, sulle quali punteremo la nostra attenzione.

Innanzitutto è interessante notare come non vi sia assoluta chiarezza nemmeno sugli scopi dell'IA, se essa debba simulare, emulare, riprodurre o coadiuvare l'intelligenza umana; inoltre è aperta e controversa la discussione sulla natura dell'intelligenza naturale, che dovrebbe, in un modo o nell'altro, essere una sorta di modello per la ricerca. Questo argomento, in particolare, è molto fecondo, in quanto si distingue a sua volta in molteplici riflessioni che, per sommi capi, riguardano: coscienza, intenzionalità e consapevolezza, anima, natura della mente, creatività e fantasia, irrazionale e astratto.

Altro problema aperto è la definizione della natura e del ruolo del linguaggio, una delle mete più ardue per l'IA, in quanto implica l'analisi di processi a tutt'oggi ignoti, quali la comprensione del testo scritto e l'elaborazione del significato di un discorso, sia proferito che ascoltato, come unità.

Inoltre troviamo le difficoltà che pone la matematica con il suo studio dei sistemi formali e degli algoritmi, a cui per ora si riconduce ogni computer, le cui implicazioni non sono a tutt'oggi ben chiare.

Infine dobbiamo menzionare le problematiche riguardanti la stessa Informatica e lo sviluppo pratico delle elaborazioni teoriche.

Lungi dal voler trovare una soluzione a questi quesiti, proveremo ad esporne i concetti fondamentali per offrire una panoramica non sulle conquiste dell'IA dal 1956 ad oggi, ma sulle domande, spesso di fondamentale interesse non solo in campo informatico, che mantengono viva e in crescita questa disciplina.

Procederemo dunque stilando un elenco di questioni aperte ed analizzando, per quanto possibile, le cause delle difficoltà nella loro soluzione, le proposte a tutt'ora presentate e la loro influenza sull'effettivo sviluppo dell'IA :

- Che cos'è l'intelligenza ?
- Quali obiettivi può porsi una ricerca in IA ?
- Quale rapporto esiste tra mente e cervello ?
- Possono le nuove scoperte sul caos e sulla fisica quantistica influenzare lo studio della mente ?
- Quali legami possono stabilirsi tra macchina e mente ?
- La mente è algoritmica ?
- Che rapporto possono avere calcolabilità e logica con i processi mentali ?
- E con l'IA ?
- Che cos'è la creatività ? Può essere *implementata* su una macchina ?
- Come conosciamo? Può una macchina conoscere ?
- Quale importanza ha il linguaggio nel processo conoscitivo ?
- Che cosa sono la coscienza, la consapevolezza, l'intenzionalità ?
- Come è possibile rilevarle in una macchina ?
- E' possibile un'autocoscienza meccanica ?
- L'intelligenza in una macchina deve godere di statuto giuridico ed etico ?

Questi problemi riguardano principalmente due aree, la MENTE e la MACCHINA, in cui saranno divisi i singoli capitoli riguardanti le questioni principali, che possono essere ridotte a :

1. Natura dell'Intelligenza
2. Rapporto mente-cervello e mente-macchina
3. Natura e Processi della Conoscenza
4. Autocoscienza

NATURA DELL'INTELLIGENZA

La trattazione di questo argomento è indispensabile al fine di un discorso sull'IA, poiché gran parte della ricerca è stata ed è tuttora profondamente influenzata dalle opinioni che i ricercatori a capo dei vari progetti hanno in proposito. Con ciò non vogliamo tanto insinuare che gli scienziati coinvolti in questo settore si lascino guidare dai propri pregiudizi; quanto piuttosto desideriamo sottolineare il ruolo fondamentale giocato da quella che può essere definita la *filosofia ingenua degli scienziati*: questa è la tendenza ad impostare il proprio lavoro su convinzioni filosofiche, non sempre esplicitamente dichiarate, non frutto di una consapevole riflessione filosofica, ma di una personale sensibilità strutturata in una certa precomprensione delle cose. Quello che ora è stato riferito unicamente alla filosofia, sarà poi esteso anche ai contributi di altre discipline che studiano il problema dell'intelligenza. Il concetto sarà più chiaro in seguito, analizzando le teorie di alcuni esperti alla luce delle riflessioni e degli studi che sono stati condotti in filosofia, psicologia, neurofisiologia, logica a proposito della natura dell'intelligenza.

MENTE

L'intelligenza è sempre stata il tratto distintivo dell'uomo, che lo distingue dagli altri animali (o "bestie") e lo pone in una posizione privilegiata di dominio sulla natura. E per questo, per lungo tempo, non si è sentito il bisogno di darne una definizione, forse avvertendo più o meno inconsciamente le difficoltà insite in un simile proposito. Vediamo ora alcuni approcci alla questione dell'intelligenza ed esaminiamone i tratti salienti e le aporie che lasciano senza risposta o addirittura sollevano.

L'animale razionale

Aristotele, nella celebre definizione di essere umano come "animale razionale", pone l'essere animale come *genere prossimo* e l'essere razionale come *differenza ultima* dell'uomo. Ciò significa che il *plus* di cui l'uomo gode rispetto agli animali è, secondo lo Stagirita, la razionalità ovvero sia la capacità di apprendere, di dedurre e di argomentare. Egli non dice "animale intelligente", cioè animale che comprende, ma che ragiona. Questa definizione è intimamente legata al concetto di *logos* (appunto "ordine, ragione") e quindi di logica ("ciò che è inerente al *logos*"); ad essi dunque si ricollega Aristotele nel definire l'uomo e quindi nel caratterizzare l'intelligenza.

Dunque, la questione principale è se l'intelligenza possa veramente essere ridotta all'esercizio della ragione e della logica, derivando tutte le altre forme di pensiero da un'ipotetica evoluzione degli istinti animali. Posta in questi termini, questa potrebbe apparire una domanda retorica alla quale si è tentati di rispondere immediatamente NO, più per senso della dignità umana che per motivi plausibili; eppure la posizione di Aristotele, e con lui di buona parte della greicità, non è così assurda, se si pensa la considerazione in cui venivano tenute nell'antichità la matematica e la filosofia, ovvero sia le scienze in cui più emergevano le capacità logiche¹.

¹ Questo è vero per la filosofia soprattutto per quanto riguarda l'opera di Aristotele. Una simile posizione era ovviamente diffusa nel mondo greco anche prima dello Stagirita, ma egli ha contribuito a chiarificarne alcuni aspetti. Infine ricordiamo che è sempre molto azzardato operare simili semplificazioni di vedute in un ambito così vasto come la filosofia greca

Inoltre emerge un'altra difficoltà: la struttura fondamentale del pensiero è razionale o *pre-razionale*², il nostro consueto pensare si basa su processi deduttivi (a cui noi siamo abituati e che quindi non notiamo) o su processi intuitivi e immediati (occasionalmente sostituiti da deduzione e inferenza quando sia necessario) ?

Infine bisogna prestare attenzione ad un particolare abbastanza significativo: abbiamo finora identificato il pensare con l'essere intelligenti, dunque pensiero e intelligenza sono la stessa identica attività ? O vige tra i due un qualche ordinamento gerarchico ?

Pensiero e Intelligenza

Sembra a questo punto il caso di definire con chiarezza che cosa vogliamo intendere (in questa sede) con i termini pensiero e intelligenza, contrapposti nel paragrafo precedente. Per l'intelligenza varrà anche per noi, nelle sue linee generali, la definizione che ne diede Aristotele, distinguendone tre fasi che diremo *concetto, giudizio e argomentazione*, ovvero sia la capacità di conoscere e comprendere nelle loro caratteristiche e relazioni le "cose".

Queste tre fasi, o più correttamente *operazioni* (parola che assume importanza nel definire la visione aristotelica dell'intelligenza in un contesto moderno) si inseriscono in una sistemazione più ampia dei processi mentali in ambito *conoscitivo*, dove troviamo che la mente *percepisce sensorialmente, spiega razionalmente e comprende intellettualmente*; in questa distinzione l'accento è posto sui livelli a cui si svolgono le attività mentali, distinguendo rigorosamente le conoscenze da esse ricavate: su un primo livello troviamo la conoscenza sensoriale, ottenibile attraverso i sensi e consistente di percezioni, ovverosia la risposta alla domanda "Che cosa è?" (*ratio quae* o *quid*); quindi vediamo comparire il tentativo di spiegare la realtà, di cui abbiamo avuto percezione, attraverso la ragione, cioè la scienza, che si occupa delle cause contingenti e del modo in cui esse si esplicano e risponde alle domande "Da che cosa è causato? In che modo?" (*ratio quia* o *quomodo*); infine la suprema attività dell'intelletto è la conoscenza delle cause prime e ultime della realtà, la comprensione dell'universale e del necessario, al di là del particolare e del contingente, ambito tipicamente riservato alla filosofia, che cerca risposta alla domanda "Perché esiste? In forza di che cosa esiste?" (*ratio propter quid*).

All'interno di queste attività, come conseguenza, ma anche come sottoprocesso, si collocano le tre operazioni suddette: l'*apprensione concettuale*, il *giudizio intellettuale*, il *ragionamento razionale*; esse rimandano non solo a tre differenti profondità dell'attività intellettuale, ma anche a tre diverse considerazioni sul valore della conoscenza, vale a dire sul giudizio (termine questo non legato al precedente) *gnoseologico* che possiamo applicare ai tre tipi di conoscenza che derivano da queste tre operazioni: il concetto può essere *adeguato o meno*, il giudizio può essere *vero o falso*, il ragionamento *corretto o scorretto*.

Da quel poco che abbiamo esposto della concezione aristotelico-tomistica emerge chiaramente che l'intelligenza è definita essenzialmente in base al comportamento : noi *formuliamo* o *esprimiamo* o *comunichiamo* concetti, giudizi e argomentazioni; inoltre l'intelligenza è esaminata unicamente nella sua simbiosi con la conoscenza, alla quale certamente essa fornisce un importante contributo, ma dalla quale potrebbe essere distinta con profitto.

Riferendoci poi al pensiero vorremo intendere tutte le attività mentali che costituiscono le capacità che abbiamo appena ascrivito all'intelligenza: la capacità di apprendere concetti, formulare giudizi e costruire ragionamenti è legata alla comunicazione della conoscenza (o quantomeno ad una sistemazione in una forma logica del sapere), ma la mente non opera solo su questo piano di deduzione-induzione-abduzione, essa passa attraverso meccanismi di collegamenti casuali (o comunque prefazionali) e di passaggi intuitivi di cui il sistema concetto-giudizio-argomentazione è l'espressione finale a scopo di comunicazione o di cosciente archi-

² Abbiamo preferito questo termine al consueto *irrazionale* per intendere lo stato embrionale o potenziale di razionalità non come caos, disordine, imprecisione, ma come la dipendenza da processi intuitivi o immediati.

viazione della conoscenza. Tutti i meccanismi, i processi, che si svolgono nella mente, che siano o meno sotto il controllo della consapevolezza, noi li chiameremo pensiero, chiaramente definendo l'intelligenza come quella parte di pensiero che opera al fine di dare una sistemazione in veste logica al sapere.

Abbiamo così introdotto, quasi di nascosto, il tema, fondamentale, della consapevolezza (a cui dedicheremo una trattazione approfondita più avanti) come differenza significativa tra il pensiero in generale e l'intelligenza in particolare. Questo prelude ad una domanda di grandissima importanza per un discorso sull'IA : è possibile intelligenza senza consapevolezza ? Ci preme inoltre sottolineare che le brevi considerazioni svolte sul rapporto intelligenza-pensiero rispondono solo parzialmente alle domande poste al termine del precedente paragrafo; infatti l'introduzione della consapevolezza come elemento che trasforma il pensiero in intelligenza è solo una delle possibili risposte e non è in sé completa né esauriente; inoltre il rapporto che intercorre tra i due non è riducibile ad una semplice inclusione, ma deve essere esplicitato meglio, considerando soprattutto che finora abbiamo dato una definizione dell'intelligenza strettamente teoretica e presa a prestito dalla filosofia. Altre scienze cosiddette "della mente", quali la psicologia, danno altre definizioni, più utili al loro campo di studi, giungendo ad esempio a definizioni operative, inerenti l'azione dell'uomo nel mondo, contrapposte alla definizione teoretica precedentemente fornita, che riguarda la conoscenza dell'uomo a proposito del realtà.

L'intelligenza e l'esperienza

Un comportamento intelligente è spesso associato all'esperienza, alla saggezza ed è comune equiparare l'aggettivo "intelligente" a "esperto", quando si tratta di campi specifici o di questioni tecniche. Riflettiamo un momento sulla parola "esperto": deriva dal latino *experior* che significa sperimentare, provare, che a sua volta viene dal greco *peira* che significa fare la prova e contiene una radice che significa "oltre, al di là" e contemporaneamente "attraverso, passando all'interno". Se invece ci concentriamo sul termine "intelligenza" troviamo degli elementi significativi già nell'etimo latino: *inter-legere* ma anche *intus-legere* che significano leggere-tra e leggere-dentro cioè trovare le relazioni tra le cose, ma anche trovare la natura, comprendere l'essenza delle cose; si tenga inoltre presente che *legere* in latino aveva anche il significato di raccogliere, estrarre e quindi possiamo mettere in relazione intelligenza (estrarre il significato dall'interno delle cose e comprenderne le relazioni) con esperienza (aver provato la realtà passandovi attraverso e conoscere per questo come agirvi) in quanto entrambe attinenti al penetrare all'interno dell'oggetto di conoscenza, sebbene con una differenza da sottolineare: l'intelligenza è (in questa determinata accezione) la comprensione della natura più profonda delle cose nel loro essere e nel loro agire e interagire, l'esperienza riguarda il relazionarsi alle realtà in campo pratico.

Possiamo a questo punto azzardarci a collocare l'esperienza come una fonte alternativa di conoscenza, ovviamente legata, nel suo manifestarsi pratico, ad un modo dell'intelligenza, quella che chiameremo *intelligenza adattativa*, ovvero la capacità di imparare dall'esperienza.

Questo particolare modo del conoscere è utile per operare una distinzione tra comportamento intelligente e comportamento genetico (o istintuale) negli animali: alcune specie compiono azioni che, per un uomo, sarebbero segno di intelligenza e comprensione della realtà, come ad esempio le dighe dei castori, i formicai, alveari e termitai degli insetti, l'uso di "strumenti" e altro; in simili casi si sarebbe propensi ad attribuire una sorta di intelligenza a questi animali, ma basta provare a cambiare alcune condizioni ambientali o interferire con queste operazioni, magari cambiando la posizione di un qualche elemento, e si vedrà, ripetendo l'intervento più volte, che la reazione sarà sempre la stessa (solitamente l'animale riprende il lavoro da una fase precedente e ripete azioni già fatte, senza motivo).

Solo alcune specie più vicine all'uomo geneticamente (solitamente tra i primati³) hanno la capacità di imparare dal passato e cambiare quindi il proprio comportamento in funzione di quel che è accaduto. Questa capacità però non si avvicina nemmeno lontanamente alla profonda abilità della mente umana di confrontare ad ogni momento la previsione (o il desiderio) con la realtà e di analizzare le eventuali discordanze rilevandone le cause.

Le difficoltà emergono quando si tenti di indagare questo processo: in sé esso è molto simile al metodo sperimentale che sta alla base della scienza moderna, con la differenza che, mentre in campo scientifico la riflessione sui motivi dell'eventuale discordanza tra teoria e dati reali è condotta coscientemente, il confronto tra aspettativa/previsione e il fatto accaduto è per buona parte svolto non coscientemente, in *background* per usare un termine informatico.

E' quindi molto complicato definire l'effettivo svolgersi di questa comparazione, i cui risultati vengono poi presentati e utilizzati dall'intelligenza adattativa. Secondo la distinzione introdotta prima, possiamo dire che la valutazione è affidata ad un processo del pensiero che entra nella sfera dell'intelligenza solo nella sua ultima fase di riconoscimento e applicazione delle conoscenze trovate.

Del resto, si potrebbe dire che ogni apprendimento è legato ad un adattarsi della mente all'esperienza fatta, seppure non quel peculiare procedimento di aspettativa-tentativo-risultato-esame delle differenze; per questo motivo c'è chi identifica la capacità di apprendere con l'intelligenza adattativa⁴. Non assumeremo questo punto di vista, non tanto per motivi di tipo scientifico o psicologico, quanto per la facilità di trattazione che segue dal distinguere la capacità di imparare ad interagire in maniera migliore con la realtà, dalla più generica abilità di apprendere concetti e informazioni, collegandoli e applicandoli alla realtà.

Dunque troviamo alcune questioni aperte : l'esperienza è compresa totalmente nell'intelligenza in un modo non ancora compreso o effettivamente esiste una fase di "elaborazione" precedente all'intelligenza ? E' possibile che l'apprendimento generico, pur distinto come detto dall'intelligenza adattativa, ne segua gli schemi anche solo in modo generico? Dunque possiamo stabilire un legame diretto tra sapere e saper fare senza mediazioni oppure serve un termine mediante ?

Inoltre l'esperienza ha forti legami con quella facoltà della mente che chiamiamo memoria, su cui le domande sono ancora di più e più sostanziali, per le quali rimandiamo alla parte dedicata alla natura ed ai processi della conoscenza, di cui la memoria è il principale strumento.

La creatività e l'invenzione

Prima di affrontare il tema proposto dall'intitolazione del paragrafo, riteniamo opportuno parlare di *intelligenza creativa* che, in psicologia (soprattutto trattando dell'apprendimento), viene di solito contrapposta all'intelligenza adattativa di cui abbiamo trattato finora. Non va confusa con la creatività, con cui ha comunque forti legami.

La fase successiva all'apprendimento è di solito la verifica della correttezza dei contenuti appresi, solitamente operata attraverso test e prove (più o meno sperimentali); quando però l'interrogazione sia autonoma e verta su questioni che non hanno risposta nota o facilmente derivabile (in base a quanto appreso), usciamo dal campo dell'apprendimento e giungiamo all'intelligenza creativa: essa consiste nel mettere in discussione e vagliare criticamente i contenuti appresi, cercando di ampliarne la portata o ricavarne generalizzazioni o varianti. In pratica, tutto ciò consiste in domande quali : "Perché è necessaria questa condizione? Cosa accadrebbe se...?".

³ Per equità, bisogna dire che anche i topi, i cani, i cavalli e i marsuini presentano capacità adattative.

⁴ Si veda ad esempio "Mente umana e cervelli elettronici", Donald G. Fink, Ed. Zanichelli 1967 in cui peraltro si può trovare una buona trattazione dell'intelligenza apparente di alcuni animali e una conseguente definizione di intelligenza naturale.

L'intelligenza creativa è tipica dell'uomo ed è all'origine della nascita delle scienze: i primi tipi di conoscenza dell'uomo sono stati la sensazione e l'esperienza, su di esse l'uomo ha cominciato ad interrogarsi, mutando il contesto e cercando di generalizzare quello che nel suo ricordo riguardava situazioni con caratteri simili. Da qui hanno avuto origine riflessioni sulla natura, sul numero, sulla misura, originando la fisica, l'aritmetica e la geometria, che più avanti sarebbero sbocciate nella diade Filosofia e Matematica che ha segnato la storia della Grecia, dal medioevo ellenico al periodo ellenistico.

Questo processo appartiene interamente alla sfera della coscienza e quindi dell'intelligenza, ma altrettanto difficile da descrivere del precedente: la difficoltà non risiede nell'identificare il processo tramite cui si dà risposta alle domande suddette, poiché lo schema generale è abbastanza semplice e riguarda l'induzione e l'inferenza, mentre lo schema particolare è troppo variabile per avere una descrizione sintetica; quello che pone problemi (e non ce lo si era aspettato) è il determinare quali domande porsi e su quali linee di ricerca insistere. Capita spesso che non si sappia proseguire nello sviluppare autonomamente un argomento, ma di solito ci si pone domande corrette e che meritano risposta.

In parallelo a questo processo di interrogazione nelle scienze, troviamo in campo artistico la creatività o invenzione. Certo anche nelle scienze, qualora le domande poste necessitino di idee a tal punto innovative da implicare l'introduzione di nuovi concetti fondanti e nuove simbologie, si può parlare di creatività, invenzione o scoperta a seconda di quale corrente di pensiero si scelga di seguire a proposito della natura della Scienza e in particolare quando si tratti di Matematica, dove il conflitto tra le varie posizioni filosofiche è molto sentito dall'inizio del XX secolo.

Il momento creativo-inventivo è sempre stato avvolto da una sorta di aura: nell'antichità, l'origine di una idea nuova, creativa era talmente inspiegabile che veniva attribuita all'intervento di divinità, le Muse, che rappresentavano la fonte per l'ispirazione di ogni artista. Anche in età moderna, quando un artista viene chiamato ad esprimersi sulla propria attività inventiva, fornisce risposte insoddisfacenti e parziali, oppure allusive e poco chiare; spesso questi sostiene che l'idea esistesse già e che egli abbia avuto solo il merito di raggiungerla, come dichiarava Michelangelo a proposito delle proprie statue o Mozart delle proprie composizioni. Altre volte, l'attività creativa è vissuta dall'artista in uno stato di semincoscienza, alla maniera di Schubert, di Dalì o di Pollock, seppure in modi differenti. Infine c'è chi concepisce il creare come un paziente lavoro ragionato e logico su un materiale "grezzo", che sarà uno schizzo, una melodia o un canovaccio; questo è solo un modo di spostare il problema, che allora riguarderà l'origine di quel materiale "grezzo". Vi sono artisti estremamente prolifici, mentre altri hanno compiuto solo poche opere; per alcuni l'arte era un tormento, per altri un sollievo; alcuni hanno continuato a lavorare dalla gioventù alla vecchiaia inoltrata, fino agli ultimi istanti, altri si sono consumati in pochi anni creativi, seguiti da lunghi "silenzi", altri ancora hanno creato, in pochi anni di vita, centinaia di opere; tutti, seppur vivendola in diversi modi, sono arrivati a dare della creatività una stessa identica descrizione, che può essere riassunta efficacemente nelle parole del pittore P. Klee : *"La mia mano è uno strumento guidato da una sfera remota e non è la mia testa che opera nel mio lavoro; è qualcosa di diverso..."*.

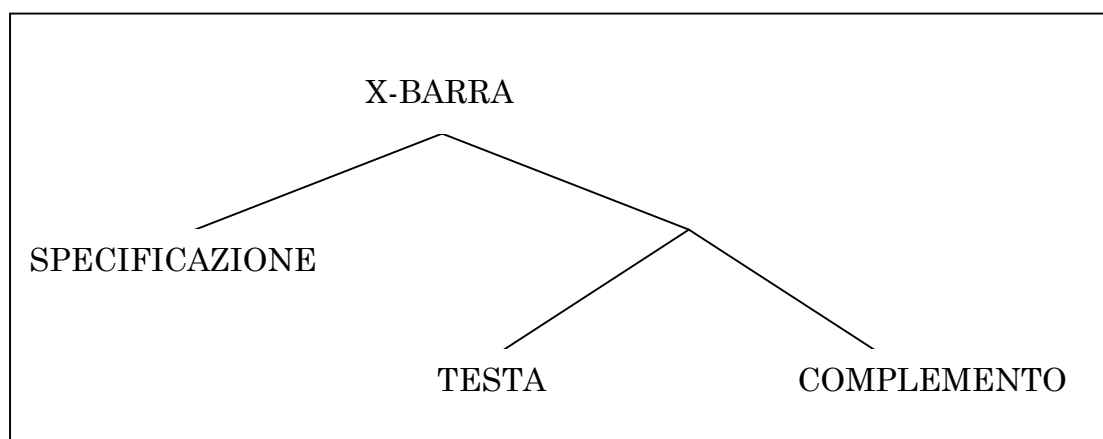
Dunque la creatività è l'aspetto che presenta fin qui maggiori difficoltà, pur non lasciando dubbi sulla sua partecipazione a ciò che chiamiamo "intelligenza". Le domande sono molteplici : dove comincia la creatività e dove finisce l'intelligenza creativa? Questo confine esiste effettivamente? Qual'è l'origine delle nuove idee? Cosa permette a certi individui di avere una tale capacità creativa da essere definiti artisti? Come si distingue una "buona" idea da una "cattiva"?

Il linguaggio e la lingua

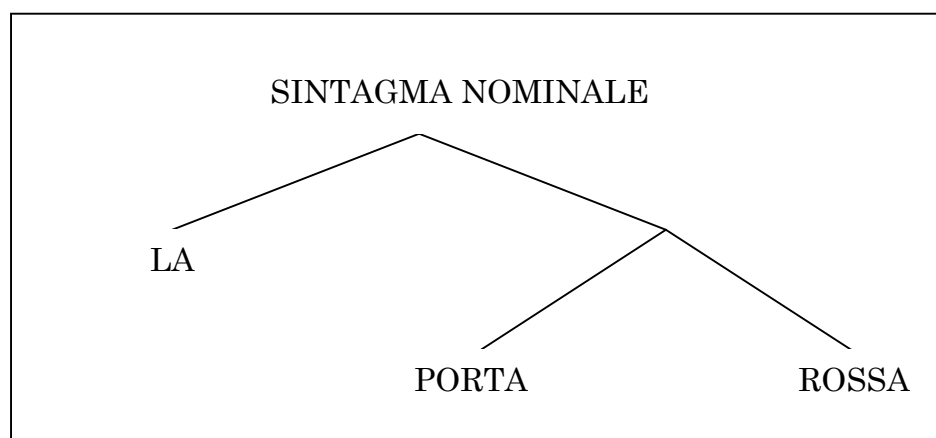
Non è un caso che il delfino, uno dei mammiferi più intelligenti, abbia probabilmente sviluppato una sorta di linguaggio per comunicare con i propri simili anche dati complicati e situazioni complesse; il linguaggio è certamente una delle caratteristiche che più decisamente differenziano l'uomo dagli altri animali ed è quindi intuitivamente collegabile con l'intelligenza. Proprio per questo il linguaggio è uno dei misteri più profondi dell'intelligenza umana ed è stato per questo collegato (e non a torto, se consideriamo i risultati ottenuti) con l'altra caratteristica "forte" esclusiva dell'uomo, la capacità di astrazione.

Il problema fondamentale del linguaggio è spiegare una simile abbondanza di strutture, relazioni, significanze in base a pochi elementi che siano comuni ad ogni lingua e riescano ad evitare che nel nostro cervello debbano essere immagazzinati miliardi di possibili strutture di frase e collegamenti semantici.

Una delle teorie più fondate e accettate sulla natura dell'*intelligenza linguistica*⁵ è stata proposta da Noam Chomsky in *Syntactic structures* (1957) e viene denominata *teoria della x-barra*; quest'ultima è una struttura ternaria in cui si dispongono una testa, una specificazione ed un complemento in un rapporto che non è di subordinazione, ma di consequenzialità logica e vicinanza o parentela semantica. Questa struttura permette di organizzare, ad esempio, un sintagma nominale di tipo semplice, quale *la porta rossa*:



**Fig 1. -
X-barra
generica**



**Fig 2. - X-
barra rappre-
sentate il sin-
tagma nomi-
nale *la porta
rossa*.**

Ovviamente ogni posto può essere occupato da una struttura simile, ottenendo una x-barra ramificata e costruendo un periodo o anche un intero discorso.

⁵ Così definiamo quella "parte" dell'intelligenza umana che si occupa del linguaggio, della sua organizzazione e della sua espressione.

Questa teoria propone anche una sistemazione del significato, secondo uno schema simile, ma la sua concezione originale era destinata unicamente a chiarire la struttura del linguaggio nella sua organizzazione sintattica.

Inoltre troviamo un'altra importante materia di discussione nella linguistica (ovvero la scienza che studia il linguaggio) che riguarda l'origine, l'evoluzione e le motivazioni⁶ del linguaggio; apripista fu sempre Chomsky che nel 1975 pubblicò *Reflections on language* in cui tratta le varie ipotesi sulla nascita del linguaggio umano, i modi in cui esso si è modificato e specializzato e l'importanza che questo ha avuto nella storia della specie umana. Per esempio citiamo, senza scendere nel dettaglio, l'ipotesi dell'*Eva linguistica*, ovvero un singolo antenato di sesso femminile vissuto 15.000-20.000 anni fa che avrebbe per prima sviluppato il linguaggio cosiddetto simbolico (che possiamo chiamare, per via allusiva, evitando spiegazioni che ci porterebbero a divagare eccessivamente, astratto).

Si tenga presente, inoltre, che finora abbiamo considerato solo il problema sintattico e non l'aspetto semantico del linguaggio: il significato dei termini di ogni lingua ha un legame stretto con la conoscenza che abbiamo e che *possiamo avere* del mondo che ci circonda.

Nei due lavori precedentemente citati (ed in particolare nel secondo) Chomsky arriva a suggerire un profondo legame tra l'intelligenza e le capacità del pensiero umano, vedendo nello sviluppo del linguaggio il momento in cui la mente umana ha compiuto il "grande salto" che ci separa dalle scimmie antropomorfe. Tanto per citare un esempio, le teorie di Chomsky sul linguaggio sono uno dei principali argomenti che usa Keith Devlin ne "*Il gene della matematica*" per spiegare l'origine delle capacità matematiche.

Il linguaggio sembra avere dunque un ruolo fondamentale nel definire l'intelligenza.

Qui occorre allora introdurre una specificazione: un linguaggio deve essere in grado di esprimere concetti, giudizi, ragionamenti in maniere quanto si vuole complessa, senza limiti teorici. Per questo non faremo rientrare sotto questo nome i *linguaggi animali*, i *linguaggi informativi*, il *linguaggio del corpo*, i vari *linguaggi in codice*, qualora non siano solo una trascrizione del linguaggio naturale.

Inoltre è doveroso accennare, per ragioni di completezza, all'ipotesi del *mentalese*, un linguaggio originario in cui si esprime la mente di ogni essere umano, che viene poi tradotto nel linguaggio parlato, fornendo un tramite tra mente e lingua e risolvendo diverse aporie; appunto per questo l'esistenza di un tale "linguaggio della mente" è messa in dubbio, mancando ogni sostegno sperimentale o plausibile, se non la comodità proprio di un'ipotesi *ad hoc*.

Dunque possiamo trovare innumerevoli questioni, tra cui: quale rapporto esiste tra linguaggio e intelligenza? Come è strutturato il linguaggio? Come è nato? In che modo influisce sulla nostra comprensione della realtà?

Rimane ancora da trattare il tema della *lingua*, ovvero del particolare insieme di segni o suoni attraverso cui un linguaggio si esprime. Se è vero che la struttura di base è identica in ogni uomo, gli elementi che la vanno a riempire sono assai diversi, non ovviamente nel ruolo, ma nel significante. Non è azzardato dire che la lingua influenza pesantemente i processi di conoscenza, determinando quali concetti e idee siano facilmente esprimibili e quali non lo siano. Per queste ragioni sono nati i linguaggi settoriali, che in realtà sono sotto-lingue, e per questa ragione è così difficile tradurre un'opera di valore letterario da una lingua all'altra. Del resto, capita non di rado di ravvisare alcune caratteristiche stereotipate di un popolo in aspetti particolari della loro lingua.

Anche qui possiamo porci varie domande sul ruolo della lingua nei processi conoscitivi e sulle relazioni della lingua con l'intelligenza, come nel caso precedente.

⁶ L'utilità evoluzionistica del particolare tipo di struttura linguistica e mentale che noi oggi possediamo.

MACCHINA

La panoramica che abbiamo svolto ha riguardato prevalentemente la filosofia, la linguistica e la psicologia e ci ha fornito una base concettuale per ora e nel nostro caso sufficiente, quantunque essa non esaurisca affatto gli argomenti proposti, ad affrontare l'esposizione degli studi che nel campo di IA sono stati svolti sulla natura dell'intelligenza e su come possa essere sintetizzata, suddivisa, organizzata e quindi riprodotta.

In questo studio, che è stato intrapreso fin dagli inizi dai pionieri di tale disciplina, troveremo la necessità di introdurre altre concezioni psicologiche o filosofiche, che non hanno avuto separata trattazione nelle pagine precedenti a causa della loro grande complessità e dei loro molteplici legami con altre parti delle scienze della mente, che ci avrebbero portato troppo lontano dall'argomento della nostra esposizione. Pertanto, di esse verrà fornita una esposizione "intuitiva" (se così possiamo dire), ovvero senza rigore terminologico o formale, qualora non sia espressamente richiesto, lasciando in nota le indicazioni bibliografiche per ulteriori approfondimenti.

Possiamo fin da ora anticipare che molte delle domande esposte al termine dei precedenti paragrafi non troveranno risposta in questa seconda sezione ed anzi ad esse se ne aggiungeranno altre, suggerite dal confronto con l'informatica e la cibernetica.

Il Test di Turing e la macchina intelligente

Nei primi anni di sviluppo dell'IA era necessario un metro di giudizio per determinare quando un computer potesse essere detto "intelligente"; spinti da tale necessità, gli studiosi di tali problemi trascurarono le questioni di natura psicologica (e in parte anche filosofica) che una simile definizione operativa comportava e salutarono con entusiasmo un articolo del celebre matematico Alan Turing, intitolato *Computing Machinery and Intelligence* apparso sulla rivista "Mind" nel 1950. In questo articolo vide la luce il test di Turing, che avrebbe guidato la ricerca in IA fino alla fine degli anni Settanta.

Turing proponeva questo modo per verificare l'intelligenza di un computer: poniamo un essere umano con una vasta cultura (nel ruolo di giudice) in una stanza con due monitor e due tastiere; i monitor e le tastiere sono collegati ad apparecchiature simili in due stanze separate. In una si trova un essere umano, nell'altra un computer. Il giudice umano può, tramite le tastiere, porre domande all'occupante dell'una o dell'altra stanza e può vedere le sue risposte sul corrispondente monitor, allo scopo di determinare in quale stanza si trovi un computer. Se non potrà prendere una decisione o addirittura sbaglierà nel giudicare quale sia la macchina, il computer in questione sarà detto intelligente. Il test di Turing è, da questo punto di vista, la prova di una vera e propria teoria scientifica per la quale vale il principio di falsificazione di Popper.

Dietro questo procedimento che può anche apparire eccessivamente macchinoso, si nasconde una precisa convinzione psicologica, che, semplificando un poco, può essere ricondotta alla scuola del *Behaviorismo*⁷ ("comportamentismo") che presta attenzione in particolar modo alle corrispondenze di stimoli e risposte. In sostanza la convinzione di Turing era la seguente: *"un computer è intelligente quando il suo comportamento è tale che, se fosse attribuito ad un essere umano, guadagnerebbe a questo un riconoscimento di intelligenza"*.

Altra teoria psicologica che ha fortemente influenzato lo sviluppo dell'IA è la reazione al Behaviorismo, ovvero il *cognitivismo* che attinge a piene mani dalla Cibernetica e

⁷ La scuola behaviorista si è sviluppata agli inizi del secolo XX in America, mantenendo una posizione di egemonia nello studio della Psicologia fino ad oltre la metà del '900 (periodo degli scritti di Turing in proposito), nonostante illustri psicologi abbiano operato al di fuori di esso anche in quel periodo (ad es. J. Piaget). La convinzione principale di questa corrente è la sostanziale non-realtà di tutta quella parte dell'esperienza umana che non è direttamente osservabile, come la coscienza e la mente.

dall'Informatica per trovare metodi tramite cui spiegare le attività mentali che innervano il comportamento osservabile.

Negli anni Settanta il test di Turing è stato posto in secondo piano dall'attenzione che veniva sempre più rivolta al problema della conoscenza e dell'apprendimento, campo nel quale questo test non aveva grande validità. Inoltre, Turing stesso afferma l'assurdità del domandarsi se un computer possa pensare, puntando la propria attenzione sulla capacità di un computer di "ingannare" un osservatore fingendo di essere un essere umano; perciò l'interesse verso i processi di pensiero che si andava destando verso l'inizio degli anni Ottanta non trovava risposta in questo *modus operandi*. Il test venne allora assunto a metodo per determinare anche la presenza di conoscenza e pensiero, non solo di comportamento intelligente, divenendo parte della cosiddetta IFIA (ipotesi forte dell'intelligenza artificiale) di cui tratteremo più ampiamente in seguito.

Pochi anni dopo il filosofo J. Searle iniziò a diffondere le proprie idee di critica all'ipotesi forte, affermando che il successo nel test di Turing non implicava né pensiero, né conoscenza, né quindi intelligenza, intesa però in senso cognitivista, seppur con qualche modifica; egli propose l'esperimento mentale della "stanza cinese": se poniamo un italiano in una stanza con il materiale per scrivere e un manuale in italiano su come manipolare i simboli per dare delle risposte alle domande (in cinese) che gli verranno presentate per iscritto attraverso una feritoia, egli probabilmente darà tutte risposte corrette, ma di certo non potremo affermare che egli sappia il cinese.

Questa conclusione è viziata da un particolare che Searle non prende in considerazione e che esamineremo più avanti.

Questo particolare argomento non presenta nuovi interrogativi, ma semplicemente mostra come la già nominata filosofia ingenua degli scienziati influenzi il loro procedere, soprattutto in campi per così dire "confinanti" con filosofia e psicologia. Infatti, dagli anni Settanta gli informatici hanno cominciato a studiare più approfonditamente le varie teorie psicologiche sulla natura dell'intelligenza e della conoscenza, trovando varie affinità con il già nominato *cognitivism*. L'unica questione aperta e non ancora sollevata è se esista una prova con cui determinare se una certa macchina sia o meno intelligente.

Algoritmi, logica e calcolabilità

Poiché un criterio operativo per stabilire l'intelligenza non è ancora stato trovato, né si intravedono speranze in un futuro prossimo di trovarlo, l'IA ha spostato i propri sforzi su un campo teorico, analizzando le componenti funzionali fondamentali di un computer e cercando una possibile corrispondenza con i meccanismi e i dinamismi del pensiero umano.

Il primo e più evidente elemento su cui si basa il funzionamento di un calcolatore è l'idea di algoritmo⁸, derivata dalla matematica, nella quale indica la sequenza di operazioni da eseguire per raggiungere un certo risultato. Il termine moderno di programma comprende in sé l'idea di algoritmo e quasi nulla altro.

Diamo ora qualche definizione che ci sarà utile in seguito per trattare con più precisione di questi argomenti. Gli algoritmi vengono tradizionalmente divisi in due categorie in base alla caratteristica della ricorsività; un algoritmo è ricorsivo quando una delle istruzioni che lo compongono prevede l'esecuzione dell'algoritmo stesso, come ad esempio succede nello scomporre un numero in fattori primi: per ottenere questo scopo potremmo scrivere un algoritmo quale quello in figura 3.

⁸Il termine deriva dal nome di un matematico arabo del IX secolo d.C. Abu Gia'far Muhammad ibn Musa *al-Khwarazmi*, autore, verso l'825 d.C., di un precursore dei manuali di algebra rinascimentali intitolato *Kitab al-Giabr wa'l-muqabalah*.

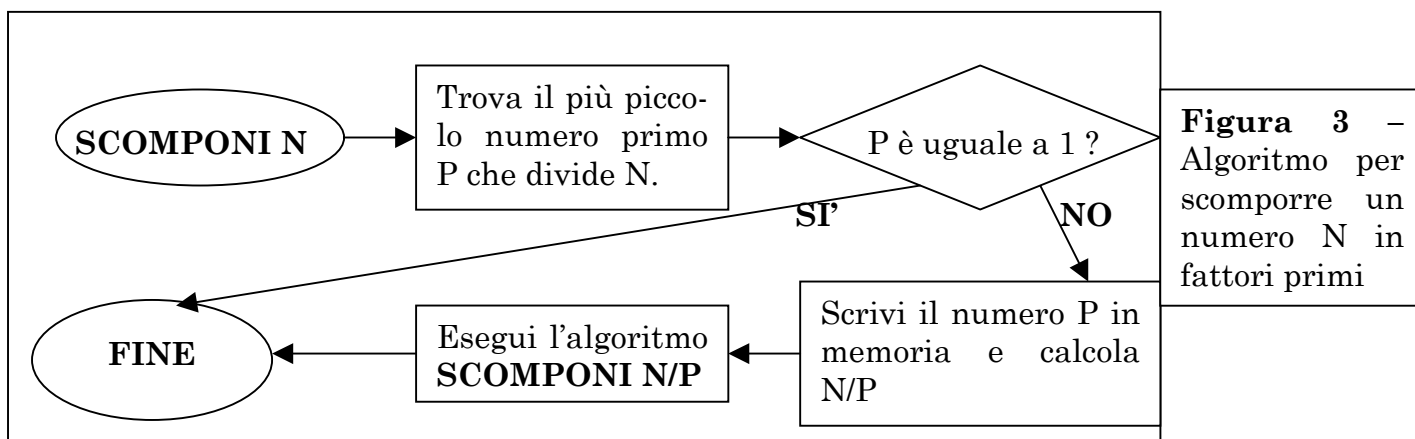


Figura 3 – Algoritmo per scomporre un numero N in fattori primi

In termini informatici, questo programma chiama un'istanza di se stesso fornendo un dato più abbordabile e così via finché il numero fornito a una delle istanze non è primo; allora l'ultima istanza non ne chiama una ulteriore e chiude la propria esecuzione e così tutte fino alla esecuzione originale del programma.

Di fianco a questa definizione, troviamo una sua generalizzazione che è la seguente: un algoritmo è ricorsivo quando una parte delle istruzioni che lo compongono è ripetuta (una o più volte). Più nel dettaglio, si pone anche una distinzione sulla natura della ripetizione (*ciclo* in termini informatici): se la ripetizione avviene un numero prefissato di volte l'algoritmo è detto *primitivo* ricorsivo; se la ripetizione avviene un numero ignoto a priori, ma comunque finito di volte, ad esempio se viene posta una condizione come **RIPETI SE $i > 0$** del cui verificarsi *prima o poi* noi siamo certi, l'algoritmo prende il nome di *generale* ricorsivo (o anche solo ricorsivo); se infine non siamo certi né del numero di ripetizioni né se questo numero sia o meno finito, cioè quando non sappiamo se la condizione che regola il ciclo si verifica prima o poi, l'algoritmo è detto *parziale* ricorsivo.

Un esempio di algoritmo primitivo ricorsivo è l'elevamento al cubo⁹ di un numero n, che consiste nella ripetizione due volte del comando **MOLTIPLICA PER n**; un esempio di algoritmo generale ricorsivo è quello già proposto sulla scomposizione di un numero in fattori primi; un esempio di algoritmo parziale ricorsivo è più complicato da presentare: se definiamo una successione A

$$A_n = \begin{cases} A_{n-1}/2 & \text{se } A_{n-1} \text{ è pari} \\ 3A_{n-1} + 1 & \text{se } A_{n-1} \text{ è dispari} \end{cases}$$

possiamo vedere che per molti valori di A_1 essa conduce dopo un certo numero di passi al valore 1 (ad es. : $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$; oppure $3 \rightarrow 10 \rightarrow 5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1$); ora, vogliamo sapere se per un certo valore di A_1 , la successione giunge a 1 e scriviamo un algoritmo che si fermi appena raggiunge 1. Bene, non possiamo sapere se, per un dato valore, l'algoritmo terminerà o meno¹⁰.

L'appartenenza ad una delle prime due categorie di algoritmi è indicata con il termine *calcolabilità*.

E' qui il caso di fare una breve digressione nella digressione e parlare dei rapporti tra logica, matematica e informatica. Se i matematici, i logici e i filosofi sono essenzialmente concordi nel ritenere che la logica e la matematica siano per molti versi equivalenti, ma non equipollenti, il ruolo dell'informatica è quanto meno dubbio: la logica può sicuramente essere formalizzata, ma in tal maniera si perde quell'aspetto semantico che permette di cogliere immediatamente la necessità di certe premesse e la verità di certe conclusioni, mentre la matematica può essere sottoposta ad una completa revisione logica, ma così facendo si può solo concludere

⁹ O a qualunque altra potenza più alta di 2.

¹⁰ La successione $3n-1$ è tuttora oggetto di studi approfonditi da parte di molti teorici dei numeri, in quanto non si è ancora riusciti a determinare se tutti i numeri giungano prima o poi a 1 e quali eventualmente non vi arrivino. La ricorsività parziale dell'algoritmo non è dunque dovuta ad una sua particolare formulazione, ma alla natura del problema.

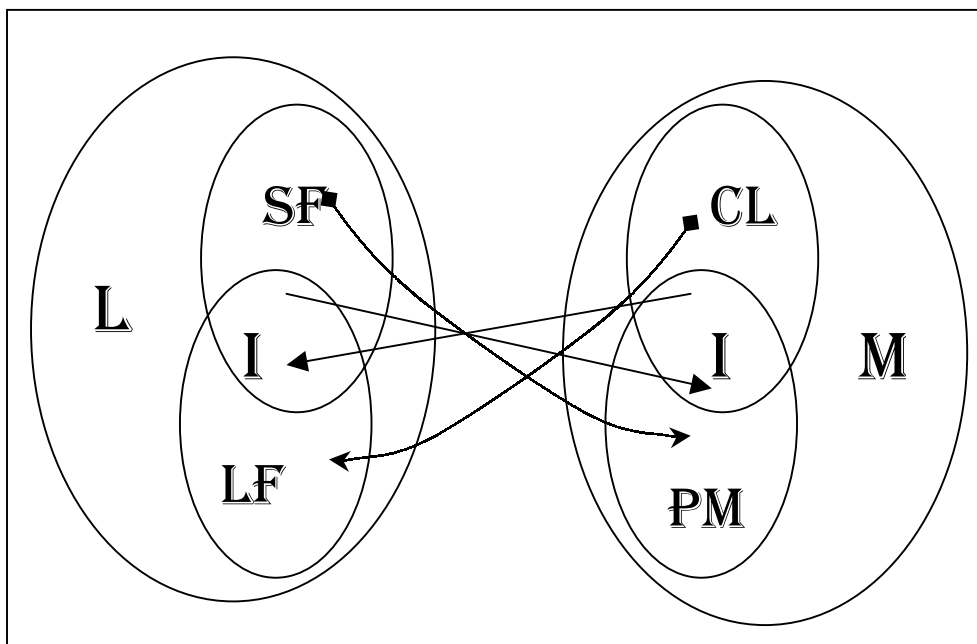


Figura 4 – La logica (L) ammette al proprio interno una parte di sistemazione formale (SF) applicabile ai principi della matematica (PM) in matematica (M). Viceversa in matematica troviamo il calcolo logico (CL) che può rappresentare una parte della logica, la logica formale (LF). La parte comune alle due suddivisioni è l'informatica (I).

che la matematica non ha alcuna ragione di essere così com'è e non in altro modo; d'altra parte l'informatica è sicuramente una applicazione di molte teorie matematiche e di buona parte della logica formale, ma è essenzialmente rappresentabile senza una evidente perdita di senso in entrambe, fatto di cui non si può affermare l'inverso.

Ora, l'algoritmo matematico e informatico è composto essenzialmente da operazioni matematiche o facilmente matematizzabili e da procedure logiche; ed ecco che fa la sua comparsa anche la logica, qui intesa non come logica sillogistica (proposizionale), né come logica booleana, ma come logica formale, ovvero quella teoria del discorso formale che studia il suo svilupparsi in relazione alla verità, alla coerenza ed alla completezza, distinguendone un aspetto semantico ed uno sintattico. Vediamo brevemente questi argomenti.

Innanzitutto, diciamo *sistema formale* un sistema logico di predicati (enunciati logici) che contenga un gruppo di *assiomi*, ovvero predicati assunti convenzionalmente come veri ("Tutti gli angoli retti sono tra loro uguali", IV postulato di Euclide), una *procedura di decisione* che permetta di decidere se un predicato è ben formulato (potremmo intuitivamente dire "se ha senso": un enunciato quale "gli angoli interni di una circonferenza hanno volume pari al quadrato della base" non ha senso geometrico, mentre "la somma degli angoli interni di un triangolo è tre angoli retti" ha senso, sebbene non sia corretta), una serie di *produzioni* o *regole di inferenza*, che permettono di ricavare, applicandole, enunciati veri da enunciati veri ("se $a \in A$ e $A \subseteq B$ allora $a \in B$ "¹¹ o una qualunque altra forma di sillogismo o ragionamento deduttivo, ad esempio).

Un enunciato derivato dagli assiomi tramite le regole di inferenza è detto teorema del sistema formale.

Un enunciato è vero se è ottenibile tramite le regole di inferenza dagli assiomi.

In un sistema formale si parla di coerenza se tra tutti gli enunciati ben formulati derivati dagli assiomi non ve ne siano di reciprocamente contraddittori. Un sistema è completo quando tutti gli enunciati che in esso sono veri sono ottenibili tramite le regole di inferenza.

A questo punto sorge una difficoltà sul concetto di verità : in effetti la verità è un concetto privo di senso se il sistema non viene riferito ad un modello, ovvero sia se agli elementi del linguaggio del sistema non vengono associati elementi di un sottoinsieme della realtà (even-

¹¹ "Se Socrate ($=a$) è un($=\in$) uomo($=A$), se tutti gli uomini($=A$) sono($=\subseteq$) mortali ($=B$), allora Socrate ($=a$) è un ($=\in$) mortale($=B$)".

tualmente sottoposta ad astrazione e generalizzazione), su cui sia possibile eseguire verifica degli enunciati del sistema.

Da questo punto di vista, vale una versione adattata della teoria epistemologica di Popper : gli enunciati logici, quando non “parlano di qualcosa”, quando non sono riferiti (corrispondono biunivocamente) ad oggetti reali, non sono né veri, né falsi, non sono scienza; solo la possibilità di falsificarli concede ad essi la proprietà della verità o della falsità. Ovviamente questo accostamento va preso con cautela, in quanto questa corrispondenza tra la teoria popperiana e il problema della verità non è completa.

Così viene introdotta la distinzione tra l'aspetto sintattico (le questioni riguardanti la forma) e l'aspetto semantico (le questioni riguardanti il significato): la coerenza è sintattica, la verità è semantica, la completezza appartiene ad entrambi i campi, seppure ha una dimensione prevalentemente sintattica, si giustifica semanticamente.

Gli studi sulla natura e le proprietà del concetto di algoritmo sono strettamente legati allo sviluppo ad inizio secolo della suddetta logica formale, grazie a nomi quali Turing, Tarski, Church.

Uno dei problemi più interessanti di questo campo di indagine è determinare se può esistere un algoritmo che riconosca se un enunciato è teorema in un determinato sistema formale¹². Presenteremo ora alcuni risultati importanti, senza scendere nel dettaglio, né tanto meno giustificarli formalmente:

- Teorema di CHURCH-TURING-TARSKI : non esiste una procedura generale ricorsiva per determinare la verità o la falsità di un teorema all'interno di un sistema formale equivalente ai PM¹³ di Russell e Whitehead o all'aritmetica di Peano.
- Teorema di GÖDEL (1930) : nel calcolo logico di primo ordine¹⁴ la correttezza semantica equivale alla derivabilità sintattica.
- Primo teorema di GÖDEL (1931) : un sistema logico formale equivalente ai PM e quindi all'aritmetica è sintatticamente incompleto.
- Secondo teorema di GÖDEL (1931) : in tutte le teorie formali (=sistemi formali) abbastanza complesse da contenere l'aritmetica non è derivabile l'asserto che dichiara la loro coerenza.
- Teorema di TARSKI : in un sistema formale la nozione di verità relativa ad esso non è definibile internamente.
- Tesi di CHURCH-TURING: il pensiero è equivalente ad un algoritmo e ogni algoritmo può essere riprodotto su di un elaboratore.

Questi risultati (escluso l'ultimo, che infatti riporta la dizione di “TESI”) hanno dimostrato i limiti dei sistemi logici formali, qualora siano abbastanza potenti.

Simili limitazioni influenzano fortemente l'IA, in quanto l'elaboratore è una versione (non sappiamo se ridotta o isomorfa) di un sistema formale e quindi di un sistema algoritmico, soggetto a tutte queste carenze in campo semantico.

Se dunque la tesi di CT regge, il pensiero deve avere limitazioni simili ? Questo potrebbe essere, giustificato dal fatto che noi non ce ne accorgiamo semplicemente perché non possiamo uscire da noi stessi e guardarci dall'esterno.

Dunque sorgono le seguenti domande: i limiti di un sistema formale si applicano realmente agli elaboratori o c'è la possibilità che, oltre una certa complessità, i risultati suddetti non ab-

¹² Questo problema, “problema della scelta” fu uno degli storici problemi che il grande matematico tedesco David Hilbert propose al Congresso Internazionale dei Matematici a Parigi nel 1900 (“Decimo problema di Hilbert”).

¹³ *Principia Mathematica*, opera ed anche sistema formale con cui R. e W. rifondarono la matematica su nuove basi logiche, dopo le intuizioni di Frege e Peano.

¹⁴ Il calcolo logico dei predicati del primo ordine permette di applicare i quantificatori “esiste” e “per ogni” solo alle variabili, mentre in quello del secondo ordine possono essere estesi anche ai predicati. L'aritmetica è assiomaticizzabile solo in un calcolo del secondo ordine.

biano più validità ? Come è possibile inquadrare l'attività mentale all'interno di un sistema logico-formale ? Per un computer è allora possibile un ragionamento o un intuizione metalogica come per l'uomo ? Esiste un algoritmo non implementabile ? Esiste un pensiero non algoritmico ?

Programma intelligente ?

Il problema che abbiamo presentato con la tesi di CT non si esaurisce certo a quelle questioni che abbiamo sollevato, ma anzi merita una trattazione a parte, tanto grande è la sua importanza nel campo dell'intelligenza artificiale.

Una delle molte formulazioni equivalenti¹⁵ di questa tesi è che non esista nessun tipo di algoritmo più potente (ovvero che possa proporre soluzione ad una più vasta categoria di problemi) dell'algoritmo parziale ricorsivo (che ovviamente include in sé come casi degeneri anche i due precedenti); questo pone due strade di interpretazione per quel che riguarda l'attività mentale e l'intelligenza. Delle due l'una, o la mente umana è una macchina equivalente ad un computer (una macchina di Turing) su cui girano innumerevoli e complessi programmi ricorsivi, o l'attività mentale è metalogica e metaalgorimica e quindi esistono processi che potremmo chiamare *intuitivi* che trascendono la dimensione formale e giungono a conclusioni praticamente valide in prima approssimazione di cui non è sempre possibile una giustificazione deduttiva.

Ora la questione è vitale per uno studio dell'IA: questa divisione netta che non permette sfumature e posizioni intermedie ha prodotto due filoni di ricerca e di pensiero all'interno della cibernetica e dell'IA, connotati da significative differenze non solo sul piano dell'impostazione teorica, ma anche dal punto di vista delle pratiche operative.

Da una parte, c'è chi sostiene una completa equivalenza e corrispondenza tra i processi mentali e gli algoritmi e quindi crede fermamente che l'intelligenza sia unicamente contenuta nella formulazione algoritmica e logica dei programmi; dall'altra troviamo chi, sostenendo non solo la presenza di "computazioni alogiche"¹⁶, ma addirittura la loro esclusiva sui processi mentali, ritiene completamente vano ogni tentativo di riprodurre intelligenza, mirando piuttosto a riprodurre comportamenti intelligenti, *emulando*, ma non *simulando* l'intelligenza umana.

Con questo non volgiamo dire che non vi siano posizioni intermedie o che non si possano trovare, soprattutto in campo operativo, concezioni eclettiche o addirittura ambigue, che concedono importanza ai risultati raggiunti e agli obiettivi prefissi piuttosto che alla teoria generale sottostante.

Dunque, troviamo qui una sola importantissima domanda, che ha però varie formulazioni: ogni operazione eseguita dalla mente può essere implementata su una macchina opportunamente programmata? Un algoritmo è dunque la sede e l'origine dell'intelligenza? Un computer può simulare l'uomo? Esiste un livello superiore di elaborazione rispetto agli algoritmi ricorsivi parziali?

La casualità e l'euristica

Tra le posizioni intermedie nominate prima rientra a pieno titolo quella che pone, alla base delle computazioni alogiche il fattore casuale che sarebbe introdotto dall'enorme complessità del cervello e dalla natura fisica dell'attività cerebrale. Questo potrebbe spiegare gli errori e le particolarità del pensiero, come le associazioni improvvise e le idee inspiegabili. Del resto il

¹⁵ D. R. Hofstadter ne presenta svariate alternative nel Capitolo XVII di "*Gödel, Escher, Bach: un'eterna ghirlanda brillante*".

¹⁶ Con questo efficace ossimoro si vogliono indicare procedimenti che prescindendo da un sistema formale raggiungono risultati sostanzialmente validi e danno di norma risposte simili a problemi simili.

caso da solo non basta, ma anzi complica la faccenda, obbligandoci a riflettere sul perché le associazioni casuali siano sempre o spesso esatte, o comunque molto plausibili.

Ovviamente la teoria del caso non è così semplice come le righe precedenti potevano far supporre.

Uno dei concetti più interessanti introdotto da questo approccio all'intelligenza è quello di reti semantiche: il nostro pensiero lavorerebbe appunto in base a queste reti semantiche, che sono raggruppamenti di idee in maniera più o meno complessa, in base a relazioni causali, somiglianze, antitesi, vicinanze spaziali o temporali, affinità e altro; grazie a queste impostazioni anche i fattori casuali non potrebbero portare il pensiero su una strada esageratamente errata. Inoltre si suppone vi sia una struttura di controllo che vagli la possibilità di correttezza delle idee ottenute tramite l'aggregazione casuale in base a riflessioni coscienti già avvenute e all'esperienza già immagazzinata e che trasmetta le ipotesi plausibili alla sfera del pensiero cosciente e dell'intelligenza per una verifica logica.

In informatica viene introdotto un concetto molto simile che è quello di euristica e, più di recente, di logica fuzzy¹⁷. In questo modo vengono sfruttati anche nella programmazione, tramite artifici matematici, i concetti di simile, più o meno affine ed in generale tutte quelle classificazioni che non possono essere trattate in termini di logica binaria e di computazione discreta.

Senza voler scendere nei dettagli, l'euristica fornisce, tramite elaborazioni matematiche di percentuali e variazioni occasionali, criteri di somiglianza e di adeguatezza a criteri non completamente oggettivi e matematizzabili.

La logica fuzzy si spinge oltre e tenta di implementare concetti quali bello, caldo freddo, alto, basso, gradevole, spiacevole, tramite complessi calcoli.

Così facendo si ottiene certamente una buona emulazione, come dimostrano risultati a dir poco eccezionali nel settore dell'Home Tech (tecnologia per la casa, quindi cucine, climatizzatori e altro computerizzati), ma non si può certo giungere ad una simulazione, poiché le computazioni matematiche necessarie richiedono agli elaboratori tempi elevati, seppure accettabili, di solito inferiori al minuto, ma comunque impensabili per la mente umana, più lenta di un elaboratore a calcolare, ma rapidissima ad accorgersi se l'ambiente è *troppo* caldo.

Con questo non vogliamo escludere l'euristica dalla ricerca dell'IA forte, che si prefigge di emulare il pensiero umano, solo cerchiamo di evidenziare come l'attuale realizzazione pratica dei principi euristici o fuzzy sia certamente inadeguata e nemmeno lontanamente simile alla gestione (eventuale) di questi criteri nei processi mentali umani.

Dunque ci si presentano fundamentalmente due questioni: il caso è un elemento effettivamente significativo nell'attività mentale? L'euristica è un suo surrogato accettabile, ovvero, c'è un modo migliore di produrre sui computer questo elemento casuale?

¹⁷ Dall'inglese *fuzzy* : sfumato, sfocato, indistinto; in questo tipo di logica l'appartenenza di un elemento ad una classe non è una funzione binaria (sì/no), ma è una funzione reale, che permette di definire appartenenze, appunto, sfumate, imprecise, rendendo possibile definire un elemento appartenente più o meno di un altro ad una classe.

HARDWARE E SOFTWARE

In informatica *hardware* indica quella parte dell'elaboratore che viene definita "rigida, dura" (hard), ovvero sia le sue componenti fisiche, mentre *software* indica la parte più "malleabile", ovvero i programmi che vengono eseguiti dalla macchina¹⁸. E' abbastanza intuitivo comprendere che il software sarà organizzato in vari livelli di competenze: un programma per produrre animazioni si occuperà del colore di ogni singolo punto luminoso (pixel) dello schermo, ma di certo non regolerà i processi fisici che permettono di visualizzare l'immagine, come un programma di elaborazione testi si preoccuperà di quali caratteri stampare e di come stamparli, ma non invierà alla stampante comandi su come far uscire l'inchiostro dalla cartuccia o su come spostare la testina di scrittura.

Un simile termine di paragone è stato molto sfruttato nelle ricerche sulla mente e sull'intelligenza artificiale, facendo sorgere dubbi sulla "sede" dell'intelligenza; insomma, siamo intelligenti per il nostro hardware (il cervello) o per il nostro software (i processi mentali) ?

MENTE

La domanda potrebbe sembrare retorica e verrebbe spontaneo rispondere che l'importante sono i processi mentali, ma bisogna tener presente che affermare che l'intelligenza abbia la sua causa nel cervello non vuol dire negare la loro presenza, ma ricondurne l'origine alla complessità neurofisiologica del cervello.

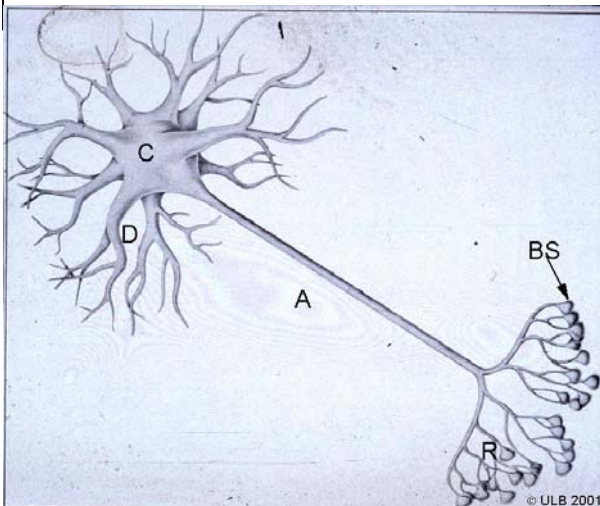
Quindi è innanzi tutto necessario cercare di capire qualcosa di più sulla natura del cervello, partendo dalle cellule che lo compongono, i neuroni.

I neuroni

I neuroni sono le cellule che compongono il cervello e il sistema nervoso; spesso sono anche chiamati fibre nervose.

Figura 5

Schema di un neurone.



Vediamone più nel dettaglio la struttura : per prima cosa occorre precisare che queste cellule lavorano tramite impulsi elettrici, ovvero creando differenze di potenziale. Altra precisazione che occorre fare è che il neurone ammette solo due stati: attivo e inattivo. Quindi, in seguito alla ricezione di impulsi il neurone può solo trasmettere a sua volta o non trasmettere.

La differenza di potenziale che si misura in un neurone inattivo è di -70 millivolt¹⁹, mentre un neurone attivo trasmette un impulso di +20 mV.

La carica per tale impulso è fornita dal meccanismo della pompa sodio-potassio: il centro del neurone (C in figura) è formato da un nucleo e dal citoplasma ed è detto *pirenoforo*, al suo interno si trovano cloruro di sodio e cloruro di potassio,

dissociati in ioni K^+ , Na^+ e Cl^- ; attorno al neurone si trovano gli stessi sali, solo con concentrazioni di ioni diverse rispetto all'interno.

¹⁸ Esiste anche il *firmware*, ovvero quei programmi che sono "scritti" direttamente nei componenti elettronici dell'elaboratore e permettono di gestire direttamente le componenti fisiche più elementari (gestione di basso livello).

¹⁹ Questo valore è interessante, poiché non solo è costante nel sistema nervoso umano, ma si incontra anche nei fenomeni di elettricità cellulare che riguardano i vegetali e gli altri animali.

La membrana di separazione, costellata di porte-sodio e porte-potassio, contribuisce a tenere questo rapporto tra gli ioni, detto equilibrio osmotico, che crea la differenza di potenziale suddetta. Questo meccanismo è detto pompa sodio-potassio ed avviene anche in altre cellule. Quando un neurone deve emettere un impulso, la pompa si inverte e crea un potenziale di segno opposto, come già detto.

L'impulso viene trasmesso lungo un canale, indicato in figura con A e chiamato *assone* o *neurite* o *cilindrassa*. Questo canale può presentare ramificazioni (R) che devono tutte terminare con i cosiddetti *bottoni sinaptici* (BS), che mettono in comunicazione il neurone con il successivo che riceve l'impulso tramite ramificazioni, situate sul lato opposto rispetto all'assone e dette *dendriti* (indicate con D). Tra i bottoni sinaptici e le dendriti c'è un piccolo spazio detto *fessura sinaptica*, attraverso il quale l'impulso deve trasferirsi all'altro neurone, tramite una sostanza detta *neurotrasmettitore*, che viene emessa dai bottoni sinaptici ed eccita le dendriti. Alcune *sinapsi* (collegamento bottone sinaptico-dendrite) sono dette eccitatorie, perché causano, eccitate, una differenza di potenziale positiva nel ricevente, stimolandolo ad emettere a sua volta; altre vengono chiamate inibitorie, poiché producono, eccitate, una differenza di potenziale negativa che ostacola l'emissione da parte del ricevente. Quando il bilancio delle varie dendriti supera un valore detto valore di soglia, il neurone ricevente scarica a sua volta lungo il proprio assone e il processo ricomincia.

Dopo aver trasmesso l'impulso, il neurone deve ristabilire la ddp di -70 mV e quindi c'è un periodo in cui il neurone non risponde agli impulsi; questa si chiama fase "refrattaria".

La velocità media di un impulso lungo una fibra nervosa è di 100-120 m/s, ma nei processi di visione e di udito, questa si abbassa anche a valori attorno ai 25 m/s.

Ogni neurone è però collegato con centinaia di altri neuroni tramite dendriti e ramificazioni dell'assone, quindi un singolo neurone, scaricando, contribuirà a eccitare o inibire un gran numero di altri neuroni; ovviamente, per superare il livello di soglia è di solito necessario che più neuroni scarichino contemporaneamente verso un altro, apportando i loro vari contributi. Dunque vediamo che il neurone non giustifica nella sua struttura l'intelligenza, poiché, se escludiamo i collegamenti multipli (che però riguardano la struttura del cervello e non del neurone), esso si comporta come un transistor, con due stati possibili ben definiti, ed anzi è assai più lento: nei cavi di un conduttore di rame la corrente elettrica viaggia ad una velocità vicina a 100000km/s, un terzo della velocità della luce.

La nostra indagine dovrà allora spingersi alla configurazione ed ai rapporti strutturali tra i neuroni.

La società della mente²⁰

Il passo successivo è quello di cercare uno o più livelli di organizzazione intermedi tra il neurone singolo e il cervello nel suo complesso. Questa strada è stata intrapresa in entrambe le direzioni, da una parte scomponendo le attività mentali in sotto procedure e compiti più semplici, che sia infine possibile ridurre a semplici interazioni fra i neuroni, dall'altra cercando indizi, nell'attività cerebrale di gruppi più o meno autonomi di neuroni, che agiscano in sincrono o comunque concordemente.

Entrambe le strade hanno portato a determinare vari livelli di complessità, ma entrambe si fermano quando devono spiegare il passaggio dal singolo neurone ad un gruppo (finora solo ipotetico) di pochi neuroni.

Un buon esempio per spiegare questa concezione del cervello è fornito da D.R. Hofstadter nel Capitolo XI e nel Dialogo "...*mirmecofuga*" in "Gödel, Escher, Bach: un'eterna ghirlanda bril-

²⁰ Titolo di un famoso libro di Marvin Minsky, in cui egli tenta di dare una spiegazione dell'intelligenza scomponendone le facoltà in agenzie e sottoagenzie, fino a livelli molto semplici, affidabili, a suo dire, all'attività di ristrette zone del cervello.

lante”, in cui l’autore pone un efficace parallelismo con l’immaginaria struttura di un formicaio intelligente, nel quale le formiche non sono affatto intelligenti, ma si organizzano in gruppetti, squadre, segnali, simboli, caste e fanno sì che il dottor Formichiere trovi nel barone di Monteformica Furio-Caio (un formicaio) un interlocutore gradevolissimo.

Queste considerazioni portano a discutere di due teorie che in campo di scienze della mente hanno un’importanza fondamentale: l’olismo e il riduzionismo.

Molto semplicemente essi possono essere definiti tramite una frase per ciascuno che esponga la visione che questi due termini indicano.

Olismo: “L’intero è maggiore della somma delle sue parti”.

Riduzionismo: “L’intero si riduce alla somma delle sue parti”.

Un punto di vista olistico sulla questione che stiamo affrontando implica il considerare il cervello come qualcosa di più rispetto alla somma dei miliardi di neuroni che lo compongono; il riduzionismo vorrebbe invece che da una descrizione a basso livello della struttura cerebrale si possano dedurre le caratteristiche globali e trovare le cause e i meccanismi delle sue attività.

L’olismo implica un’impossibilità di conoscere, almeno per ora, la causa dell’intelligenza, non avendo noi individuato altro metodo di studio del cervello se non l’indagine sui neuroni; il riduzionismo, d’altra parte, richiede un’immensa mole di lavoro, che unisca progressivamente i risultati raggiunti per un neurone, studiando le combinazioni di gruppi sempre più ampi di cellule e cercando di vedere a quale livello corrispondano determinate facoltà mentali, presupponendo che noi si sia già classificate tutte le attività mentali di base (cioè quelle attività “semplici” che compongono, unendosi e combinandosi, attività di alto livello, come la visione).

A questo punto, si presentano varie questioni: innanzi tutto è necessario proseguire almeno in modo sommario lo studio della struttura cerebrale su scala più grande; inoltre sarà utile una panoramica sulle teorie che cercano di spiegare le relazioni mente-cervello; infine dovremo affrontare questo tema anche dalla parte del software mentale, di cui si conosce molto poco e su cui le congetture rimangono ancora molto vaghe.

La struttura del cervello

La prima precisazione obbligatoria è che il cervello è, propriamente parlando, una parte dell’organo centrale del sistema nervoso, detto encefalo e diviso in tre parti: cervello, cervelletto e ipotalamo.



Figura 6 – Le aree principali del cervello

Il cervello è ulteriormente diviso in due emisferi ed in alcuni lobi, rivestito da una “buccia” stratificata detta corteccia cerebrale. Grazie alle moderne tecniche di scansione ai positroni, si è in grado di riconoscere con buona precisione quali aree siano attive in quale momento, ma questo ci dice comunque assai poco su un eventuale corrispondenza tra certe zone del cervello e certe attività o percezioni.

Ad esempio è noto che il movimento è legato ai comandi del cervelletto da cui si dipartono i treni di nervi verso l’emisfero sinistro. Esperimenti

fatti rilevando l'attività cerebrale durante l'ascolto di musica, ha rivelato una preponderante partecipazione del lobo temporale, ma in modi assai vari ed insieme ad altre aree del cervello, quale il lobo prefrontale.

Un fattore interessante da tener presente è che i cervelli umani, al contrario dei cervelli di forme di vita meno complesse (ad esempio i lombrichi) non sono *isomorfi*, cioè non è possibile ritrovare la stessa cellula nervosa in due diversi cervelli umani, mentre è possibile farlo con i lombrichi; in altro modo potremmo dire che i lombrichi hanno tutti uno stesso cervello, ovvero che esiste un solo lombrico. Eppure nel cervello umano vi sono somiglianze notevoli su scala appena più grande del neurone. Forse questo può implicare che le interconnessioni tra neuroni e le loro particolari disposizioni (ciò che rende impossibile ritrovare in due encefali umani la stessa cellula) siano una traccia delle conoscenze e delle nozioni dell'individuo? Ovvero, le differenze mentali tra gli individui si riflettono sulla struttura fisica del cervello? E' possibile individuare sottoinsiemi di neuroni che sovrintendono a particolari processi conoscitivi, mnemonici o altri aspetti della vita mentale individuale?

A queste domande hanno cercato di dare risposta molti studiosi, tra i quali spicca come pioniere Karl Lashley che nel 1920 condusse esperimenti sulla capacità di ricordare dei ratti: le cavie venivano poste in un labirinto con del cibo all'altra estremità e l'esperienza veniva ripetuta finché il ratto non imparava la strada, quindi ad alcuni di essi venivano asportate zone di corteccia; fu rilevato che queste cavie menomate avevano difficoltà a muoversi e a coordinarsi e la difficoltà era tanto più grande quanto maggiore era la parte di corteccia asportata. Ciò non intaccava però la capacità di ricordare. Lashley concluse in un articolo del 1950 che la memoria non era possibile, mentre in seguito questo risultato è stato reinterpretato come segno della diffusione di certi ricordi su vaste aree corticali e interne, con duplicazioni e ridondanze volte a preservare l'informazione.

Nello stesso periodo in cui Lashley terminava le sue esperienze (anni '40), il neurochirurgo Wilder Penfield iniziava a sperimentare in Canada una nuova tecnica di indagine: egli stimolò con elettrodi vari punti della corteccia cerebrale di pazienti umani, mandando impulsi simili a quelli che sarebbero potuti provenire da un altro neurone; i pazienti poi riferirono di aver avvertito ronzii, rumori, musiche, o di aver visto luci, colori, movimenti, oppure in alcuni casi di aver rivissuto scene passate ed esperienze più o meno remote. Penfield si accorse che le aree che potevano causare simili ricordi erano ristrettissime, pochi neuroni. Questi risultati corroborarono la seconda interpretazione dei risultati di Lashley, e cioè che i ricordi avessero vari duplicati in varie zone corticali e interne.

Negli anni '60 e '70 gli esperimenti di David Hubel e Torsten Wiesel sulla visione hanno localizzato alcuni interessantissimi percorsi neurali, che gli esperimenti di Lashley sembravano aver negato. Essi hanno individuato aree cerebrali che si occupano di particolari configurazioni geometriche semplici, ad esempio la presenza di un margine verticale che separi due zone a diversa luminosità (area 17); in queste aree si trovano neuroni (cellule complesse) che reagiscono ognuno ad una particolare situazione, con una scarica nulla, debole o forte. Questi risultati hanno portato a supporre che vi siano strutture dinamiche di neuroni che si comportino, all'attivazione come una presunta "cellula della nonna" cioè una cellula molto complessa che reagisca con diversi gradi di scarica ai vari gradi di "presenza" visiva della nonna : oggetti di sua proprietà, una foto, la nonna di spalle, la nonna di fronte. Ovviamente questa è solo una scherzosa ipotesi, ma le attuali teorie dei percorsi neurali tendono a seguire questo orientamento.

Teorie della mente

Le teorie della mente si sono sviluppate in ambito prettamente filosofico fino ai primi anni del Novecento, quando il problema del rapporto tra cervello e mente ha interessato anche la psicologia; quindi, la questione è divenuta interessante anche per gli studiosi di intelligenza arti-

ficiale, in quanto premessa non solo teorica ai loro lavori. In questo periodo di multidisciplinarietà queste teorie sono andate dividendosi secondo le due correnti summenzionate, l'olismo e il riduzionismo²¹, occupandosi più scientificamente della produzione dei processi di pensiero dai processi fisico-chimici del cervello.

Le prime teorie della mente si sviluppano quando acquista importanza il problema del rapporto corpo-spirito in campo conoscitivo, sotto forma di relazione tra rappresentazione e cosa in sé, ovvero sia con la filosofia cartesiana che sviluppa una posizione di *dualismo interazionista*: la mente e il corpo sono due *res* diverse tra loro e vengono in reciproca azione causale tramite la ghiandola pineale, un particolare organo.

L'immediata conseguenza è l'*occasionalismo* di altri razionalisti, come Spinoza, secondo cui le due *res* non hanno alcuna interazione, ma ognuna procede per conto proprio, mentre Dio si assicura che alla stimolazione degli organi di senso corrisponda una sensazione o che al pensiero di muoversi corrisponda un movimento. Altri si sono rifugiati nel *panpsichismo* che suppone una presenza spirituale in ogni realtà fisica, facendo cadere il problema del rapporto *res extensa - res cogitans*. Una teoria che ha perdurato per parecchi tempo è l'*epifenomenismo*, secondo cui, se anche ad un certo livello di aggregazione vivente la materia produce attività mentali, queste non hanno alcuna conseguenza sul mondo materiale ed hanno con esso solo un aggancio passivo (ne sono prodotte, ma non vi agiscono); da questa concezione nasce il *comportamentismo*, che nega la necessità di un'attività mentale di origine non materiale.

Infine ha avuto un vario sviluppo il *monismo* che prevede l'assolutizzazione di una delle due *res*, affermando l'esistenza dei soli stati fisici del cervello o della sola realtà mentale; le sue derivazioni più longeve comprendono il *monismo meccanicista* che studia i processi fisici e chimici dei singoli elementi del cervello, rinunciando senza affermarlo esplicitamente ad uno studio di processi mentali superiori, e il *monismo emergentista* che funzioni del pensiero a fenomeni biologici.

Quest'ultima teoria è a tutt'oggi la più promettente ed è sostenuta nella sua formulazione più feconda da Gregory Bateson in *Mente e Natura* del 1979 : "Una mente è un aggregato di parti o componenti interagenti. [...] La spiegazione dei fenomeni mentali deve sempre trovarsi nell'organizzazione e nell'interazione di parti multiple." Anche la natura per Bateson non è nient'altro che una struttura connettiva e connessa in cui è inserita la struttura mentale.

Un'altra teoria che si è affermata negli ultimi anni del XX secolo è la teoria di K. R. Popper, che egli espone in vari scritti tra cui il volume primo de *L'Io e il suo cervello*, scritto a 4 mani con John Eccles, neurofisiologo premio nobel; qui Popper sostiene una posizione di *interazionismo psico-fisico* nel quale egli accetta la teoria dell'evoluzione darwiniana da cui partono i materialisti contemporanei, per poi però affermare che la Natura abbia prodotto per sue volte una "vera novità" non insita nel materiale di partenza, prima la vita dalla materia inorganica, poi la mente dalla materia viva non cosciente, infine la creatività e il linguaggio dalle menti primitive dei mammiferi superiori. Questa posizione popperiana è quanto meno ambigua, in quanto Popper, pur non rinunciando alla scienza e alla radice biologica del cervello, tenta di recuperare quell'umanesimo, quella dignità umana che erano andate perdute con i materialisti e i fiscalisti.

Il software mentale

Mentre sull'hardware dell'intelligenza c'è comunque molto da dire, nonostante i grandi dubbi lasciati aperti e le aree non ancora esplorate, sul software della nostra mente non possiamo dire più di tanto. Questo perché non sappiamo dove "leggere" le ipotetiche istruzioni che la

²¹ L'olismo e il riduzionismo non interessano solo questo particolare campo di ricerca, ma sono due modi di pensare e di intendere la realtà che permeano, esplicitamente o implicitamente, vaste aree del sapere.

nostra mente segue. Le supposizioni in questo campo si rifanno, da una parte, alla matematica e alla logica, dall'altra, alla psicologia ed alle neuroscienze.

Il contributo della logica e della matematica riguarda, come è ben comprensibile, la sfera del ragionamento deduttivo e in generale delle applicazioni scientifiche del pensiero, seppure ancora non abbia saputo né potuto fornire una risposta al problema, già citato della creatività e più in generale del pensiero divergente, che è in fondo la caratteristica che Turing cercava di individuare nella macchina per definirla intelligente, ovvero sia la possibilità del pensiero ipotetico-deduttivo, opposto alla coazione a ripetere, identificata nel pensiero convergente.

La logica ha supposto che almeno parte dell'attività mentale umana nel campo del ragionamento avvenga in base a processi che implicano algoritmi o calcoli logici del secondo ordine; come abbiamo già sottolineato, non si conosce per ora un più potente calcolo logico predicativo, ovvero sia non si è ancora trovato una caratteristica algoritmica più efficace nel risolvere problemi della ricorsività parziale.

Le neuroscienze vengono in aiuto al problema del software mentale con una schietta fede nel metodo sperimentale e con una concezione fondamentalmente empirica dell'indagine in questo campo, soprattutto grazie alla suddetta impostazione monista emergentista che ha condizionato buona parte degli studi finora svolti; nonostante questo forte legame con la sperimentazione e un implicito rifiuto della costruzione di teorie, dobbiamo riconoscere a questa impostazione la capacità di fare da "ponte" tra l'insostenibilità di una concezione materialistica comportamentista e la necessità di collegarsi alla struttura biologica, chimica e fisica del cervello e dell'essere umano in generale.

I risultati più significativi sono stati osservati nello studio della registrazione delle percezioni sensoriali e nella loro prima elaborazione da parte della mente. Importante è la teoria di Marr sulla visione, elaborata tra mille aporie da un punto di vista teorico-filosofico, ma assai utile e pratica nella descrizione e nello studio dell'attività cerebrale al momento di uno stimolo visivo coscientemente registrato; da segnalare anche l'applicazione delle teorie di Chomsky sul linguaggio ad un modello di "simboli neuronici", ovvero di aggregati dinamici e interconnessi di neuroni, che si comportano come cellule nervose giganti con una gamma di risposta più ampia del semplice sì/no neuronale.

A questo punto, è lecito interrogarsi anche solo sull'esistenza di questo software mentale, il che però equivale ad interrogarsi sull'esistenza di una mente altro dal corpo e sostanzialmente ad esso non riconducibile in tutte le sue attività tramite collegamenti causali di tipo materiale.

Certo è più diffuso il dubbio sulla complessità di questo software mentale: esso è legato ad attività semplici, suddiviso in miriadi di compiti elementari che combinati danno luogo a manifestazioni più complesse o consta di "programmi" dettagliati anche per i più alti processi di pensiero?

Ancora una volta è doveroso citare *La società della mente* di Marvin Minsky, in cui l'autore offre una sistemazione di un possibile software mentale in termini riduzionisti, sostenendo che l'intelligenza e i processi superiori, su cui primeggia l'autocoscienza, emergono come unione e azione simultanea di queste parti in sé semplici e implementabili su un computer (almeno in linea di principio).

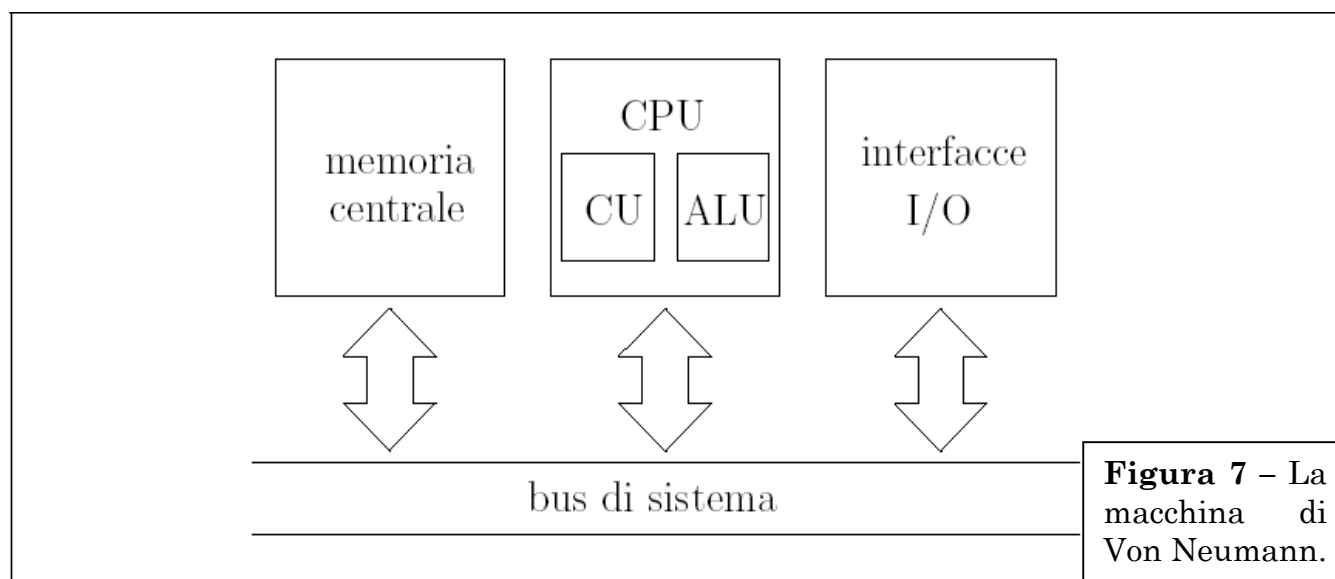
MACCHINA

Per quanto riguarda la macchina, questo discorso è al tempo stesso antico e nuovo, poiché da quando l'algoritmo ha guadagnato indipendenza dalla configurazione tecnica delle componenti elettroniche, divenendo modificabile senza necessità di intervenire sui circuiti dell'elaboratore, il rapporto hardware-software è uno dei motivi di discussione maggiori in ambito informatico; eppure, la questione della dipendenza della qualità dei risultati ottenibili

in IA dal livello e dalle specificità dei supporti materiali non è stata presa seriamente in considerazione come una linea su cui eseguire ricerche se non in tempi molto recenti. Questo ritardo va attribuito allo stato della tecnologia che fino a dieci anni fa rendeva computer che non rientrassero nella *strozzatura di Von Neumann*²² un sogno da Star Trek; oggi invece possiamo parlare di computer quantici e molecolari, anche se a livello di sperimentazione.

Per queste ragioni e anche per il fatto che il software impiegato nelle ricerche di intelligenza artificiale è ad un altissimo livello tecnico e richiede, per essere compreso anche solo nei suoi caratteri generali, una enorme quantità di concetti che non possiamo dare per scontati, punteremo la nostra attenzione soprattutto sull'hardware, mentre solo in ultimo ci soffermeremo sui rapporti con il software, che, sia detto per inciso, presenta ultimamente meno spunti innovativi rispetto all'hardware. Cominceremo col fornire una veloce descrizione dell'architettura generale di un elaboratore elettronico, per poi esporre le due più significative linee di ricerca, già citate sopra, e considerarne le possibilità e i vantaggi.

L'hardware attuale



Lo schema in figura si riferisce ad una macchina progettata nel 1947 da Von Neumann, Burks e Goldstine a Princeton per eseguire calcoli a grande velocità. Nonostante la tecnologia dei computer abbia conosciuto grandi progressi, questo schema (detto *architettura classica dell'elaboratore*) rimane ancora valido come descrizione dell'hardware di un calcolatore ad un certo grado di astrazione.

Le componenti principali sono tre: la *memoria centrale* (RAM, ROM, cache), il *microprocessore* (CPU), le *interfacce input/output* (I/O); queste tre componenti sono tra loro collegate dal *bus di sistema*. Vediamo brevemente il loro funzionamento.

La CPU (*Central Processing Unit* – unità di elaborazione centrale) ha la funzione di controllare il funzionamento del computer e di sovrintendere tutte le operazioni che si compiono al suo interno; essa interpreta ed esegue le istruzioni di un programma tramite due sotto-unità funzionali che sono la *ALU* (*Arithmetic Logical Unit* – unità logico-aritmetica), che esegue le operazioni aritmetiche e le operazioni logiche di confronto tra i dati, e la *CU* (*Control Unit* – unità di controllo) che segue il flusso delle istruzioni e coordina il lavoro degli altri componenti dell'elaboratore.

²² La *strozzatura di Von Neumann* è la limitazione a processare, uno alla volta, un numero finito ma non limitato di celle dato, in cui sia possibile scrivere o leggere uno tra i simboli di un certo insieme finito, con operazioni discrete e finite. Una macchina simile è detta anche *macchina di Turing*, poiché Turing la usò come esperimento mentale nelle sue teorie sulla commutabilità e la ricorsività.

La memoria centrale è un dispositivo di memorizzazione dati che si divide in tre tipi: la memoria RAM (*Random Acces Memory* – memoria ad accesso casuale) nella quale il processore può accedere, per leggere o scrivere, a qualunque punto della memoria (*casuale* viene contrapposto a *sequenziale*); la memoria ROM (*Read Only Memory* – memoria di sola lettura) nella quale vengono conservate informazioni postevi alla fabbricazione del computer che riguardano quei programmi di *firmware*²³ per la gestione a basso livello della macchina; la memoria *cache* (in inglese *nascondiglio*), una memoria ad alta velocità (ed alto costo) in cui possono essere trascritti alcuni dati che necessitano di essere trattati rapidamente o molte volte nel corso di un algoritmo.

I dispositivi di input/output (ricezione/trasmissione) servono a mettere il computer in comunicazione con l'utente, permettendo a questo di inserire dati tramite tastiera, mouse, scanner, penna ottica, microfono, telecamera ed altre periferiche, e di conoscere la risposta della macchina tramite schermo, stampante, altoparlanti, modem, plotter, etc.

Il bus è l'insieme dei collegamenti elettrici tra le parti ora elencate che permette di trasferire dati dall'una all'altra, regolandosi secondo i tempi del *clock* (orologio) della CPU, solitamente con una frequenza di circa 14 MHz²⁴, poi aumentato tramite sintonizzatore.

L'architettura moderna di un calcolatore è molto più complicata: nella CPU esistono spazi di memoria detti registri, ad altissima velocità, circuiti di segnalazione di eventi detti interruptus; le memorie RAM, ROM e cache si classificano in vari tipi a seconda dell'utilizzo o delle caratteristiche tecniche; i dispositivi I/O dialogano con controller di flusso e interfacce fisiche che permettono loro di sintonizzarsi con il clock della CPU e le sue modalità di azione; esistono almeno 3 tipi di bus, di solito a doppio canale (quindi in pratica 6 bus).

Inoltre esistono delle unità di memoria fissa, per memorizzare dati a lungo termine, dette memorie di massa (dischi rigidi, CD-ROM, DVD, dischi ottici e magneto-ottici, etc.)

Vediamo ora brevemente il funzionamento, sempre in linee generali, di un elaboratore: esso si divide in due fasi generali, denominate FETCH e EXECUTE²⁵; nella prima il processore cerca e carica in memoria l'istruzione da eseguire e gli eventuali dati che essa necessita, nella seconda questa istruzione viene eseguita dall'ALU, nel caso di una operazione matematicologica, o dalla CU nel caso di una istruzione di memoria o di I/O.

Ovviamente anche questo schema generale viene complicato nell'architettura moderna per ottenere migliori prestazioni e minori possibilità di errore; inoltre sono state sviluppate e formalizzate nel corso degli anni molte tecniche di gestione della memoria e delle istruzioni, che qui non abbiamo spazio per trattare, per le quali rimandiamo ai testi indicati nella bibliografia.

E' evidente che, nonostante alcune somiglianze, troviamo notevoli differenze tra l'hardware mentale e quello di un elaboratore : innanzi tutto il computer opera in modo *seriale*, cioè esegue un'istruzione dopo l'altra e non può dedicarsi nello stesso istante a due compiti, mentre il cervello lavora principalmente in *parallelo*, ovvero troviamo più neuroni o gruppi di neuroni in azione contemporaneamente; inoltre nel cervello non compaiono distinzioni apparenti (a livello anatomico) tra *memoria* e *processore*, ma ogni neurone può indifferentemente essere usato per ritenere un'informazione o assolvere ad un "calcolo"; il numero di neuroni sorpassa vistosamente il numero di transistor, mentre la velocità di un componente elettronico è di molto superiore a quella di un impulso nervoso.

Dunque dobbiamo chiederci se sia lecito sperare che, con un software adeguato, un computer possa eguagliare una mente nonostante queste diversità; inoltre, è possibile modificare queste

²³Vedi nota 17.

²⁴ La frequenza esatta, 14.318 MHz, dipende dal tipo di quarzo utilizzato nei primi computer per controllare l'oscillazione del circuito integrato del clock; tale cristallo è stato mantenuto anche negli elaboratori più moderni per ragioni di coerenza storica e per il basso costo.

²⁵ to *fetch* : andare a cercare, prendere e portare; to *execute* : eseguire.

caratteristiche tecnologiche in modo da rendere il calcolatore più simile strutturalmente al cervello ?

Le reti neurali

Come abbiamo detto nell'introduzione, nei primi anni dell'IA si evidenziarono due correnti; una, che vide i suoi padri fondatori in Mc Carthy, Minsky e Shannon, puntò sullo sviluppo di software senza attribuire importanza all'hardware, utilizzando come base fisica la macchina di Von Neumann; l'altra, proposta da Mc Culloch e Pitts nel loro già citato articolo del 1943, si basò sulla costruzione di hardware più simili alla struttura del cervello umano. Da questa seconda impostazione nacquero tutti i successivi studi sulle reti neurali (RN).

Esse sono reti di elaboratori (detti *nodi*) collegati tra loro da cavi elettrici; questi nodi si comportano come i neuroni: se la somma degli impulsi supera una certa soglia, il nodo manda a sua volta un impulso. Ogni entrata del nodo ha un suo valore o *peso*, che può essere modificato in base all'esperienza o all'intervento umano, rafforzando o indebolendo la connessione tra due determinati nodi.

Una RN viene inizialmente impostata con un algoritmo di apprendimento e con pesi assegnati a caso alle varie connessioni; successivamente essa viene sottoposta a stimoli e, dopo un eventuale intervento del programmatore, l'algoritmo di apprendimento modifica i pesi delle connessioni a seconda dello stimolo ricevuto.

Esistono due tecniche principali di apprendimento: *supervisionato* e *non-supervisionato*. Nel primo caso la rete esegue la propria elaborazione e quindi confronta il proprio output con quello previsto dal programmatore, correggendo così i pesi delle connessioni; nel secondo, la rete esegue svariate elaborazioni su un gran numero di stimoli, trovando essa stessa regolarità, differenze e somiglianza e modificando "a proprio giudizio" i pesi delle connessioni, senza un confronto con il programmatore.

Oggi le RN sono formate da più strati di nodi interconnessi, di modo che l'elaborazione possa estendersi anche a classificazioni complesse, considerando varie interrelazioni.

Le reti più famose sono il *Multilayer Perceptron* e la *rete di Kohonen*, la prima è supervisionata e multistrato, la seconda è non-supervisionata e ad un singolo strato. La rete di Kohonen si basa sull'osservazione che nel cervello un neurone ha connessioni molto forti con i vicini e più deboli (anzi, a volte inibitorie) con i più distanti; Kohonen ha pensato allora di utilizzare una griglia bidimensionale in cui tutti i neuroni siano interconnessi e l'apprendimento avvenga con la modificazione dei pesi dei soli neuroni vicini. Ovviamente la scelta di quali siano i neuroni vicini non è così semplice, mentre è assai meno complicato scrivere un algoritmo di apprendimento.

I vantaggi delle reti sono il ragionare per similitudini e differenze ed il reagire in modo appropriato al nuovo e all'imprevisto: una RN ben strutturata, abituata a fornire B come risposta allo stimolo A, trovandosi di fronte allo stimolo A' (simile ad A) risponderà B oppure B', fornendo quindi o la stessa risposta o una risposta simile; inoltre l'architettura di una rete è assai simile alla presunta conformazione della corteccia e del cervello interno, lasciando sperare che questo possa essere un passo verso la realizzazione di una mente elettronica. Il dubbio è : oltre a questo abbozzo di ragionamento *analogico* (per analogie e differenze), è possibile ottenere da una RN sufficientemente evoluta un processo di pensiero divergente?

I computer molecolari

Nel 1994 lo scienziato californiano Leonard Adleman ha pubblicato su *Science* un articolo in cui proponeva e descriveva praticamente come si potesse trovare l'*hamiltoniana di un grafo*²⁶

²⁶ Un grafo è un insieme di punti collegati da archi; un'hamiltoniana (più precisamente un circuito di Hamilton) è un percorso che passi una volta sola per ogni punto del grafo, utilizzando solo gli archi del grafo.

sfruttando la capacità di calcolo delle molecole di DNA. Da allora sono stati avviati molteplici progetti in tutto il mondo per la realizzazione di computer molecolari, ovvero sia di “macchine” (tra virgolette, poiché, in effetti, questi dispositivi non presentano parti meccaniche o elettroniche) basate sulle caratteristiche combinatorie²⁷ del DNA; inoltre, è nata in parallelo una ricerca volta a produrre circuiti molecolari che possano sostituire quelli integrati al silicene attualmente in uso, risparmiando energia e spazio.

Gli obiettivi generali sono dunque la velocità, la dimensione ridotta e l'efficienza energetica.

Queste due linee si chiamano *computazione molecolare* e *elettronica molecolare*.

I risultati più interessanti per quanto riguarda l'IA li possiamo trovare nella computazione molecolare, in quanto essa cambia totalmente il modo di concepire il funzionamento di un calcolatore: in un computer molecolare troviamo trilioni di molecole che eseguono la stessa computazione contemporaneamente e forniscono possibili soluzioni; ovviamente l'utilità di questo dispositivo è da vedere in una sua applicazione a problemi che abbiano una molteplicità di soluzioni, non solo in termini matematici, ma anche euristici, cioè non solo più soluzioni certe, ma anche più soluzioni probabili o adeguate.

Simili progetti sono attivi in Giappone, in Israele, negli USA, in Europa e stanno dando notevoli risultati, il più incoraggiante dei quali è il computer a DNA ed enzimi costruito da Ehud Shapiro al Weizmann Institute of Science a Rehovot in Israele. Esso consiste in un cilindro di vetro trasparente in cui vengono miscelati in soluzione molecole “software”, molecole “carburante” e molecole “dati”, tutte di DNA, insieme ad enzimi che catalizzano opportune reazioni. Per ora il computer può solo rispondere sì o no ad una domanda e la verifica dell'output deve essere eseguita tramite l'analisi di sezioni di DNA, quindi in modo poco comodo. I problemi principali sono l'utilizzo della memoria e l'implementazione di funzioni elevate.

Di certo, un simile schema di computazione si avvicina molto ai meccanismi cerebrali, per quanto ne sappiamo ora; viene da chiedersi se effettivamente questa ricerca potrà svilupparsi ulteriormente, producendo, se non un sostituto dei computer al silicene, un dispositivo ad essi complementare, più proficuo nel riprodurre l'intelligenza su un elaboratore. Vi sono pareri fortemente negativi, come quello espresso dal gruppo di ricerca di Tokyo, che sembrano però smentiti dal recente successo di Shapiro.

I computer quantici

L'idea di un computer basato sui principi della meccanica quantistica fu proposto negli anni '70 da alcuni scienziati e ricercatori americani, sulla base della considerazione che, se ogni anno diminuivano le dimensioni dei circuiti, ben presto queste avrebbero raggiunto il livello degli atomi e quindi i processi all'interno sarebbero caduti sotto le leggi della meccanica quantistica.

Come abbiamo visto, esiste un modello astratto di struttura per il calcolatore, detto macchina di Von Neumann; ebbene, esiste un corrispondente progetto per un elaboratore quantico, realizzato da Richard P. Feynman nel 1982, che dimostrò come esso potesse fungere da simulatore per processi fisici quantistici; più tardi nel 1985 David Deutsch comprese che il progetto di Feynman poteva condurre alla realizzazione di un elaboratore quantico universale, cioè in grado teoricamente di svolgere qualunque algoritmo, e dimostrò che ogni processo fisico poteva essere simulato perfettamente su un simile dispositivo, stabilendone quindi la superiorità rispetto ad un tradizionale computer al silicio.

Ulteriori sviluppi riguardarono unicamente teoriche applicazioni di tale computer ad alcuni problemi matematici di poca importanza, finché nel 1994 Peter Shor pubblicò un articolo in cui descriveva un algoritmo di scomposizione in fattori primi di grandi numeri per un computer quantico. Questo ravvivò l'interesse dei governi e della comunità scientifica, principalmen-

²⁷ Ovverossia la possibilità di esprimere un grande numero di dati cambiando la disposizione di pochi componenti.

te perché i più sofisticati algoritmi di crittografia sono tuttora basati sulla difficoltà a scomporre grandi numeri.

Vediamo ora, per quanto è possibile, i concetti su cui si basa questo elaboratore. Al posto dei tradizionali bit troviamo i *qubit*; mentre un bit (un “interruttore”) può assumere solo i due valori 1 e 0, un qubit, formato da alcuni particolari atomi, può assumere i valori 1,0 e tutte le combinazioni (o sovrapposizioni) possibili di questi. La spiegazione di questo apparentemente strano comportamento è che la fisica quantistica prevede che siano contemporaneamente possibili più stati per una singola particella, la quale *collasserebbe* poi in uno stato ben preciso nel momento in cui fosse osservata, ovverosia ne fossero misurate alcune grandezze.

Quindi un computer quantico potrebbe contemporaneamente operare su molteplici stati, fornendo poi al momento dell’osservazione una sola risposta. Teoricamente un computer quantico può essere completamente simulato da un computer tradizionale, ma con tempi talmente lunghi da essere improponibili; questo è l’effettivo vantaggio di un elaboratore quantico, la velocità con cui può compiere le proprie operazioni, grazie alla speciale elaborazione in parallelo dovuta alla sovrapposizione degli stati.

Le operazioni che attualmente vengono svolte sui bit dalle porte logiche booleane AND, OR, NOT, etc, vengono sostituite con onde elettromagnetiche che modifichino le particelle.

Al momento attuale i principali ostacoli alla costruzione di un computer quantico sono la *correzione degli errori*, la *decoerenza*, la *tecnologia necessaria*; la correzione degli errori è necessaria poiché un sistema quantico ha la tendenza a decadere in uno stato caotico (o incoerente), perdendo quindi le informazioni in esso registrate, quando viene a contatto con il sistema ambiente e questo comportamento è detto decoerenza; nel 1998 è stato realizzato il primo elaboratore quantico, grazie alla triplicazione di ogni dato ed al controllo degli errori tramite la rilevazione della rotazione delle particelle e la tecnica di risonanza magnetica nucleare, nei laboratori di Los Alamos. La tecnologia impiegata in questo progetto è però tutt’altro che accessibile e difficilmente impiegabile in massiccia quantità per la produzione di elaboratori con elevate capacità.

L’interrogativo inerente alla nostra discussione è se questa particolare elaborazione in parallelo possa essere in qualche modo simile a quella cerebrale o comunque più utile per tentare di imitare l’intelligenza; inoltre, visti i recenti studi sulle connessioni tra neurodinamica e quantistica, un computer quantistico potrebbe essere più vicino alla natura del cervello più di quanto non si pensi.

Software per l’IA

Visti i molteplici fronti sui quali si concentra la ricerca nell’intelligenza artificiale, non ci è possibile fornire un solo schema generale di funzionamento per una ipotetica mente elettronica: l’informatica non è ancora arrivata a concepire un unico sistema che eguagli o che approssimi la mente umana, piuttosto si è dedicata a produrre varie tipologie di programma per riprodurre un particolare aspetto dell’attività mentale umana.

Esamineremo di seguito, a titolo esemplificativo, la struttura generale di un Sistema Esperto, cioè un programma che possa acquisire tramite il “dialogo” con un uomo, esperto in un determinato settore, le conoscenze necessarie a risolvere problemi riguardanti quel campo specifico, eventualmente richiedendo all’interrogante chiarimenti.

Abbiamo scelto questo esempio perché, pur semplice nelle sue linee generali, è uno dei campi in cui l’IA ha fatto i maggiori progressi negli ultimi anni, insieme alla percezione sensoriale, la cui spiegazione richiede però nozioni altamente tecniche di robotica, elettronica e informatica teorica.

Un Sistema Esperto (SE) è formato da due parti principali: una componente procedurale - algoritmica e una componente dichiarativa – rappresentativa; la prima raccoglie le definizioni degli algoritmi che intervengono nelle procedure di risoluzione dei problemi, la seconda con-

tiene la descrizione formalizzata (ovvero rappresentata) che il SE dovrà risolvere. Entrambe queste componenti si suddividono a loro volta. All'interno della seconda possiamo trovare sia la rappresentazione delle conoscenze di base indispensabili per delimitare il campo di competenze del SE, detta base di conoscenza, sia la specifica del particolare problema da risolvere, detta condizione di obiettivo. La prima componente (detta anche *guscio*) si divide in 4 parti: un *motore inferenziale*, composto dagli algoritmi generali che permettono di creare l'algoritmo per risolvere il problema assegnato; un sistema di gestione delle basi di conoscenza che struttura le informazioni; un modulo di acquisizione della conoscenza, che serve al SE per aggiornare la propria conoscenza in base all'esperienza; un interfaccia con l'utente, che permette a questo di porre domande al sistema e al sistema di chiedere chiarimenti.

Il motore inferenziale si compone di 3 tipi di algoritmi: gli algoritmi di ricerca operativa, associativa e di apprendimento; i primi cercano tra le conoscenze quelle che possono essere utili nel risolvere il problema, i secondi dispongono tali conoscenze a formare un algoritmo di soluzione, i terzi modificano in seguito ad esperienze acquisite i due precedenti tipi di algoritmo.

La base di conoscenza è un modo per memorizzare le informazioni, tramite due strutture, una sottobase di oggetti contenente conoscenza statica, sotto forma di fattualizzazione di conoscenze astratte, e una sottobase di azioni, contenente conoscenza dinamica, ovvero elementi che possono modificare la conoscenza stessa.

Questa base di conoscenza è la struttura innovativa del SE, visto che questo schema di funzionamento era già stato proposto per una macchina ben più ambiziosa, il solutore universale di problemi. Una base di conoscenza deve essere *aperta* (cioè anche ciò che non è contenuto nel suo interno può essere vero), *completa* (in cui siano possibili sia enunciati affermativi, sia negativi), *deduttiva* (presenza di regole che permettano di inferire nuova informazione), *dichiarativa* (aggiornamento in tempo reale della coscienza in seguito a cambiamento del mondo esterno), *interrogabile con ambiguità* (senza bisogno di fare domande precise, lasciando al SE di eliminare eventuali ambiguità).

Ovviamente non abbiamo detto nulla riguardo alla struttura dei singoli algoritmi, poiché non ne esiste né una versione ufficiale, né uno schema di fondo per tutti uguale, ma ogni programmatore, sceglie le soluzioni che meglio si adattano al problema particolare.

Ebbene, se a questo livello possono ancora apparire delle somiglianze piuttosto forti tra il procedere di un sistema esperto e il lavorare di una mente umana, un'analisi più dettagliata del motore inferenziale rivelerebbe meccanismi talmente sofisticati che è difficile credere che nella nostra mente la Natura non abbia trovato un sistema più comodo per ottenere gli stessi risultati.

Rapporti hardware-software

Fino ad ora uno degli obiettivi primari della produzione informatica delle grandi aziende è stata la portabilità, ovvero l'indipendenza dalla configurazione della macchina su cui un programma gira; perciò è quanto meno prevedibile una pressoché totale indipendenza del software dall'hardware.

Del resto non si vedono, con le tecnologie diffuse al giorno d'oggi, motivi validi per cui progettare un programma di IA destinato ad un particolare elaboratore di cui possa sfruttare al meglio le caratteristiche.

E' probabile che le due nuove linee di sviluppo nominate prima porteranno alla scrittura di software dedicato esplicitamente ad un certo tipo di computer, a seconda delle potenzialità di ogni tipo di elaboratore; anzi, nel campo dei computer quantici, Shor aveva già scritto alcuni programmi pensati proprio per sfruttare la complessità e la capacità di elaborazione in parallelo fornite dalla fisica quantistica.

A questo punto, vi sono varie domande da porsi: una migliore interazione tra software ed hardware potrebbe essere utile per lo studio dell'IA? Su questa strada, si potrebbero raggiun-

gere buoni risultati pur restando legati alla macchina di Von Neumann? Non è possibile progettare algoritmi ad alte prestazioni che trattino direttamente con l'hardware per risparmiare tempo e, a volte, anche denaro ? Non è possibile strutturare il software in modo che rispecchi maggiormente quei modelli di operazione che sono i nostri processi mentali, nella visione di Minsky o di Hofstadter ?

Simili interrogativi sono legati a doppio filo con lo studio del funzionamento della mente o del cervello nel campo delle scienze della mente.

NATURA E PROCESSI DELLA CONOSCENZA

La conoscenza viene divisa tradizionalmente in sensoriale e intellettuale, a seconda della “provenienza”; lo studio della natura del conoscere e del modo in cui noi conosciamo è una questione fondante per tutta la ricerca sull’IA. I processi di apprendimento dei Sistemi Esperti e i processi di visione o più in generale di sensazione dei robot sono una minima applicazione pratica della vasta teoria della conoscenza che è stata sviluppata dall’informatica e dalla cibernetica negli ultimi decenni. Questa ovviamente è una teoria funzionale, o, se fosse possibile dirlo, una teoria della pratica della conoscenza; quindi è ben distinta dalla gnoseologia, la parte della filosofia che tradizionalmente si occupa del conoscere umano, individuandone la natura e i modi. Proporremo dunque un breve percorso tra le concezioni gnoseologiche a nostro avviso più significative, per poi soffermarci sulla presunta natura della sensazione e dell’intellectione a livello neurale; quindi ci occuperemo del modo in cui queste nozioni sono state incorporate dall’IA, tramite i Sistemi Esperti e i Robot.

MENTE

Per secoli, anche prima che sorgesse il problema di trovare una spiegazione biologica al nostro conoscere il mondo, questo problema ha assillato chiunque si sia misurato con la riflessione sull’uomo e sul sapere. Per questo uno dei settori tradizionali della filosofia è la gnoseologia, che affronta proprio il tema del conoscere, nei suoi due aspetti della conoscenza tramite i sensi e della conoscenza tramite l’intelletto. In tempi più recenti il problema si è spostato sul piano scientifico, interrogando gli studiosi non più sulla natura della conoscenza, ma sui meccanismi biologici che la sostengono.

Ebbene, queste considerazioni sono importantissime per le ricerche sull’IA, in quanto una buona parte di tali studi sono rivolti ad ottenere un programma in grado di imparare, non solo da dati già codificati, per l’inserimento nella base di conoscenza, ma dati inseriti in un discorso, dati da astrarre e da comprendere; in futuro l’attenzione si sposterà sull’astrazione e la comprensione dei dati forniti dai sensi, che per ora vengono utilizzati solo dai robot, ma solo perché essi hanno obiettivi generalmente semplici e immediati, che non prevedono scelte ambigue o decisioni, ma solo procedure algoritmiche, complesse, ma determinate.

La storia della conoscenza

Il problema della conoscenza è un tema che attraversa la storia del pensiero umano dalle prime riflessioni dei Presocratici alle più recenti teorie di Hubert Dreyfus (autore di alcuni testi ottimistici quali *What computers can’t do*). La conoscenza è sempre stata distinta in sensoriale e intellettuale, ovvero derivata dai sensi e derivata dall’intelletto; attraverso i secoli molti filosofi hanno tentato di fornire una teoria che spiegasse come noi conosciamo la realtà, fornendo spesso un principio di base su cui la conoscenza potesse basarsi, come l’*anamnesi* platonica, il *parallelismo* spinoziano, il *giudizio sintetico a priori* kantiano, l’*intuizione* di Husserl, l’*esser-ci* di Heidegger, l’*analogia formale di fatti e proposizioni* di Wittgenstein.

In mezzo a queste differenti impostazioni, troviamo un tratto costante in quasi tutta la gnoseologia: alla base della conoscenza riconosciamo il fenomeno dell’apparire²⁸; ne troviamo traccia nel *non latere* agostiniano, nella concezione aristotelico - tomistica dell’identificarsi con il conosciuto, nella teoria hobbesiana in cui l’apparire è inizialmente attribuito alla sensazione, nella necessità husserliana di *andare alle cose*, descrivendo ciò che appare, prima di costruire un sistema.

²⁸ Fenomeno e apparire hanno radice comune, dal greco φαίνεσθαι : apparire; con questi termini è qui inteso ciò che appare, ciò che si fa presente al conoscente.

Vi sono certo filosofi che non prestano attenzione all'apparire, ma anche questi, quando parlano di conoscenza, si esprimono di solito in termini di vedere, come fa ad esempio Cartesio per il quale è possibile dire che "conoscere è vedere".

Per quanto concerne l'obiettivo che ci siamo proposti, non è necessario tracciare una dettagliata storia della gnoseologia: basterà evidenziare alcune tappe importanti, tenendo ben presente la centralità dell'apparire sopra esposta.

Nella filosofia antica possiamo trovare due momenti importanti nel pensiero di Platone e di Aristotele: il primo sottolinea l'importanza di una conoscenza che non sia vincolata alla sensibilità, ma all'intelletto e la teorizza introducendo il mondo delle idee, che noi conosciamo attraverso il ricordo (anamnesi) di ciò che l'anima ha potuto vedere prima di incarnarsi; il secondo introduce una unità conoscitiva, ma non di essere, tra il conoscente e il conosciuto, attribuendo all'intelletto umano un ruolo passivo (ricettivo) e postulando l'esistenza di una componente attiva, che ponga in atto il divenire conoscitivo dell'intelletto passivo nel conosciuto. Quindi, mentre Platone introduce l'idea, Aristotele introduce la forma.

Altra tappa importante è la filosofia di S. Agostino, il quale definisce la sensazione *passio corporis non latens animam*²⁹, mentre vede la conoscenza intellettiva come frutto dell'illuminazione divina.

Giungiamo, attraverso la ripresa di Aristotele ad opera di S. Tommaso, all'età moderna, in cui Galileo distingue tra qualità primarie, effettivamente presenti nell'oggetto, e qualità secondarie, causate nei nostri organi di senso dalle qualità primarie. Tra le prime l'estensione, la forma, il movimento, tra le seconde il colore, il suono, il sapore; questa distinzione rimane a fondamento di buona parte della successiva filosofia, la troviamo ad esempio in Cartesio e in Locke.

In questo periodo si sviluppano le due correnti dell'empirismo e del razionalismo; la prima ha notevoli conseguenze nel problema della conoscenza, riducendo progressivamente la sua area di applicazione fino a giungere con Hume ad affermare che noi possiamo conoscere solo le sensazioni, idee e percezioni, e che l'esistenza dei corpi o della sostanza spirituale è garantita solo dalla fede. Il razionalismo, invece, ha uno sviluppo a nostro avviso irrilevante per il problema che stiamo affrontando.

Si giunge così a Kant, che caratterizza il fenomeno come l'applicazione di 12 categorie a priori dell'intelletto al materiale sensibile organizzato nelle intuizioni a priori dei spazio e tempo, affermando che condizione indispensabile al nostro conoscere è la sensazione e limite inevitabile sono le strutture a priori della sensibilità e dell'intelletto, da cui noi non possiamo prescindere nel conoscere. Dopo Kant la gnoseologia ha un periodo di stasi, dovuto all'impostazione particolare dell'Idealismo che vede la ragione umana in grado di conoscere compiutamente l'assoluto.

Significativa la parentesi positivista che, nella foga di negare ogni metafisica, costituisce la metafisica della scienza e fonda ogni conoscenza sul dato reale, sull'esperienza e sulle sensazioni.

Nell'età contemporanea, troviamo di particolare interesse l'impostazione di Husserl, che afferma l'importanza dell'intuizione prima di tentare qualunque sistemazione conoscitiva; anche la ripresa operata da Heidegger, che pone alla base della conoscenza non più l'intuizione dell'oggetto, ma l'esser-ci (*Dasein*), l'essere immersi nella società e nei rapporti con altri esseri umani. Su altre posizioni troviamo Wittgenstein, che riconosce nella logica il denominatore comune della realtà: i fatti trascendono la logica, ma non possono prescindere da essa, che stabilisce la necessaria analogia tra fatti e proposizioni che li descrivono.

Possiamo vedere che la gnoseologia oscilla tra l'idea e la sensazione, l'oggettivo e il soggettivo; questo ci fornisce la possibilità di interrogarci su quale teoria della conoscenza possa meglio

²⁹"Modificazione del corpo che non sfugge all'anima"

integrarsi con l'IA, ma nel farlo si noterà che alle diverse linee di ricerca corrispondono diverse teorie gnoseologiche, creando la possibilità che futuri sviluppi della cibernetica e dell'informatica diano conferma o smentita ad alcune posizioni filosofiche.

La sensazione

I nostri sensi sono uno dei più complessi sistemi di ricezione-elaborazione mai prodotti, dalla natura o dall'uomo. Per rendercene conto esaminiamo il processo della visione

La luce raggiunge gli occhi e colpisce la retina, una struttura formata da tre strati di cellule che si trova sul fondo del bulbo oculare. La luce penetra fino al terzo strato. Qui si trovano cellule dette bastoncelli, che sono sensibili alle variazioni di luminosità, e coni, che reagiscono a luce di diverse lunghezze d'onda corrispondenti ai colori rosso, verde e blu. La luce scolora il pigmento presente in queste cellule. Questo genera un segnale che viene trasmesso alle cellule del secondo strato e da queste ad altre cellule ancora nello strato più esterno. Gli assoni di queste cellule si uniscono per formare il nervo ottico.

I milioni di neuroni del nervo ottico arrivano a una regione del cervello nota come chiasma ottico. Qui i neuroni che portano i segnali provenienti dalla parte sinistra della retina di ciascun occhio si incontrano e proseguono paralleli fino alla parte sinistra del cervello. Analogamente, i segnali provenienti dalla parte destra di ciascuna retina si uniscono e viaggiano verso la parte destra. Gli impulsi arrivano poi a un "ripetitore" situato nel talamo, e da lì i neuroni successivi li trasmettono a una zona nella parte posteriore del cervello nota come corteccia visiva.

Elementi diversi dell'informazione visiva viaggiano su strade parallele. I ricercatori sanno ora che la corteccia visiva primaria, insieme a un'area vicina, funge un po' da "ufficio postale" in quanto seleziona, smista e integra le molte informazioni portate dai neuroni. Una terza area percepisce i contorni, ad esempio i margini di un oggetto, e i movimenti. Una quarta area riconosce sia la forma che il colore, mentre una quinta aggiorna di continuo mappe di dati visivi per seguire i movimenti. Le ricerche attuali fanno pensare che ci siano ben 30 aree cerebrali diverse che elaborano le informazioni visive raccolte dagli occhi. Ma in che modo queste aree collaborano per dar vita a un'immagine? Come fa la mente a "vedere"?

L'occhio raccoglie informazioni per il cervello, ma è evidentemente la corteccia a elaborare le informazioni che il cervello riceve. Se scattate una foto, questa rivelerà particolari dell'intera scena. Se invece guardate la stessa scena con gli occhi, osservate consapevolmente solo quella parte su cui fissate l'attenzione. In che modo il cervello faccia questo rimane un mistero. C'è chi ritiene che questo sia il risultato di un'integrazione progressiva dell'informazione visiva in cosiddette "aree di convergenza", che permettono di paragonare ciò che si vede con ciò che si

conosce già. Altri suggeriscono che quando non si vede qualcosa di evidente, è semplicemente perché i neuroni che controllano la visione cosciente non producono impulsi.

La descrizione si ferma a questo livello perché da qui in poi interviene l'intelligenza a comprendere l'immagine, ad analizzarla, paragonarla e catalogarla con sistemi che sono tuttora un mistero, nonostante i molti studi condotti sull'argomento. Di altri sensi si sa molto meno, in quanto è più difficile distinguere la parte della sensibilità da quella dell'intelligenza.

Vogliamo ora puntare l'attenzione sul sistema che ci permette di unire le varie sensazioni per averne una unitaria: ad esempio, quando incontriamo



Figura 8 – Sistema talamo-corticale

una persona che ci stringe la mano salutandoci, noi dobbiamo concertare la sensazione tattile con quella uditiva e quella visiva. Questo processo viene garantito dal sistema talamo-corticale: la corteccia contiene aree dedicate ai cinque sensi che sono collegate ad una zona interna detta talamo, composta da un sistema specifico e da uno non-specifico; da entrambi parte un collegamento per ogni parte della corteccia. Il sistema non-specifico invia impulsi a ogni zona della corteccia con frequenza di 40 Hz, mentre il sistema specifico invia solo se viene stimolato dall'arrivo di una sensazione; quando entrambi gli stimoli giungono ad una determinata zona della corteccia, questa rimanda l'impulso al talamo e noi avvertiamo la sensazione.

Questo sistema non spiega l'elaborazione delle sensazioni, campo in cui gli studi, come già accennato sono ad un livello ancora embrionale, lasciando aperti alcuni interrogativi: esistono aree dedicate specificamente all'analisi di particolari configurazioni sensoriali (visive, uditive, etc.)? Esistono aree (eventualmente dinamiche) che sovrintendono alla fusione di determinate sensazioni per riconoscere un oggetto (colore, sapore, odore per riconoscere un frutto, ad esempio)?

L'astrazione

L'astrazione è solitamente intesa come capacità di separare le caratteristiche universali da quelle individuali, ottenendo dal dato reale, sensoriale, particolare, il concetto universale, intellettuale che rappresenta una classe di oggetti tra loro analoghi per quanto riguarda una certa caratteristica. Ad esempio dall'esperienza che abbiamo di Pietro, Paolo, Giovanni, Marco, etc. possiamo astrarre il concetto di *uomo*, che indica una classe a cui tutti i suddetti appartengono in quanto tutti sono simili per quel che concerne l'essere uomo.

L'astrazione si basa dunque sul cogliere le somiglianze e le differenze tra più individui (realtà particolari) e porta a creare classificazioni di oggetti a cui riferire poi la nostra futura esperienza.

E' importante sottolineare che l'astrazione non è un processo intenzionale, almeno nella sua forma naturale (o totale, secondo la terminologia della filosofia scolastica), ovvero il considerare tutto l'oggetto, nella sua interezza, ma in modo indeterminato, prescindendo dalla sua individualità, come nell'esempio precedente; questo è un carattere importante, poiché conferma in un certo modo l'idea che l'astrazione sia alla base di tutta l'attività intellettuale umana.

Possiamo affermare questo grazie a due osservazioni: l'intelligenza "lavora" solo con concetti universali (con un diverso grado di universalità, se vogliamo, ma comunque universali) e l'universale è prodotto dall'astrazione, mentre il senso ci fornisce il particolare.

Esiste anche un secondo modo di astrazione, che è l'astrazione artificiale (o formale, sempre secondo gli scolastici), che consiste nel astrarre ulteriormente, da un tutto già generalizzato, una particolare caratteristica o determinazione: proseguendo nell'esempio di prima, dal concetto di uomo, potremmo isolare la determinazione di attività affettiva e volitiva.

Questa astrazione è a volte attivata dall'intenzione di astrarre, qualora si voglia conoscere più pienamente una certa realtà, studiandone le singole determinazioni, che devono essere allora astratte dal concetto generale e non sempre si presentano con evidenza alla mente dello studioso.

Questa attività è il passaggio intermedio tra la sensibilità e l'intelletto, o almeno ha buoni caratteri per esserlo, almeno a livello psicologico, se non a livello filosofico.

Finora la neurofisiologia non è riuscita ad individuare una spiegazione accettabile di tale processo, sebbene siano recentemente emerse delle teorie che sembrerebbero poter spiegare, almeno in generale un simile processo: supponendo il cervello un sistema caotico di impulsi nervosi, uno stimolo sensoriale introdurrebbe uno squilibrio nella uniformità che è propria del caos e creerebbe la possibilità di una degenerazione in uno stato particolare, più ordinato, detto attrattore; attrattori simili significherebbero somiglianze nella percezione sensoriale, fi-

no a diventare una sorta di catalogazione; allo stesso modo un'idea o un'altra potrebbero produrre altri particolari attrattori, magari in aree diverse del cervello, stabilendo somiglianze e differenze.

Il quesito principale è dunque inerente alla natura del sistema e del criterio che presiedono all'astrazione e, come è intuibile, il rapporto che c'è tra l'astrazione e gli altri processi dell'intelligenza.

La comprensione

Abbiamo scelto di intitolare così questo paragrafo, invece di proseguire con *L'intelligenza*, per limitare il campo di indagine alle questioni più problematiche nel campo dell'IA: i processi logici che avvengono nel cervello e che trovano corrispondenza in molte attività intellettive, seppur ancora argomento di studio nel loro effettivo realizzarsi a livello neurale, sono facilmente riproducibili su un computer; ciò che da problemi alle macchine è la comprensione.

Dobbiamo però specificare cosa intendere con questo termine:

- la comprensione dei dati sensoriali, ovvero il riconoscimento visivo di volti e oggetti, uditivo di voci, suoni e rumori;
- la comprensione dei significati dei dati sensoriali, cioè la capacità di interpretare il linguaggio scritto e orale;
- la comprensione di sé, ovvero la coscienza del proprio essere, del proprio sentire e del proprio pensare.

Non analizzeremo questi punti nella loro interezza, poiché per ciascuno di essi si trovano facilmente decine di riferimenti in filosofia, psicologia, neurofisiologia e persino in fisica e in matematica; una simile esposizione eccede gli scopi e le possibilità di questo scritto. Dunque ci limiteremo a sottolineare i punti problematici di maggiore importanza, senza entrare nei dettagli delle singole questioni.

Per quanto riguarda i dati sensoriali, la loro comprensione è in parte dovuta, per quel che se ne sa ora, a processi che rientrano ancora nel dominio della sensazione (prima dell'arrivo degli impulsi al sistema specifico del talamo), in parte ad attività di astrazione e in parte a processi che sono ulteriori rispetto all'astrazione, quindi uno studio dettagliato di questo punto impegna la psicologia su vari livelli e ottiene aiuti dalla neurofisiologia solo per alcuni punti ben delimitati.

Si è giunti a determinare dove e come avvengano alcune preliminari e molto grezze elaborazioni dell'informazione all'interno della corteccia, volte ad esempio a determinare la posizione di corpi scuri o chiari su fondo contrastante o la presenza di particolari frequenze molto evidenti o di picchi di suono in un rumore di fondo sufficientemente basso; ma tutto questo non spiega come noi possiamo riconoscere un volto da un altro o un oggetto da un altro, prima ancora di abbinarvi un'idea, un nome, un'azione, un ricordo. E questo è il problema principale quando si tratti di lavorare con i dati sensoriali su una macchina.

Per quel che riguarda la comprensione del significato di particolari dati sensoriali, la ricerca è tuttora ferma a risultati poco felici: sono state individuate le due aree del cervello responsabili della comprensione e della produzione del linguaggio, ma è ignoto il loro funzionamento, se non a livello molto generale; la nostra attenzione si incentra sul linguaggio poiché, seppure in qualità di ipotesi, in IA si tende a valorizzare l'idea che le attività di riconoscimento di schemi astratti, significati non immediati (non come fumo -> fuoco, ma come luce rossa -> stop) e, in generale, un'interpretazione ad alto livello dei dati sensoriali, siano legate alle abilità linguistiche. I meccanismi di produzione e comprensione del linguaggio sono uno dei misteri meno penetrati delle neuroscienze, poiché non si è ancora compreso dove o come un flusso continuo di suoni diventi un insieme discreto di segni e quindi un esiguo gruppo di idee (o viceversa); molti credono che, una volta sciolto questo nodo, il problema del linguaggio sarà essenzialmente risolto.

L'ultimo punto è, nell'opinione di molti, il più importante e viene unanimemente ritenuto il più complesso da analizzare, comprendere ed eventualmente riprodurre. Nell'uomo, da un punto di vista neurofisiologico, la funzione di "senso interno" è svolta dal sistema non-specifico del talamo, che periodicamente manda stimoli a tutta la corteccia, stimolato a sua volta dagli avvenimenti interni (al cervello); questa però non è una spiegazione dei fenomeni che normalmente passano sotto il nome di coscienza (o più specificamente autocoscienza), è solo un minimo riscontro fisico dei fenomeni mentali. Sebbene si abbia un'idea, seppure approssimata, di come il cervello senta e elabori, non si è ancora riusciti a trovare un corrispettivo, a qualunque livello, del *sapere di sentire* e del *sapere di pensare*, per non parlare del *sapere di esistere*, non come corpo, ma anche e soprattutto come coscienza; questa comprensione è forse ciò che permette alla mente umana alcuni processi di pensiero ancora inspiegati: da un punto di vista matematico la mente è un sistema che si autocomprende infinitamente, cioè un sistema che capisce se stesso e sa trascendersi e assimilarsi, senza comunque potersi osservare, infatti io so di sentire, ma so di sapere di sentire e so di sapere di sapere di sentire, etc.; quindi non siamo limitati ad un certo numero di livelli di comprensione e solitamente li trascendiamo tutti, infiniti quanti sono, giungendo ad affermare di sapere di esistere, percependo una costanza nel continuo fluire di questo sapere e questo sentire e nel progressivo trapassare ogni forma di coscienza di un particolare sapere o sentire; la coscienza di questa costante è l'autocoscienza, poiché non è certo errato identificare l'io in questo permanere. Questa impostazione permette di spiegare la coscienza, ma perde ogni utilità in campo scientifico, quando si debba trovare questo infinito, che parrebbe essere in qualche modo misterioso potenziale e attuale insieme. Infatti un sistema matematico così descritto sarebbe inconciliabile con gli attuali risultati in logica e in teoria degli insiemi, dovendo da una parte essere ampliato all'infinito dalla comprensione di se stesso e dall'altra portare ad uno stato finito e determinato di autocoscienza, ovvero di conoscenza e coscienza di sé.

MACCHINA

In questa parte vorremmo presentare quanto è stato fatto e quanto ci si propone di fare per ottenere dei computer conoscenti e forse coscienti. Questa panoramica sarà assai breve, in quanto, oltre ad alcuni punti che dovremo tacere per l'elevata complessità tecnica (soprattutto per quanto riguarda i sensi), troveremo che ben poco si è ottenuto in tali campi di significativo e definitivo.

I sensi di un computer

I *sensi* di un computer sono quei dispositivi che gli permettono di ricevere informazioni dall'esterno, informandolo sull'ambiente circostante. In questo senso la tastiera, il mouse, lo scanner, le videocamere digitali, i microfoni, sono organi di senso, sebbene non trovino tutti corrispondenza con i sensi dell'uomo.

Nello studio e nella progettazione dei robot, campo decennale della cibernetica e recentemente sfruttato anche dall'IA, si tenta invece di dotare la macchina di sensi simili a quelli umani. E' interessante notare che i progressi più significativi sono stati ottenuti nello sviluppo di sensibilità tattile e olfattiva-gustativa, ovvero quei sensi che nell'uomo sono, per così dire meno sfruttati sul piano intellettuale. Questo indica quanto sia difficile separare la pura attività di percezione sensoriale dalla attività di analisi e sintesi operata dal cervello per comprendere lo stimolo. Inoltre, parte dell'elaborazione viene svolta al livello degli organi di senso ed è questo che i ricercatori tentano di ottenere; nella retina vi sono cellule sensibili più alla luminosità che al colore o viceversa, mentre sotto la pelle vi sono diversi tipi di nervi, per registrare varie sensazioni tattili, ed infine, la lingua è divisa in zone che rispondono alle diverse componenti del gusto. Questa specializzazione delle unità sensoriali è stata tentata senza molto successo e si è dovuto affidare queste operazioni ad una elaborazione preliminare a livello di software;

comunque, bisogna segnalare che ora possiamo costruire dei robot che riconoscano diversi oggetti (le cui caratteristiche devono però essere immagazzinate in memoria) che differiscono per pochi particolari, grazie a vista, tatto e olfatto-gusto, sebbene sia necessario precisare che le distinzioni più precise vengono fatte in base ai dati tattili e olfattivi-gustativi, mentre dalla vista si ottengono ancora dati piuttosto grossolani.

Il senso dell'udito è quello che presenta i maggiori problemi, in quanto pare totalmente affidato alla successiva elaborazione del cervello, senza alcun processo preliminare a livello dell'orecchio, se non la regolazione della tensione del timpano.

Gli espedienti tecnologici utilizzati sono tra i più vari e si dividono in tentativi di imitazione degli organi umani e pesante sfruttamento delle nuove tecnologie, per battere l'ingegnosità della natura con la quantità di informazioni e la velocità di elaborazione.

L'astrazione

Questo passaggio che prima abbiamo sottolineato come importante anello di congiunzione tra il senso e l'intelletto, manca totalmente nelle macchine e non è nemmeno in fase di sviluppo: non si è ancora riusciti a produrre un programma che riesca ad astrarre da un oggetto un concetto, anche in termini di classificazione o di confronto. Le reti neurali, che pure ottengono buoni risultati per quanto riguarda il riconoscimento di oggetti e situazioni, non utilizzano questa fase, ma passano immediatamente dallo stimolo alla risposta, senza classificare lo stimolo e dedurne la risposta. Probabilmente, il processo astrattivo è necessario in sistemi complessi in cui lo stimolo e la risposta sono composti da migliaia di azioni elementari, tra loro concatenate e relazionate; il problema principale è che l'astrazione dovrebbe produrre l'universale (cioè la classe, l'insieme) senza avere nulla in precedenza, mentre su una macchina è necessario introdurre delle classi base o dei metodi di confronto o degli esempi di classificazione o altro materiale che faccia da fondamento alle successive inserzioni di oggetti.

Inoltre, il normale funzionamento del computer consiste nel tenere in memoria i singoli oggetti, mettendo a disposizione un nome di comodo con cui riferirsi a un certo gruppo, ma mantenendo il "ricordo" di ogni singolo individuo, senza riassumerne le caratteristiche nel nome della classificazione.

E' assai probabile che l'astrazione sia una delle prove del fuoco per l'IA: se mai verrà risolta e riprodotta in modo accettabile senza però portare a qualche traccia di intelligenza (in qualunque senso la si voglia intendere), molti si riterranno giustificati a credere che una macchina possa solo emulare l'uomo, senza simularlo.

Memoria e riconnessioni

La comprensione è un tema talmente vasto e in alcune sue parti così poco scientifico che è pressoché inutile parlare di ricerche sulla comprensione intellettuale di un computer, piuttosto si tende a ricondurre questa abilità a due componenti, importanti ma non uniche, che sono la memoria e la capacità di ricombinare i dati acquisiti in modi nuovi e relazionati all'esperienza.

Certo, queste facoltà dipendono strettamente dall'astrazione e quindi il nostro discorso sarà forzato e parziale, guardando più ai risultati ottenuti che non ai meccanismi utilizzati e avvallando un certo indirizzo comportamentistico da cui queste ricerche ad alto livello non sono comunque scovre.

La memoria di un computer intesa in senso classico è assolutamente inutile per l'IA, in quanto presto si riempirebbe inesorabilmente dei soli dati sensoriali; per questo sono state elaborate tecniche di classificazione anche molto complesse che registrino le analogie e permettano riferimenti incrociati multipli, di modo che un'esperienza possa essere ricostruita in base ad elementi di esperienze passate, prestando attenzione anche a non registrare troppi dettagli inutili e a non spingere troppo in alto le generalizzazioni.

Risultati significativi sono stati ottenuti dalle reti neurali, ma non si tratta di vera e propria memoria, anche se è in fase di test un tentativo di unire le capacità di risposta delle reti neurali con i moderni programmi di memorizzazione per livelli e riferimenti, creando conformazioni e raggruppamenti di nodi in relazione a particolari caratteristiche e variandone la risposta a seconda della determinazione della caratteristica (una zona addetta al colore che dia varie risposte a seconda del colore visto, ad esempio).

Qui entra in gioco l'altra facoltà suddetta, la capacità di riconnessione, legata non solo ai collegamenti tra i singoli neuroni, come già avviene nelle reti neurali, ma anche ai legami tra le idee, i concetti, le situazioni. Questa capacità della mente umana è probabilmente legata all'elaborazione parallela che il cervello può operare e quindi alla contemporanea attivazione (per somiglianza, quindi, in certo modo, per abitudine) di più aree cerebrali, legate a vari simboli.

AUTOCSCIENZA

Preferiremmo tralasciare lo studio dell'autocoscienza e sostituirlo con una serie di questioni sulla possibilità dell'intelligenza artificiale. Questo perché il problema, oltre ad avere radici filosofiche profonde e complesse, trova agganci con tutto quanto trattato finora e rappresenta in certo modo il cuore della ricerca sull'IA, sebbene non ci sia alcun progetto che abbia come suo scopo precipuo la produzione di un computer autocosciente.

Vedremo che le questioni poste, pur partendo da ciò che abbiamo esposto in queste pagine, toccheranno spesso più o meno evidentemente il tema del sapere di sapere e del sapere di essere, oltre alla capacità di trascendersi e autocomprendersi, di cui abbiamo accennato in precedenza.

Struttureremo questa parte proponendo 10 tesi a favore dell'IA e 10 risposte che le mettono in dubbio.

1. L'IA è possibile nella sua ipotesi debole in quanto quella parte dell'attività mentale che viene ritenuta non algoritmica è facilmente implementabile tramite procedure euristiche (cioè tramite l'introduzione di fattori pseudo-casuali nei calcoli) e moduli di logica fuzzy, seppur con alti costi tecnici.

Assunta questa ipotesi della matematicizzabilità anche di molteplici varianti in deterministiche, il problema è quello di vedere il loro possibile numero. Se esso non è infinito e quindi indefinibile, di nuovo, l'IA ricadrebbe sotto la prospettiva di un calcolatore comunque programmato tecnologicamente.

Infatti l'ipotesi debole afferma che la macchina può emulare, ma non simulare, e quindi tale problema, della equivalenza di un sistema finito con una macchina, intacca l'ipotesi forte, ma non quella qui sostenuta che rimane valida, seppur di non notevole importanza.

2. La pretesa di uniformità delle procedure della macchina ai processi mentali è inconsistente (e quindi non fornisce un valido metro di giudizio), in quanto ancora non conosciamo con precisione la natura dei secondi e quindi non possiamo istruire paragoni se non su supposizioni e teorie non confermate.

D'accordo, tuttavia il consenso rafforza di nuovo la congettura dell'incommensurabilità assoluta tra IA e IN.

Questa è una critica la cui validità è legata alle convinzioni di ciascuno e quindi perde parte della sua validità quando la si consideri da un punto di vista oggettivo, chiedendo una spiegazione dettagliata di quali elementi rafforzino la detta congettura.

3. L'intenzionalità, spesso vista come mancante alla macchina, non è in alcun modo rilevabile empiricamente e oggettivamente tramite procedure scientifiche di tipo fisico-matematico: noi ci attribuiamo come caratteristica fondamentale questa intenzionalità sulla base del nostro senso interiore, ma in realtà non abbiamo alcun modo di provare agli altri (né a noi stessi) che siamo esseri umani liberi e intelligenti anziché automi perfettamente programmati ad ingannare se stessi e gli altri.

A questa ipotesi si può rispondere osservando che l'intenzionalità del soggetto umano, sia essa cognitiva o volitiva, proprio perché sfugge a una determinazione oggettiva in senso fisico-matematico, esprime e rivela la sua natura propriamente umana di non finitezza ed è rilevabile in modo scientifico in senso analogico, ricorrendo cioè a metodi che per rigore e oggettività si differenziano dal modello matematico-fisico. Questo, lungi dall'essere una ragione contro l'IN, è invece una delle sue giustificazioni più rigorose. In altre parole prendiamo coscienza di noi stessi come enti viventi, senzienti, intelligenti e volenti in un modo che il calcolatore non potrebbe mai concepire e realizzare proprio perché automa tutto costruito e non soggetto umano. Si aggiunga che un automa anche perfezionatissimo non po-

trebbe mai far queste osservazioni in quanto incapace di autoriflessione e di autocoscienza, oggetto sia di fenomenologia, sia di logica, sia di ontologia. Questa risposta parte dalla psicologia ma apre inevitabilmente all'esperienza morale come caratteristica specifica propria della persona umana: soggetto capace di intendere il vero e di volere il bene o anche drammaticamente di contraddirli; "non faccio il bene che voglio, ma faccio il male che non voglio". Qualsiasi computer, essendo sine macula originali conceptum, non può nemmeno peccare.

Certamente è vero che in alcuni campi l'intenzionalità è valutata con pretese di oggettività secondo metodi analogici, ma questo suggerisce un legame tra analogicità, intenzionalità e autocoscienza, però non lo prova.

4. Per quanto riguarda la struttura fisica del supporto all'algoritmo: la struttura della mente umana si basa su cellule che hanno un comportamento da interruttori e che sono molto più lenti dei circuiti integrati di un computer, quindi è lecito ritrovare la differenza qualitativa tra questa e la macchina nella complessità della struttura e delle interconnessioni, oltre che nella abilità di lavorare in parallelo, senza limitarsi al processo sequenziale della strozzatura di Von Neumann.

Alla quarta ipotesi, fondata sulla maggior complessità della mente umana rispetto a quella attuale di un artefatto, si può rispondere che la possibilità di raggiungere in ipotesi per via tecnologica le complessità della prima non fa nessuna obiezione alla diversità di natura specifica per il fatto che trattasi di complessità puramente quantitativa e materiale. Tuttavia anche nel caso in cui il potenziamento quantitativo della complessità materiale facesse insorgere la novità di un salto di qualità e quindi di un mutamento di natura specifica al punto da "creare" per via tecnologica la mente umana, tale progetto si presenta quantomeno infinito proprio per la natura anche quantitativa dell'oggetto da costruire.

Da ultimo è interessante che il teorema di Gödel insegna che nessun sistema assiomatico è in grado di autofondarsi se non rimandando ad un sistema, ulteriore ed esterno che lo fonda, all'infinito. Ciò mostra come anche in matematica l'uomo non riesca a "comprendere" esaustivamente ciò che la sua mente produce. In tal senso la mente umana pare non sopportare alcuna chiusura, ma essere invece strutturalmente aperta all'infinito.

Qui, possiamo solo evidenziare come anche i matematici siano in dubbio sulla possibilità di applicare il T. di Gödel a sistemi così complessi e non totalmente noti come la mente: forse la complessità offre una via di fuga che tramite la possibilità del sistema di trascendersi porta a intuire la verità senza doverla dimostrare internamente. Questo comunque lascia valide le altre critiche.

5. La creatività e la fantasia che vengono anch'esse viste come tratto distintivo dell'intelligenza umana nella sua manifestazione del pensiero divergente, possono essere riprodotte su un calcolatore, almeno in teoria, avendo a disposizione le adeguate risorse computazionali ed energetiche.

Ma chi mette a disposizione tali risorse, stabilendone o meno l'adeguatezza? Certamente una mente che vede il problema oltre il problema, a distanza, dal di fuori. Una mente quindi capace di trascendere il dato in un modo ancora una volta così autoriflessivo da essere impossibile a un manufatto tutto immanente.

Facciamo solo notare come l'ultima affermazione sia ancora una volta sostenuta da convinimenti personali e teorie filosofiche.

6. L'autocoscienza non difetta alle macchine, unicamente non è implementabile in quel livello di intreccio tra piano fisico e piano interiore che si vede nell'essere umano, ma può comunque essere spinta in profondità, acquisendo una chiarezza anche superiore a quella possibile per l'introspezione umana.

Posto anche che si possa produrre l'autocoscienza per via tecnologica, un'autocoscienza prodotta è sempre qualcosa di fatto e finito, un intelletto passivo, incapace di svolgere quell'attività di trascendenza sopra richiamata. Un'autocoscienza cioè che non può diventare auto all'infinito come accade all'io dell'uomo. Una coscienza cioè che non può dire "io" a livello di successivi non finiti orizzonti.

Come molte critiche all'IA, anche questa si basa sul presupposto dell'inferiorità della macchina e delle caratteristiche non finite della mente umana; questi aspetti non sono ancora stati provati, anche se fortemente sostenuti.

7. L'IA è possibile, alla luce degli ultimi studi sulla teoria degli attrattori caotici che potrebbero giocare un ruolo molto importante nei processi cerebrali e che permetterebbero di matematizzare in maniera abbastanza completa le attività neurotiche, fornendo una buona base per le indagini di neurofisiologia. Questo permetterebbe, una volta che la tecnologia avesse fornito complessità di interconnessione sufficiente, di impostare matematicamente il controllo di una simile "mente elettronica".

Si risponde rimandando alle ipotesi precedenti. Nessuna complessificazione materiale supera il livello di sommatoria quantitativa per farsi prodotto qualitativamente diverso in quanto capace di produrre altro da sé.

L'accostamento di questa ipotesi alle precedenti è riduttiva in quanto non si parla di progressiva complessità, ma di un salto qualitativo, seppure a livello software.

8. La natura duale di anima e corpo non inficia necessariamente la possibilità di creare una mente artificiale, ma rileva unicamente che dalla materia possono prodursi fenomeni mentali che non sono materiali, secondo la tesi di Parfit, che però non nega la possibilità che questi fenomeni si producano da processi fisici, chimici e biologici.

Qui occorre essere più materialisti dei materialisti, negando l'errore di Cartesio, cioè la separazione tra anima e corpo, e vedendo l'anima come forma del corpo nelle sue dimensioni vegetativa, sensitiva, intellettuale e volitiva, riproducibili forse per via di ingegneria genetica su base comunque naturale. Ma non per via tecnologica, su base tutta artificiale.

Il fatto che Parfit sostenga che a livello fisico-chimico si pongono le basi anche delle attività mentali superiori, è una condizione necessaria, ma non sufficiente, per spiegare le medesime. Per cui come osserva Leibniz la falsità nella conoscenza non è tanto in ciò che si afferma, quanto piuttosto in ciò che si nega.

Tutta questa critica si regge sull'affermazione, forte e contestata, che l'anima sia la forma del corpo (e che l'anima esista, prima ancora).

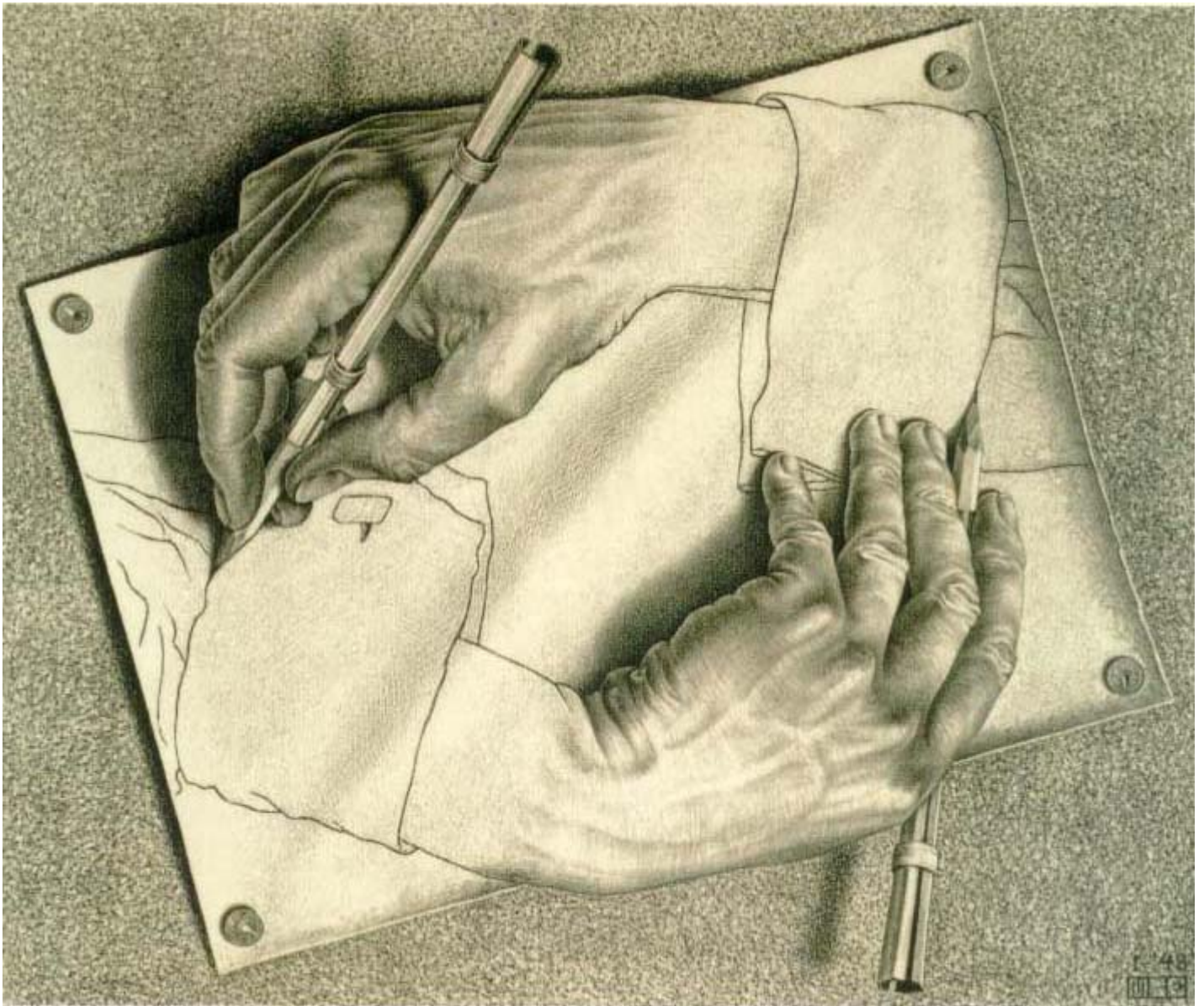
9. L'ipotesi del monismo emergentista, affermata negli ultimi decenni, fornisce una base sperimentale per comprendere i fenomeni mentali in base ai processi biologici.

Quanto a questa nona ipotesi, vale quanto appena detto. Occorre tuttavia evitare il cortocircuito per cui dal legittimo riduzionismo metodologico si passa all'illegittimo riduzionismo ontologico: poiché i fenomeni mentali sono prodotti da processi biochimici, allora non sono altro che processi biochimici.

10. Nelle opere di Dreyfus si sottolinea l'importanza del contesto sociale per lo sviluppo dell'io nelle forme dell'intenzionalità e dell'autocoscienza e si pone così una scusata per l'IA: di non poter calare una mente elettronica in relazioni sociali con altri "simili".

Qui, dove si mostra come nel nostro esserci sia strutturale il rapporto intersoggettivo umano, si può rispondere che un tipo di rapporto simile potrebbe al limite essere implementato pure nel computer, ma sarebbe sempre anzitutto l'intersoggettività umana trascendente propria degli sperimentatori.

Qui concludiamo questa sezione su autocoscienza e autoreferenza, riportando una polemica definizione di intelligenza, data da un sostenitore dell'IA, e un esempio di ciò che, secondo i detrattori, un computer non arriverà mai a fare né a comprendere .



“L'intelligenza è tutto ciò che un computer non riesce a fare”

CONCLUSIONI ?

Questo titolo ha due caratteristiche che meritano una spiegazione: innanzitutto è al plurale, poiché abbiamo sollevato, in queste pagine, così tanti interrogativi diversi che non è possibile considerare un'unica, univoca e definitiva conclusione; inoltre è una domanda, non una affermazione, in quanto non pretendiamo affatto di fornire delle conclusioni, ma semplicemente cerchiamo di riassumere le questioni sollevate finora, esponendo le possibili conclusioni che si possono trarre sul problema dell'IA, senza per ora prendere posizione.

Vogliamo, dunque, parlare delle principali ipotesi teoriche e di lavoro che sorreggono la ricerca in IA, ma non intendiamo, né potremmo in tutta onesta, avvalorarne una piuttosto che un'altra, poiché la stessa comunità scientifica non prende posizione ufficiale, nonostante molti cerchino di diffondere le proprie convinzioni, proprio perché non si hanno criteri certi e sicuri in base ai quali operare una scelta oggettiva e definitiva.

L'anima

Ebbene, l'anima è l'argomento favorito di coloro che riconoscono all'intelligenza umana una superiorità qualitativa inarrivabile e tendono a sottolineare come una macchina resti pur sempre una macchina, ovvero sia artefatto finito e determinato.

L'anima è qui intesa, in una visione tomistica, come forma del corpo e quindi come elemento di superiore unità delle parti, sia dal punto di vista fisico-organico, sia dal punto di vista intellettuale-razionale: se il cervello costituisce la base fisica del nostro pensiero e la mente gioca il ruolo di componente sensibile, intellettuale, intenzionale, volitiva, autocosciente, allora l'anima è quel principio che permette a tutto ciò di essere *uno* in quanto essere umano. In questo senso l'anima è un irrinunciabile elemento di unità, spirituale, non riducibile in termini sperimentali, ma necessario.

L'anima così pensata è una possibile conclusione delle ricerche teoriche dell'IA, poiché essa pone definitivamente la differenza qualitativa in un quid spirituale, che non è misurabile, quantificabile, né quindi riproducibile tecnologicamente e dunque segna un marcato e insuperabile limite tra l'intelligenza naturale e l'intelligenza artificiale, riconoscendo la superiorità della prima sulla seconda, in virtù dell'elemento spirituale che la forma e la costituisce unitariamente.

Nell'ipotesi dell'anima, il problema dei rapporti mente-corpo viene a cadere, come privo di senso, poiché l'anima è proprio ciò che garantisce l'unione tra i due, ma questo è solo uno spostare il problema: ora, anziché dover definire i rapporti tra mente e corpo, dovremo ricercare un possibile legame tra il corpo e questa sostanza definitivamente spirituale; tutto ciò è risolto qualora si pensi all'anima come forma del corpo, in senso aristotelico-tomistico, privando di senso questo dualismo, riconoscendolo in ogni cosa, poiché tutto è dato dall'azione della forma sulla materia.

Rimarrebbe il problema della collocazione della mente, che potrebbe essere un fenomeno dell'anima, come anche una produzione del cervello completata e resa autoriflettente dall'unificazione formale dell'anima, ma a questo punto sarebbe arduo distinguere gli apporti del cervello dagli apporti dell'anima e diverrebbe comunque irrilevante la riproduzione della mente così concepita su un computer, poiché non sarebbe più intelligenza, ma solo capacità di pensiero logico o analogico, finite, limitate e ignare di essere un Sé.

Abbiamo già accennato a come la questione dell'autocoscienza venga risolta affidando questa all'esclusivo contributo dell'anima, in quanto elemento unificante e quindi in grado di far cogliere alla mente la propria "ità", il proprio esistere come realtà singola, autonoma e cosciente.

Anche il problema della conoscenza verrebbe affidato all'anima, almeno nella sua parte di coglimento della realtà sensibile e comprensione delle realtà intellettuali, poiché essa fornirebbe il collegamento tra il sentire del corpo e l'*intelligere* della mente³⁰.

L'accettazione di questa ipotesi svuoterebbe dunque di qualsiasi ulteriore significato la ricerca secondo l'ipotesi forte dell'IA, in quanto la possibilità di ottenere un computer intelligente dipende dalla nostra capacità di riprodurre su silicio l'anima, abilità questa che noi non possediamo e che sembra essere intrinsecamente impossibile, vista la suddetta definizione.

L'uomo di Turing

Per sviluppare alcuni dei suoi più importanti lavori, Turing si servì di una ipotetica macchina a stati finiti e memoria e tempo illimitati; venne detta Macchina di Turing. L'ipotesi dell'uomo di Turing sostiene l'equivalenza (l'isomorfismo) tra la Macchina di Turing e la mente umana e dunque la natura algoritmica del pensiero umano. In quest'ottica, l'unica differenza che può esserci tra l'intelligenza naturale e quella artificiale sta nella bontà del supporto e nella sua maggiore o minore adeguatezza al tipo di algoritmo che deve eseguire; una tale posizione permette all'IA di procedere semplicemente per prove ed errori, cercando di riprodurre sempre più funzioni ed aspetti della mente umana, così da arrivare prima o poi ad aver implementato su di un computer ogni singola abilità della mente umana e quindi ad averne ottenuto una copia algoritmica, che in base all'ipotesi assunta può essere considerata una mente a tutti gli effetti.

Chi assume una tale prospettiva rimane però privo di un qualunque criterio pratico e operativo (se non quello generale del metodo scientifico) per affrontare la ricerca, poiché questa ipotesi non può spiegare i fenomeni conoscitivi o almeno fornire un riferimento per lo studio, può solo assicurare che esiste un algoritmo ad essi equivalente che ad un determinato input (stimolo) dà il loro stesso output (risposta).

In una simile ipotesi, si dovrebbe puntare alla scoperta dell'algoritmo mentale, tramite l'unione di piccole parti, ma anche tramite la progettazione di menti in miniatura, che man mano vengano ampliate e diventino sempre più complesse. Questa è già la seconda volta che le conclusioni da trarre da questa ipotesi sono legate ad altri tipi di posizione filosofico-intellettuale: ciò mostra quanto complicato sia un tale campo d'indagine, dove intervengono svariate discipline, i cui esperti non sempre sono disposti a studiare seriamente anche solo i rudimenti delle materie che non conoscono, per poter esprimere con coerenza la propria opinione sull'argomento.

Il Tutto

Una delle posizioni da chiarire a cui facevamo precedente accenno è lo schierarsi per l'olismo o per il riduzionismo. L'olismo vuole che il tutto sia qualitativamente superiore alla somma delle parti e non riduce questo valore aggiunto alle relazioni tra le varie componenti, ma lo vede emergere proprio dal Tutto che ne risulta come peculiarità dell'oggetto completo in sé.

Da questo punto di vista la ricerca rimarrà infruttuosa finché non avrà trovato, riprodotto e unito tutti gli elementi che concorrono alla formazione dell'intelligenza umana; si noti che nulla si dice sulla natura di tali elementi, lasciando aperta la possibilità che alcuni siano introvabili o irriproducibili. Anche questo modo di vedere deve sorreggersi tramite altre ipotesi, che riguardino campi che questa lascia inesplorati.

La visione che si prospetta non è certo ottimistica, poiché non si dà modo per controllare la bontà del lavoro, per dir così, *in itinere*, né si può indicare precisamente quali fattori influ-

³⁰ Anche se si deve riconoscere che su questo punto sono stati molti i tentativi, non sempre felici, di spiegazione di questo rapporto conoscitivo tra l'anima e i sensi. Ad esempio S. Agostino definiva la sensazione *passio corporis non latens animam* (modificazione del corpo che non sfugge all'anima).

scano sulla diversità di natura del Tutto risultante. Queste difficoltà lasciano intravedere ai detrattori dell'IA la possibilità che l'ipotesi olistica giunga a negare la pretesa dell'intelligenza artificiale di equiparazione all'intelligenza naturale; ciò ha fatto così che l'olismo sia pian piano venuto a coincidere nella mente di molti con la negazione dell'IA, ma bisogna osservare che questa posizione non sancisce affatto l'impossibilità di una riproduzione della mente tramite una macchina, semplicemente pone delle condizioni ardue al suo svilupparsi operativo.

Le Parti

Esaminiamo ora l'ipotesi opposta: il riduzionismo, secondo il quale una realtà complessa si riduce alla somma delle proprie componenti. E' evidente da quanto detto a proposito dell'olismo che il riduzionismo rende la vita assai più facile all'intelligenza artificiale, in quanto le richiede di ridurre l'intelligenza naturale nelle sue componenti e di provarle singolarmente ad una ad una, unendole poi per formare una copia dell'originale; certo non dice nulla di queste componenti, se siano di natura biologica o meno, ma è certo che, applicando sempre un criterio riduzionista, si arriverà a definire l'intelligenza in base a fenomeni mentali così elementari da poter essere collegati causalmente a reazioni chimiche e fisiche del cervello.

Del resto, anche l'identificazione semplicistica del riduzionismo con l'ipotesi forte dell'IA è errata, poiché se gli *elementa* a cui si riducesse l'attività mentale non fossero implementabili su una macchina, l'intelligenza artificiale rimarrebbe impossibile, pur in una visione riduzionista.

Che cosa faccio ?

Un punto di vista consolidato da decenni e recentemente messo in dubbio è quello fornito dall'ipotesi del comportamentismo: un computer è intelligente al pari dell'uomo se agisce (risponde agli stimoli) come agirebbe, stimolato allo stesso modo, un uomo; su questa tesi si basa il Test di Turing e si è basata buona parte della ricerca svolta in IA negli anni seguenti alle prime pubblicazioni di Turing sui rapporti tra macchina, algoritmo e pensiero.

Tale punto di vista forniva un'ottima base di partenza in quanto lasciava alla psicologia e alla neurofisiologia il funzionamento della mente e del cervello e si preoccupava di trovare un particolare algoritmo che fosse in grado di fornire determinate risposte a certi stimoli (da intendere qui in senso ampio e generale, non come stimoli sensoriali e risposte fattuali, ma come stimoli e risposte informazionali), con la certezza che l'algoritmo che si fosse rivelato totalmente equivalente alla mente nella produzione di output sarebbe stato esso stesso una mente e quindi un'intelligenza.

Questo modo di pensare fornisce certo meno limitazioni all'operato dei ricercatori, ma non ne traccia nemmeno le linee guida, lasciando loro, come termine di paragone, l'agire umano.

Inoltre, nel Test di Turing, si può anche aver l'impressione che la macchina stia cercando di ingannare l'uomo e non di mostrare la propria intelligenza e quindi si può pensare che lo scopo del programmatore sia stato quello di istruire la macchina sui trucchi e i sotterfugi che possono essere utili per superare il Test, piuttosto che di programmare un elaboratore intelligente.

Come lo faccio?

La posizione opposta, detta cognitivismo, basa tutto sull'indagine dei processi di pensiero e di conoscenza che viene svolta dalla psicologia, dando quindi importanza al modo tramite cui i risultati vengono ottenuti, anzi, rendendo impossibile prescindere da esso.

Nonostante questa ipotesi fornisca anche molteplici basi teoriche sui processi mentali, la sua adozione nella pratica operativa è poco comune, in quanto decenni di programmi "non antropomorfi" e tecniche di massiccia elaborazione numerica sono difficili da abbandonare, né è

tanto semplice seguire i dettami di questa teoria, che impone al ricercatore molte scelte che sono dispendiose dal punto di vista sia dell'efficienza che della velocità di esecuzione.

L'intelligenza artificiale

Queste posizioni sono le più diffuse e consolidate; ve ne sono altre, per così dire, estreme, che assolutizzano una qualche convinzione ponendola alla base di tutto il sistema. Vi è anche chi considera l'IA solo un mezzo per produrre macchine un po' più evolute, che liberino l'uomo non solo dalle fatiche del lavoro manuale, ma anche dalla *fatica del concetto*, macchine a cui non debba essere riconosciuto lo stato di creatura intelligente, ma solo di eccellente emulatore di capacità umane (solitamente ben definite e circoscritte).

Del resto, l'IA rimane, se guardiamo al contributo che ha portato alle altre scienze e agli sviluppi che ha provocato, una delle aree di ricerca più fertili e proficue, in quanto coniuga in modo originale e nuovo discipline lontane e per molti versi indifferenti le une alle altre.

Ogni tentativo di bilancio è oltre l'obiettivo e la possibilità di queste pagine, in quanto le attività che sono state svolte dagli anni Cinquanta fino al giorno d'oggi sono una tale quantità e varietà da sconsigliare una simile opera.

Infine c'è da considerare che fino ad ora abbiamo sollevato tantissimi problemi, ma ne abbiamo risolti ben pochi, ed è questa caratteristica di apertura, di continua ricerca, di confini e limiti sempre mutevoli che rende l'IA interessante e attuale da 50 anni; forse lo scopo dell'IA non verrà mai raggiunto, forse verrà addirittura accantonato con il tempo, ma sicuramente i molti esperimenti di collaborazione tra psicologia e informatica, neurofisiologia e filosofia, matematica e scienze del pensiero, non verranno meno per lungo tempo.

Da ultimo l'ennesimo problema irrisolto, che ci assillerebbe qualora fossimo sul punto di dichiarare conclusa la ricerca dell'IA e di produrre il primo computer intelligente:

Se costruissimo un elaboratore che, una volta in funzione, dimostrasse intelligenza, pensiero, sensibilità, affettività, autocoscienza, avremmo poi il coraggio e la legittimità di spegnerlo?

GLOSSARIO

- Algoritmo** Insieme di istruzioni o di passi procedurali atti alla soluzione di uno specifico problema
- Analisi** Esame di un concetto a sistema per la divisione in singoli elementi, in modo da stabilire le relazioni funzionali tra questi elementi. Si contrappone alla sintesi.
- Analogico** riferito alla rappresentazione di grandezze per mezzo di segnali variabili con continuità.
- Apparire** In filosofia della conoscenza così si definisce quell'accadimento in cui il dato conoscibile si mostra, si rivela, appare agli organi di senso o all'intelletto.
- Architettura classica dell'elaboratore** Architettura sviluppata da Von Neumann che prevede un processore centrale, un'unità di memoria, dispositivi di I/O.
- Assioma** In un sistema formale, una proposizione che viene convenzionalmente assunta come vera e di cui ci si serve per dedurre altre proposizioni vere.
- Astrazione** Processo attraverso il quale si passa dagli oggetti particolari alla classe universale che li contiene tutti in base a determinate caratteristiche simili, isolate (=astratte) e poste a definizione della classe detta.
- Attrattori caotici (teoria degli)** Teoria delle neuroscienze che ipotizza la natura caotica dell'attività elettrica cerebrale e quindi la possibilità che uno stimolo sensoriale o intellettuale dia origine ad una modificazione dello stato caotico fino a farlo collassare in un attrattore, uno stato specialmente stabile, permanente o temporaneo; a questo processo corrisponderebbe la nostra memoria e il nostro riconoscere.
- Calcolatore (computer)** L'insieme di apparecchiature collegate fra di loro, in grado di effettuare elaborazioni - di tipo logico e aritmetico - di dati, senza che sia richiesto alcun intervento dell'uomo durante le elaborazioni stesse.
- Cervello** Parte dell'encefalo, sede della maggior parte dell'attività nervosa del corpo umano, è diviso in due emisferi e ricoperto dalla corteccia; è formato da cellule particolari dette neuroni. Viene riconosciuto come sede dell'intelligenza, ma non si è ancora riusciti a stabilire un legame tra le varie zone del cervello e le varie capacità intellettuali.
- Cibernetica** Scienza che studia i meccanismi della comunicazione e del controllo negli organismi viventi e nelle macchine.
- Coerenza** Caratteristica di un sistema formale che consiste nell'impossibilità al suo interno di dimostrare una proposizione e la sua negazione; coincide con la validità del principio di non contraddizione.
- Cognitivismo** Dottrina psicologica sulla natura dell'intelligenza e linea teorica di sviluppo e ricerca in IA; afferma che è necessario conoscere l'attività mentale per definire l'intelligenza, non solo il comportamento del soggetto.
- Completezza** Caratteristica di un sistema formale che consiste nella coincidenza tra l'insieme delle proposizioni derivabili al suo interno e l'insieme delle proposizioni vere al suo interno.
- Comportamentismo** Teoria psicologica secondo cui l'unico segno e l'unica natura dell'intelligenza è il comportamento intelligente del soggetto, lasciando in dubbio l'esistenza stessa della mente. E' in fondo una forma di monismo.
- Conoscenza** Insieme dei processi che ci permettono di assimilare e comprendere i dati del mondo esterno; tradizionalmente la conoscenza è distinta in sensoriale e intellettuale, a seconda che riguardi i sensi o l'intelletto.
- Digitale** Contrapposto ad analogico, che prevede l'uso di segnali discreti per rappresentare dati in forma di numeri e di lettere.

- Deduzione** Processo che permette, partendo dall'universale, di arrivare al particolare, procedendo da una affermazione vera ad un'altra affermazione vera. In logica matematica è quel processo che permette di passare dagli assiomi ad un teorema, ottenendo ad ogni passaggio una proposizione vera, seguendo le regole di deduzione del sistema.
- Dualismo** Posizione filosofica che afferma l'esistenza di due *res*, sostanze, quella materiale e quella spirituale; solitamente viene anche negata ogni possibilità di contatto causale tra i due.
- Euristica** Disciplina matematica che tenta di introdurre il trattamento di variabili casuali nello studio di sistemi complessi e difficilmente parametrizzabili.
- Gnoseologia** Parte della filosofia che si occupa di indagare i modi e gli oggetti della conoscenza umana. Veniva un tempo inglobata nella logica e definita *logica maior*.
- Hardware** L'insieme dei circuiti, dei dispositivi, delle unità che compongono un sistema di elaborazione.
- Informatica** Teoria e pratica del trattamento dell'informazione. Prima degli anni Novanta aveva come obiettivo principale lo sviluppo di tecniche per l'elaborazione dell'informazione, mentre nell'ultimo decennio si è sviluppata la tendenza ad approfondire i sistemi di comunicazione dell'informazione.
- Input** Termine inglese usato per indicare tutto ciò che attiene all'ingresso dei dati in un sistema di elaborazione.
- Intelligenza** Capacità di apprendere concetti, formulare giudizi e produrre ragionamenti, indagando le tre *rationes* della realtà (*quid, quomodo, propter quid* : cosa? come? a che scopo?), esplicantesi eminentemente nel comunicare questi concetti, giudizi, ragionamenti.
- Intelligenza adattativi** Particolare tipo di intelligenza, legato all'esperienza, che permette di adattare il proprio comportamento sulla base dei risultati ottenuti operando in un certo modo in una determinata situazione.
- Intelligenza Artificiale** Branca della scienza dei calcolatori che studia i concetti e i metodi di induzione- astrazione connessi al calcolatore e la rappresentazione simbolica delle conoscenze da utilizzare in tale processo. Essa si è sviluppata in risposta a questa domanda: è possibile ottenere che un calcolatore si comporti in modo che, ove si trattasse di esseri umani, sarebbe definito comportamento intelligente. Si divide in due rami, che sono l'IA debole, convinta che per una macchina non sia possibile essere intelligente, ma solo emulare il comportamento intelligente, e l'IA forte che sostiene la possibilità effettiva di un'intelligenza in un elaboratore.
- Intelligenza creativa** In psicologia, è l'abilità di porsi domande mettendo in discussione e vagliando criticamente le conoscenze apprese. Viene anche detta **pensiero divergente**.
- Logica** La logica è lo studio dell'ente di ragione, ossia del pensato in quanto pensato, nelle sue strutture formali e nei suoi caratteri di coerenza, completezza e verità. Storicamente vengono distinte varie *logiche* (simbolica, formale, aristotelica, booleana,...) che comunque rimangono sottoinsiemi o della logica matematica o della logica aristotelica.
- Logica fuzzy** Rivoluzione della disciplina della logica, introdotta da L.A. Zadeh nel 1973, che consiste nel sostituire la relazione di appartenenza con una funzione di appartenenza, che restituisca valori tra 0 e 1, indicando il grado di appartenenza di un elemento ad un insieme.
- Macchina di Turing** Particolare macchina, dotata di un numero finito di stati, in grado di leggere, riconoscere e scrivere simboli su un nastro potenzialmente infinito, su cui essa può muoversi liberamente.
- Memoria** Capacità di ritenere dati, informazioni, conoscenze; può essere definita *statica*, quando ci si riferisca ad un supporto materiale per la memorizzazione di dati digitali (la stessa carta di un libro o l'*hard disk* di un computer), oppure *dinamica*, quando essa ge-

stisca attivamente le relazioni tra i dati e le informazioni, di modo che il richiamarne uno alla memoria comporti l'attivazione di altri "vicini".

Monismo Posizione filosofica che assolutizza una delle due sostanze proposte dal monismo, prestando fede solo alla materia o solo allo spirito come causa e origine della realtà e della nostra conoscenza di essa.

Neurone Cellula che compone il sistema nervoso; riceve tramite dendriti impulsi dai neuroni collegati e scarica a sua volta se la somma è superiore al livello soglia.

Neuroscienze Quelle scienze che si preoccupano di studiare la natura del cervello e il suo legame con la mente e l'intelligenza. Oltre alla neurofisiologia, ne fanno parte anche la psicologia e la fisica (neurodinamica).

Output Indica tutto ciò che attiene all'emissione di dati in un sistema di elaborazione.

Quanto *Pacchetto* di energia discreta, caratterizzato da una dualità onda-particella e principale oggetto della fisica quantistica.

Rete Calcolatori e collegamenti in trasmissione dati che consentono ai calcolatori di comunicare tra loro e di condividere programmi, mezzi *hardware* e basi di dati e di conoscenze.

Rete neuronale Rete in cui i singoli calcolatori (detti nodi della rete) hanno solo un programma per valutare se sia il caso di emettere un impulso in base a quelli ricevuti; esse simulano il comportamento del cervello umano, rafforzando o indebolendo i collegamenti tra i nodi a seconda della risposta del programmatore alla loro elaborazione o alle somiglianze tra gli input.

Ricorsione Ripetizione di una o più istruzioni (una o più volte) di un algoritmo, effettuata di solito tramite la presenza di una istruzione che esegue l'algoritmo in cui essa stessa è inserita.

Semantico Detto di ciò che riguarda il piano dei significati, ovvero il senso di un messaggio o di una proposizione.

Sintattico Detto di ciò che riguarda la struttura dei simboli che compongono un messaggio, una proposizione, senza andare ad indagarne il significato. Il sintagma *una mela gustoso* è sintatticamente errato, il sintagma *una mela deontologica* è corretto sintatticamente, ma sul piano del significato è assurdo.

Sistema Esperto Programma di macchina che esegue un compito professionale specializzato di norma difficile al livello (e talvolta al di sopra del livello) dell'esperto umano. Poiché il loro funzionamento dipende in misura decisiva da grandi insiemi di conoscenze incorporate, i S. E. talvolta sono noti come *knowledge-based system* (sistemi basati sulle conoscenze). Poiché spesso sono usati per assistere l'esperto umano, vanno anche sotto il nome di *intelligent assistant* (assistenti intelligenti).

Sistema formale Un insieme di strutture simboliche in cui esistono un alfabeto, alcune regole di derivazione, alcuni assiomi, una regola di correttezza sintattica.

Software L'insieme dei programmi per il funzionamento di un sistema di elaborazione.

Strozzeria di Von Neumann L'insieme di limitazioni che prevede l'architettura classica dell'elaboratore, soprattutto la serialità delle operazioni.

Talamo Parte interna al cervello che riceve gli stimoli della corteccia sensoriale e quelli del senso interno. Coordina le sensazioni, garantendo la contemporaneità tra loro stesse e con il senso interno.

Teoria del caos Teoria che si occupa di sistemi (il cui comportamento è di solito descritto da equazioni differenziali) dove una piccola variazione delle condizioni iniziali comporta una enorme ed imprevedibile modificazione dello stato finale.

Test di Turing Test che tenta di appurare se un computer possa essere definito intelligente. Si basa sulla convinzione comportamentalistica che se un computer parla con un interlocutore in un modo che in un uomo verrebbe definito intelligente, allora è intelligente.

BIBLIOGRAFIA

TESTI

- Jader Jacobelli (a cura di) *Aspettando robot. Il futuro prossimo dell'Intelligenza Artificiale*, Laterza, 1987 Bari
- Francesco Lerda *Intelligenza umana e intelligenza artificiale - Est modus in rebus*, Rubettino Editore, 2002 Catanzaro
- Angelo G. Sabatini, Francesco Ianneo *Le nuove frontiere della mente*, Newton&Compton, 1996 Roma
- Donald G. Fink *Mente umana e cervelli elettronici. Introduzione all'intelligenza artificiale*, Zanichelli, Bologna 1967
- Roger Penrose *La mente nuova dell'imperatore. La mente, i computer e le leggi della fisica*, Rizzoli, 1992 Milano
- Douglas R. Hofstadter *Gödel, Escher, Bach: un' Eterna Ghirlanda Brillante*, Adelphi, 1990 Milano
- D.R. Hofstadter e D.C. Dennet (a cura di) *L'io della mente*, Adelphi, Milano 1985
- Karl R. Popper e John Eccles *L'io e il suo cervello*, A. Armando, Roma 1981
- Marvin Minsky *La società della mente*, Adelphi, 1989 Milano
- Lettura Magistrale al VI Congresso della Società italiana di Neuroscienze *Neuroni, mente, anima, algoritmo: quattro ontologie* in Luigi Lombardo Valluri *Neraluce*, Le Lettere, Firenze 2001
- J. R. Searle *La riscoperta della mente*, Boringhieri, Torino 1994
- A. Newell, H. A. Simon *Computer Science as Empirical Inquiry: Symbols ad Search*, Communications of the A. M. C. 1976
- Hubert Dreyfus *What Computers Can't Do: A Critic of Artificial Reason*, Harper & Row, New York 1972
- Margaret Boden *Artificial Intelligence and Natural Man*, Basic Books, New York 1977
- IN: Sergio Neri e Giuseppe Velardo (a cura di) *Mente, Simbolizzazione, Educazione infantile*, Servizio Scuola Materna, Messina 1996
 - Guglielmo Tamburini *Intelligenza artificiale, cognitività e modularità della mente*;
 - Marco Somalvico *Intelligenza Artificiale: aspetti e applicazioni*;
 - Giuseppe Trautteur *Simboli e calcolabilità*.
- Gregory Bateson *Mente e natura*, Adelphi, Milano 1984
- Tullio De Mauro *Capire le parole*, Edizioni Laterza, Bari 1994
- Isaac Asimov *Il cervello umano*, Bompiani, Milano 1970
- Evandro Agazzi *La logica simbolica*, La Scuola, Brescia 1964
- Umberto Bottazzini *Storia della Matematica*, UTET, Torino 1990
- Keith Devlin *Il gene della matematica*, Longanesi 2002
- Giovanni Melzi *Le idee matematiche del XX secolo*, Borla, Roma 1983
- Sofia Vanni Rovighi *Elementi di filosofia I*, La Scuola, Brescia 1962

- Sofia Vanni Rovighi *Gnoseologia*, Morcelliana, Brescia 1963
- Marco Buzzoni *Popper. La persona fra natura e cultura*, Edizioni Studium, Roma 1984
- IN: AA VV *Homo loquens. Uomo e linguaggio, pensiero, cervelli e macchine*, Ed. Studio Domenicano, Bologna 1989
 - Gianfranco Basti *Modelli connessionistici della conoscenza e teoria tomistica dell'atto cognitivo*
 - Abelardo Lobato *L'uomo e la macchina*
 - Eduardo Caianiello *Le basi fisiologiche del pensiero*
 - Francesca Rivetti Barbò *Operazioni meccaniche applicate a espressioni di pensiero e loro diversità di principio: collocata fra simboli sensibili e concetti non-estesi e a-temporali*
 - Guglielmo Tamburini *Menti e macchine: alcuni problemi epistemologici*
 - Filippo Fabrocini *Macchine che imparano e macchine che imparano ad imparare*
 - Roberto Busa *Intelligenza naturale e intelligenza artificiale*
- Sergio Galvan *Pensiero e logica* in Rivista di Filosofia Neoscolastica, N. 1 Anno LXXIX gennaio-marzo 1987 pagg 118-128, Vita e Pensiero, Milano 1987

SITI INTERNET

- <http://www.mediamente.rai.it/articoli/20020206b.asp>, *Il futuro del PC si chiama Qubit*
- <http://www.agor.mediacity.it/CervelloComeFunziona.htm> *Il cervello come funziona?*
- <http://www.psyco.com/memoriali/cervello/cervelloemulatore/cervelloemulatore.htm> *A che cosa serve il cervello?*
- <http://www.ipnosicostruttivista.it/ipnosi/mentecorpomente/emulatore.htm> *Il cervello emulatore di realtà stabili e condivise*
- <http://www.dhushara.com/book/paps/consc/brcons1.htm> *Fractal Neurodynamics and Quantum Chaos : Resolving the Mind-Brain Paradox Through Novel Biophysics*
- <http://www.cs.caltech.edu/~westside/quantum-intro.html> *The quantum computer*
- <http://www.forumteorema.it/dreyfus.htm> *Heidegger Husserl e la filosofia della mente*
- http://www.its.it/computer_science/pagine/molecol.htm *Approfondimenti. Computer molecolari*
- <http://spazioinwind.libero.it/cobas/varie/molecole.htm> *La ricerca "progetta" i computer con circuiti grandi come molecole*
- <http://www.happywebonline.it/archivio/03-2002/mul-art.htm> *Se il computer è fatto di luce*
- http://www.neuroscience.com/qua09961/htm/f_body.html *The prospects for a quantum neurobiology*
- http://news.nationalgeographic.com/news/2003/02/0224_030224_DNAcomputer.html *Computer made from DNA and Enzymes*
- http://www.thehumanbrain.org/press_sl/art81.pdf *Il cervello una macchina complessa e misteriosa ancora da decifrare*

INDICE

Introduzione	1
Natura dell'Intelligenza	3
Mente	
L'animale razionale	3
Pensiero e Intelligenza	4
L'intelligenza e l'esperienza	5
La creatività e l'invenzione	6
Il linguaggio e la lingua	7
Macchina	
Il Test di Turing e la macchina intelligente	10
Algoritmi, logica e calcolabilità	11
Programma intelligente?	15
La casualità e l'euristica	15
Hardware e Software	17
Mente	
I neuroni	17
La società della mente	18
La struttura del cervello	19
Teorie della mente	20
Il software mentale	21
Macchina	
L'hardware attuale	23
Le reti neuronali	25
I computer molecolari	25
I computer quantici	26
Software per l'IA	27
Rapporti hardware-software	28
Natura e Processi della Conoscenza	30
Mente	
La storia della conoscenza	30
La sensazione	32
L'astrazione	33

La comprensione	34
Macchina	
I sensi di un computer	35
L'astrazione	36
Memoria e riconnessoni	36
Autocoscienza	38
Conclusioni ?	42
Glossario	47
Bibliografia	51
Indice	53