### 11

### L’algoritmo e la mente[[1]](#footnote-1)\*

Giuseppe Trautteur

Preliminari

Fin da quando la scienza cognitiva si chiamava cibernetica e, in seguito, intelligenza artiﬁciale, insieme all’entusiasmo per l’apertura allo studio ﬁsico e – dalla ﬁne degli anni ‘50 – algoritmico del processo mentale era percepibile un certo disagio e addirittura una certa difﬁdenza. Ma difﬁdenza di che? Non è facile dirlo perché non è chiara la collocazione culturale e disciplinare di questa successione di avventure intellettuali, genericamente, ma non soddisfacentemente, uniﬁcate dall’indagine sul processo mentale, né le opinioni in merito sono unanimi. Un punto fermo è che la cibernetica sia stata, nonostante un periodo iniziale di intensa infatuazione da parte della comunità scientiﬁca, più una promessa non mantenuta che una disciplina reale (recentemente in Italia è stata soppressa come settore scientiﬁco-disciplinare ufﬁciale autonomo). La valutazione della situazione non è certo sempliﬁcata dal permanere del termine cibernetica in collegamento con contesti teorici “morbidi” quali la teoria generale dei sistemi, soprattutto sociali, dalla sua resurrezione nell’ambito di varianti moderne della intelligenza artiﬁciale, quali la cosiddetta AI nouvelle, che ha introdotto i concetti di *situatedness* e *grounding* e, soprattutto, dall’esplosione del termine in contesti ludico-tecnologici, sociologicamente non banali, sotto forme quali cyberpunk, cyberspace ecc. L’intelligenza artiﬁciale, poi, ha acquisita una sufﬁciente robustezza accademica solo quando si è conﬁgurata come il settore “alto” della ingegneria del software, e comunque anche adesso non vive una vita di tutto riposo nell’ambito dei diversi rami della Informatica. In effetti dal clima culturale e tecnologico che dette origine a questa speranza di un nuovo assalto alla mente, omogeneo alle scienze naturali più che a quelle umane – preﬁgurato dalla duplice collocazione di neurologia e, soprattutto, psicologia –, non è emersa l’attesa nuova disciplina scientiﬁca, bensì un nuovo dominio tecnologico: l’informatica. Qui non c’è alcuna insicurezza e difﬁdenza, se non quella del laico di fronte al nuovo e all’insolito, ma la normale, impietosa, equa dinamica di una disciplina scientiﬁca applicativa tra le altre discipline. Qui il criterio del successo è limpido: il programma gira e compie ciò per cui era stato realizzato.

Purtroppo, invece, nello studio del processo mentale, soprattutto nella fase intelligenza artiﬁciale, cioè approssimativamente dalla metà degli anni ‘60 e ﬁno all’inizio dei ‘90, sono state formulate innumerevoli promesse realizzate in maniera per così dire sghemba. Per esempio le promesse relative all’uso del linguaggio naturale in contesti di comune buon senso sono a tutt’oggi irrealizzate; invece negli ambiti della robotica, del controllo, della percezione – soprattutto visiva –, della ricerca e reperimento dati forse si sono ottenuti risultati che non erano stati né previsti né promessi, e nel classico campo dei giochi le promesse sono state mantenute in modo soddisfacente.

Questa insolita impostazione dei programmi di ricerca, quasi si trattasse di scommesse, e gli incerti esiti relativi ben giustiﬁcano una certa difﬁdenza. Si noti che l’incertezza degli esiti non si riferisce tanto al risultato esplicito, quanto alla quasi completa assenza di penetrazione conoscitiva di situazioni pur dominate dall’artefatto. E tuttavia il punto veramente dolente consiste nella mancanza di una strutturazione metodologica di questa ipotetica disciplina scientiﬁca in ﬁeri e nella assenza di sue chiaramente riconoscibili mete conoscitive, al di là della generica dichiarazione d’intenti di voler studiare la mente con metodi ﬁsico-algoritmici. Sorprende anche la relativamente modesta discussione teorica seria in merito, soprattutto se la si confronta con l’animosità della denigrazione popolare, purtroppo spesso acriticamente sottoscritta anche da brillanti cultori di discipline afﬁni. La difﬁdenza attuale, quindi, è rivolta all’imbarazzo che nasce dal dover gestire uno schema metodologico fondato sulla imitazione o simulazione artiﬁciale del processo mentale, e sul suo confronto con questi artefatti, ben diverso dallo schema cosiddetto ipotetico-deduttivo che, tacitamente o con consapevolezza epistemologica, regge l’occidente scientiﬁco dal Seicento in poi.

Situazione ben diversa, quindi, dalla difﬁdenza suscitata, poniamo, dalle pretese di Teoria del tutto, avanzate dai ﬁsici delle stringhe. La diversità consiste, forse, nel fatto che in ﬁsica i risultati, anche se parziali e rivedibili, sono incontrovertibili. Una teoria parziale e forse erronea nel frattempo dà conto in maniera soddisfacente di una parte del reale. Una tale teoria non viene considerata una mera ipotesi o una linea di pensiero sostenuta da una scuola contro qualche altra. Se c’è contrasto tra una parte e un’altra del corpus sperimentale – e c’è sempre – è l’intera comunità che porta il fardello della contraddizione. E quando quella teoria sarà riveduta si saprà perché, e in che cosa, era erronea e si capirà anche perché in un certo ambito era così soddisfacente anche se erronea. Le cose sono andate avanti così per lo meno da Tolomeo ai buchi neri. Insomma, mentre si tende a capire dettagliatamente l’organizzazione dell’Universo e, di fatto, a costruire una metaﬁsica sperimentale, indipen-dentemente dalla seriosità di ﬁlosoﬁ ed epistemologi della ﬁsica “tradizionalisti” (p.e. Mittelstaedt, Snell, Kuhn, Popper) o di personaggi più “scapestrati” (p.e. Cartwright, Feyerabend), non si rimane immersi in una discussione di opinioni, ma in una costruzione di risultati intuitivamente, concettualmente, sociologicamente e applicativamente appaganti. Sembra invece che l’indagine sulla mente alla luce e tenendo conto dei risultati scientiﬁci e ﬁlosoﬁci contemporanei non si lascia problematizzare con la chiarezza a cui tre secoli di successi scientiﬁci ci hanno abituato. Che cosa esattamente ci aspettiamo da un’indagine ﬁsico-algoritmica sulla mente?

La domanda non è retorica. Se infatti in questi anni fosse stato dimostrato il teorema di Pask che discrimina le procedure *m*-calcolabili da quelle *h*-calcolabili o la famosa teoria war (Wiener, Ashby, Rosenblatt) sui processi ﬁsici algoritmici, come, per esempio, è stato fatto con il teorema di Shannon in Teoria dell’informazione o la teoria qed dell’elettro-magnetismo quantistico, ci si sarebbe trovati in una linea di continuità con l’*habitus* della ricerca scientiﬁca, e il sapore intellettuale dei problemi oggi sarebbe stato, probabilmente, molto diverso.

Ma risultati di tal genere non sono venuti e in loro vece abbiamo una lunga serie di artefatti, rapidamente dimenticati, e, in luogo di acclaramenti su che cosa sia e come funzioni la mente, una vasta letteratura speculativa che spesso prende a soggetto precedenti speculazioni senza mai l’avallo di una misura empirica corroborante o falsiﬁcante, ovvero, come accade in ﬁsica teorica, in logica e in matematica, di una dimostrazione probante. Gli artefatti realizzati non sono esperimenti – forse perché non esiste o non è stato ancora identiﬁcato l’ente di natura corrispondente alla parola ‘mente’ – e non esiste un criterio, formalizzato o intuitivo, di che cosa costituisca un risultato scientiﬁco positivo nell’ambito di quella successione di discipline oggi approssimativamente riassunte nella scienza cognitiva.

Il confronto tra il biologico e l’artiﬁciale

Come accennato sopra, lo strumento epistemico nuovo della scienza cognitiva consiste nella imitazione, simulazione, duplicazione, e quindi nel confronto, tra i processi mentali, ma in realtà tra tutti i processi simbolici naturali, e i processi simbolici artiﬁciali. Va notato che tutto ciò è diventato possibile anche perché per la prima volta nella storia dell’umanità sono sorti verso la metà del Novecento algoritmi autonomi, cioè processi simbolici *attivi*, ﬁnora appannaggio dei soli sistemi biologici. Visitiamo quindi la teoria algoritmica della mente analizzando il confronto tra mente biologica e artefatti che ne imitano alcuni aspetti. Considereremo quattro occasioni principali del confronto tra il biologico e l’artiﬁciale.

*Capacità razionali*

La prima occasione è la più coltivata dalla ﬁlosoﬁa della mente e consiste nel confronto tra le capacità razionali della macchina e dell’uomo. In epoca classica e medioevale tale confronto era inconcepibile per la mancanza del termine *a quo*, cioè la macchina, mentre era considerato come un’ovvietà il fatto dell’inferiorità della mente umana rispetto a quella angelica o divina. In epoca moderna – si fa risalire, tradizionalmente, il punto di svolta a Cartesio – tale confronto rimase molto a lungo poco più che un’ “accademia”. Si ricordi che La Mettrie, insieme al celebrato *L’homme machine*, pubblica altre opere assai meno note, tra cui il breve *L’homme plante*, che ha più l’aspetto di un virtuosismo da scommessa che quello di una vera proposta interpretativa. La macchina “anatomicamente moderna”, se è permesso mutuare questa espressione dalla paleoantropologia, fa una prima precisa, ma fugace, apparizione in Babbage e si insedia negli studi ﬁlosoﬁci solo dopo la cosiddetta crisi dei fondamenti e la scoperta dei teoremi di limitazione. Gödel affronta questo confronto in alcuni suoi scritti (cfr. per esempio, la cosiddetta *Gibbs Lecture* [Gödel 1995]. Vedi anche la discussione in merito in [Tamburrini 1997]) nei quali è adombrata la “superiorità” della mente umana sulla macchina. Questa superiorità, ancorché percepita come qualitativa, di fatto è sempre stata discussa nell’ambito dell’estensione, in quanto il confronto è sempre consistito nel valutare il fatto che la mente umana fosse in grado di dimostrare almeno un teorema in più di un sistema formalizzato – la celebre proposizione gödeliana che sta all’origine dei teoremi di limitazione interna dei formalismi. Questa valutazione, ﬁnora apparentemente inconclusiva, ha dato origine a un dibattito che ormai ha raggiunto quasi le dimensioni di una letteratura specializzata. Il dibattito si aprì con l’articolo di J.R. Lucas “Mind, Machines, and Gödel” [Lucas 1961], è proseguito con numerosissime risposte e controrisposte, talvolta bizzarre [Penrose 1990], e continua ancor’oggi sia pure con qualche segno di stanchezza. é caratteristica di questo e di altri dibattiti del campo una certa coloritura emotiva. C’è chi sostiene la supremazia umana e chi, per contro, vuole eliminare ogni esplicita primazia. Volendone parlare, sarebbe comunque prudente usare una formulazione “inespressiva” dei risultati di limitazione la quale, senza detrarre nulla alla importanza conoscitiva dei risultati stessi, né tanto meno alla loro problematicità ﬁlosoﬁca, ne permette una valutazione forse più lucida.

La formulazione “inespressiva” consiste nella esposizione di tali risultati nei termini delle funzioni parziali ricorsive e in particolare dei problemi di decisione (per queste nozioni tecniche si possono consultare [Börger 1989] e [Odifreddi 1989]), piuttosto che in termini di enunciati dell’aritmetica formalizzata ai quali è inevitabile attribuire un’intenzionalità di senso comune: la cosiddetta interpretazione sottintesa dei segni (*token*) che si fonda sulla comprensione consapevole. Naturalmente in tutti gli ambiti scientiﬁci gli enunciati empirici e teorici vengono poi interpretati dall’operatore cognitivo umano: non si capirebbe infatti perché tali enunciati sarebbero ricercati e formulati se non fossero compresi e non portassero nuova conoscenza. Ma la difﬁcoltà che tiene aperta la discussione da sessant’anni su questo speciﬁco confronto uomo – macchina risiede nella riﬂessività dei contenuti espressi dagli enunciati che costituiscono il risultato stesso di limitazione. Non è così negli altri due grandi risultati di “limitazione” della scienza del Novecento: la meccanica quantistica con la scoperta dell’esistenza del quanto di azione, h tagliato, e la relatività ristretta con la scoperta della invarianza della velocità della luce, *c*. In entrambi questi casi è pure rilevante alla loro “scoperta” un cammino metodologico che ha dovuto considerare le modalità operazionali del processo ﬁsico di misurazione e quindi le modalità di formulazione degli enunciati stessi, ma l’interpretazione critica di questi casi, pur contemplando esplicitamente la relazione tra l’osservatore e il fenomeno osservato, non coinvolge il processo – mentale – con cui si giunge agli enunciati pertinenti. (Per la meccanica quantistica questo non è storicamente del tutto corretto. Com’è noto, vi sono interpretazioni della meccanica quantistica – ed è certo degno di attenzione il fatto che il concetto di “interpretazione” della meccanica quantistica è del tutto corrente, mentre nessuno sente il bisogno di “interpretare” la meccanica classica o la biologia molecolare – nelle quali è cruciale la coscienza dell’osservatore.)

È chiaro che questo confronto e il relativo dibattito sono alquanto inconcludenti. La *pars construens* consiste nella credibilità – che ne giustiﬁca almeno la discussione – della ipotesi algoritmica della mente. Cioè che la mente sia un algoritmo. L’oscurità di quest’ultima frase è appena rischiarata dalla sua trasposizione in: “Il processo mentale è un processo algoritmico” perché sussiste il non piccolo problema che, ontologicamente, non si sa dove collocare l’algoritmo (cfr. la sezione “Teorie e algoritmi”). La *pars*, se non *destruens*, almeno deludente è che – come nella ricerca parapsicologica si asserisce o nega con veemenza la mera esistenza di quell’eventuale *quid* non materiale –, così in questo confronto con pari, e forse non dissimilmente motivata, veemenza si asserisce o si nega l’esistenza di un *quid* non-computazionale, lasciando inesplorato il problema di cosa sia quell’ente, la mente, che poi potrebbe, o non, essere dotata dell’ulteriore *quid*. In altre, più icastiche, parole: «*This Gödel is killing me!* » [Hutton 1976].

*Giochi di abilità*

La seconda occasione di confronto è apparentemente più mondana. Si tratta del confronto, materiale questa volta, tra un calcolatore o programma – non discutiamo qui la differenza tra hardware e software – e un giocatore umano nell’ambito di un gioco di società. Non è lontano l’episodio della vittoria del programma per giocare a scacchi *Deep Blue*. Ciò che distingue questo episodio da precedenti successi dell’artiﬁciale (dama, backgammon) è la “serietà” attribuita al gioco degli scacchi. Gli scacchi costituscono un ambito intellettuale ritenuto non inferiore alla matematica o ad altre serie intraprese culturali. Allora la macchina ha superato l’uomo? La notizia non suscitò l’emozione che ci si sarebbe aspettati e, a parte la consueta osservazione a proposito della mancata autonomia della macchina in quanto programmata, si è generalmente ammesso che *Deep Blue* ha sì vinto, ma con metodi diversi e “stupidi”. Non ha mostrato intelligenza. Ha usato la cosiddetta “forza bruta”. L’uomo avrebbe un’intuizione che non può essere imitata dalla macchina. Ora si ammette generalmente che i programmi per giocare a scacchi (o ad altri giochi) non hanno accresciuto la nostra conoscenza del gioco degli scacchi. Per usare termini introdotti in linguistica si può dire che la competenza nostra (e della macchina, se pur si possa assegnare competenza alla macchina) è rimasta quella di prima, mentre la prestazione, della macchina questa volta, ha raggiunto livelli sommi. E in relazione alla prestazione l’argomento della forza bruta non regge. L’evidenza della distinzione tra forza bruta macchinica e intelligenza autentica – o intuito, o sensibilità, o creatività – umana è illusoria perché si basa inavvertitamente su un’altra distinzione. Quella tra processo mentale umano consapevole e inconscio. Osserviamo che – per una curiosa convergenza parallela tra psicanalisi e intelligenza artiﬁciale – ormai le locuzioni ‘processo mentale inconscio’ e ‘pensiero senza soggettività’ (la *no-subject mental life* di Daniel Dennett) non sono più percepite come quegli ossimori che in fondo sono. L’insegnamento, l’apprendimento, il miglioramento, di molte attività mentali, manuali e corporee umane, forse di tutte, è fondato sulla invenzione e sull’apprendimento di gesti, ancora una volta mentali, manuali e corporei, che non sono affatto, o non facilmente, accessibili alla introspezione consapevole né sono immediatamente controllabili consapevolmente. L’intuizione consapevole, e talvolta erronea, del valore di una posizione nel gioco degli scacchi rientra in questo tipo di gesti nei quali il lavorìo inconscio e il ragionamento consapevole – che qui viene inteso come capacità di resoconto verbale – si intrecciano in modi che la psicologia sperimentale sta dipanando con successo crescente, ma che rimangono ancora largamente ignoti. La forza bruta della macchina si rispecchia nell’enorme elaborazione inconscia su cui si fonda l’intuizione consapevole. Sia la macchina sia il sistema nervoso centrale umano attingono a risorse enormi. Come si vedrà nel prossimo confronto, si è sempre considerato la piccolezza delle macchine come un elemento a loro sfavore. Rimproverare loro un eccesso di forza bruta sembra contraddittorio: la macchina è inferiore all’uomo perché è piccola e anche perché ha troppa forza bruta. Naturalmente tutto si gioca sulla connotazione di ‘bruta’. In intelligenza artiﬁciale si parla propriamente di forza bruta quando una situazione di scelta viene esplorata interamente e senza l’aiuto di nessuna euristica. A questo genere di tecnica di decisione ben si attaglia l’aggettivo bruta vista la banalità della procedura e il numero generalmente ben più che astronomico dei casi da esaminare (10120 posizioni da esaminare nel gioco degli scacchi). Di fatto la forza bruta non può mai essere usata. Già il classico algoritmo detto *alpha-beta pruning* è un caso di forza bruta “mitigata”. Il numero di casi rimane elevatissimo, ma l’esistenza di una euristica sotto la forma della funzione di valutazione permette di potare l’albero di scelta in modo da rendere praticabile il numero dei casi da esaminare. Ebbene, anche l’elaborazione inconsapevole umana è certamente bruta in questo senso mitigato e abbiamo saputo duplicarla. L’intuizione consapevole, e talvolta erronea, del valore di una posizione nel gioco degli scacchi (che in gergo ﬁlosoﬁco si potrebbe caratterizzare come il *quale* di quella posizione) e la percezione somatica viscerale che le si accompagna (il marcatore somatico di Damasio) mancano in *Deep Blue*. Tuttavia la lettura dei capitoli 8 e 9 di [Damasio 1995], soprattutto le pp. 270, 277 e 291-300, potrebbe far azzardare addirittura l’ipotesi di una somiglianza più profonda dei meri esiti comportamentali e funzionali. È suggestiva l’assimilazione del marcatore somatico con la funzione di valutazione e della elaborazione che ha luogo nei lobi frontali con il cosiddetto *backtracking* e visitazione dell’albero di scelta dell’algoritmo *alpha-beta*. Naturalmente nulla di tutto ciò è avvertito consapevolmente dalla macchina, ma non confondiamo il problema della consapevolezza con quello della capacità razionale per riﬁutare in maniera preconcetta il riconoscimento di un risultato molto interessante.

*Il calcolatore e il cervello*

La terza occasione di confronto non è propriamente tra l’uomo e la macchina, ma tra il cervello dell’uomo e il calcolatore. Ciò accadeva tra la ﬁne degli anni ‘40 e l’inizio degli anni ‘50 quando la comunanza di stampo cibernetico tra biologico e artiﬁciale era presa realmente sul serio. Shannon, von Neumann e molti altri si soffermarono sul confronto più immediato: quello basato su fattori di merito oggettivi. Von Neumann in particolare [von Neumann 1957] condusse alcune valutazioni confrontando le dimensioni e i tempi di commutazione degli elementi costitutivi dei sistemi rispettivamente biologici e tecnologici. Riassunse questi confronti in fattori di merito che risultarono, quello dimensionale (ed energetico) enormemente a favore del biologico e quello temporale a favore del dispositivo artiﬁciale. L’ordine di grandezza del fattore di merito globale, prodotto dei due, era stimato da von Neumann tra 104 e 105 a favore del cervello. Quarant’anni dopo gli stessi calcoli condotti sui dispositivi moderni portano alla parità (fattore di merito uguale a uno), addirittura considerando come elemento biologico la sinapsi e non il neurone, come praticato da von Neumann, e quindi penalizzando l’artiﬁciale di un ulteriore fattore compreso tra 103 e 104, quanto è il numero medio di sinapsi per neurone. Naturalmente in queste grossolane valutazioni numeriche rimangono assenti i confronti globali tra le due entità: il calcolatore e il sistema nervoso centrale. In queste valutazioni globali, cioè dei sistemi nella loro interezza e non solo delle caratteristiche degli elementi costituenti, entrerebbero almeno tre importanti fattori: la natura discreta o continua sia dei costituenti ﬁsici elementari, sia dell’aggregato che forma il sistema nella sua interezza; la discernibilità stessa, al di là di assunzioni e categorizzazioni intuitive, di cosa costituisca un elemento la cui analisi o suddivisione sia irrilevante – o dannosa – alla comprensione del sistema nella sua globalità, piuttosto che un aggregato di entità da sondare ulteriormente (per esempio nel biologico è elementare il neurone o la sinapsi – per tacere dei microtubuli – e nell’artiﬁciale un circuito integrato, un ﬂip-ﬂop, una porta, il transistor o la giunzione?); l’architettura distribuita del sistema nervoso centrale a fronte dell’architettura cosiddetta seriale, cioè con un solo centro focale di attività, del calcolatore e, soprattutto, se questa differenza sia di natura ﬁsicamente essenziale e quindi non aggirabile, ciò che minerebbe gravemente l’ipotesi – o la speranza – funzionalista. Infatti “una teoria funzionalista non fa riferimento alle proprietà ﬁsiche del particolare sistema studiato, ma solo al *modo con cui opera*.” Questa formulazione (nostra enfasi) viene data a p. 24 del noto libro *Computation and Cognition* di Z.W. Pylyshyn [Pylyshyn 1984], quale deﬁnizione tradizionale di teoria funzionalista, ma viene ampiamente elaborata e precisata dall’autore soprattutto con la introduzione di vincoli che speciﬁchino cosa sia il “modo con cui opera” il sistema. In particolare sono analizzati i livelli di descrizione e le corrispondenze tra stati di sistemi funzionalmente equivalenti. Oltre al livello ﬁsico (o neuroﬁsiologico) devono essere considerati altri due livelli distinti: quello rappresentazionale o semantico e quello dell’elaborazione simbolica. Pylyshyn avverte che Newell [Newell 1982] chiama il primo di questi due livelli “livello della conoscenza” e il secondo “livello simbolico” (con riferimento alla ben nota *Physical Symbol System Hypothesis* [Newell and Simon 1976] e ai suoi sviluppi). Queste restrizioni sulla teoria funzionalista richiedono, come appare evidente, l’uso di altre nozioni (apparentemente considerate primitive nell’ambito della scienza cognitiva) quali: simbolo, informazione, rappresentazione, conoscenza, semantica, e, naturalmente, livello. La parola ‘funzione’ rimane però il cardine ed è interessante osservarne l’uso parallelo in biologia che sarà menzionato poco oltre.

Questo terzo confronto, tra biologia e artefatto, mette in evidenza che il moderno convive con un’emotività mitopoietica responsabile, per esempio, della distinzione disciplinare tra chimica inorganica e chimica organica e che ﬁorisce nel cosiddetto vitalismo e nella ancora non completamente dominata tendenza al ﬁnalismo nello studio dell’evoluzione. (Curioso questo contrappasso in una teoria – quella di Darwin – certamente ostile a ogni spiritualismo, ma la cui ideologia – il cosiddetto darwinismo – è infestato da forze e da sapienze immateriali surrettiziamente contrabbandate nella Natura e fenomenologicamente non dissimili dal ﬁnalismo usato nella Quinta via di S. Tommaso. Messi alle strette, normalmente, i darwinisti seri negano ogni ﬁnalismo e ogni idea di progresso migliorativo, e tuttavia si continua a percepire una come nostalgia.) Così la ﬁducia nel biologico in quanto tale, presente in spiriti peraltro rotti a ogni astuzia dialettica, per esempio in Searle, agisce come una malleveria per l’accettazione della materialità biologica della mente mentre ne riﬁuta un’incorporazione meccano-elettronica. Somenzi in [Somenzi 1967] suggerisce: “Tra l’altro, questi paragoni incrementarono l’uso nelle scienze umane di quella distinzione tra “organo” e “funzione”, che permise ai materialisti dell’Ottocento di aggirare l’ostacolo di una completa identiﬁcazione tra attività mentali e attività della materia cerebrale, e tuttora permette a molti “dualisti” di fugare i timori suscitati dalla crescente autonomia delle macchine, concedendo ad esse il possesso di organi materiali che ne assicurano il funzionamento, ma riservando allo spirito umano la ﬁssazione e il riconoscimento delle funzioni che tali organi assumeranno.” In merito a queste parole va notato in primo luogo la sagacia somenziana nel navigare tra l’atteggiamento materialista, largamente accettato nei circoli scientiﬁci, e il rispetto per chi è sensibile “[a]i timori suscitati dalla crescente autonomia delle macchine”. E in secondo luogo la variegata connotazione della parola ‘funzione’. Nel contesto del funzionalismo, come visto sopra, funzione è afﬁne a comportamento formando quasi una coppia polare con struttura. In ambito biologico è contrapposta a organo ed ha una netta afﬁnità con l’idea di scopo e con il ﬁnalismo. Non si può qui entrare nella discussione degli sforzi cibernetici per naturalizzare il concetto di ﬁnalità.

Questa emotività, ben distinta dal nobile dualismo di spirito e corpo, denso di forti risonanze etiche e religiose, ignora il fatto che la vera barriera è situata tra il materiale, nel quale cervello e macchina sono accomunati – sia pure non banalmente –, e il mentale o spirituale che, pur costituendoci persone, ci sfugge.

E ancora. Cosa dire della reticenza – non si può chiamarla altrimenti – a realizzare concretamente questo confronto cervello – calcolatore tramite il tentativo, almeno, di costruzione di un cervello artiﬁciale? Non è chiaro perché a fronte di progetti quali i grandi acceleratori di particelle, la conquista della Luna, l’esplorazione vicariante del sistema solare con sonde percettivamente molto smaliziate, il sequenziamento del genoma umano, l’anno geoﬁsico internazionale, il decennio del cervello, ecc. i tentativi rivolti alla realizzazione di un cervello macchinico (o delle sue funzioni) siano stati e continuano a essere modesti, quasi pavidi. L’alibi della complessità, nel senso della immensità del compito, era già poco convincente con la tecnologia dei tempi di von Neumann; oggi semplicemente non sta in piedi. Le realizzazioni informatiche per le comunicazioni e per le attività ludiche, ormai ancor più che per l’elaborazione scientiﬁca, ﬁnanziaria e amministrativa, comportano agglomerati di risorse ben superiori a quelle di un cervello biologico. Certamente manca un progetto, ma ancor più si percepisce una resistenza del tipo: suvvia, siamo seri; veramente pensate di fare un vero cervello? Come si sa, il ridicolo è un antidoto della paura.

*Il “non-detto”*

La quarta occasione di confronto è costitutiva della svolta dalla cibernetica alla intelligenza artiﬁciale. Il confronto assurge in questo caso a vera e propria istanza metodologica. C’è qui il convergere di due intuizioni. Da una parte l’intuizione cibernetica, fondata sulla somiglianza tra (l’organizzazione dei) circuiti nervosi e (dei) circuiti logici o elettrici; dall’altra l’intuizione, scaturita dagli studi sui fondamenti della logica e della matematica, che parte, almeno, del processo mentale pensiero sia computazionale. Il dispiegarsi della tecnologia dei calcolatori – che gli storici futuri acclareranno quanto sia indipendente dagli studi sul pensiero umano – porta inevitabilmente ai tentativi di simulazione che daranno origine alla intelligenza artiﬁciale. I tentativi cibernetici di costruire macchine “pensanti” diventano sistematici, senza che però ne venga chiarita la fondatezza epistemologica e metodologica.

Così, per esempio, ponendo mente alle celebri tartarughe di W. Grey Walter risulta chiaro che, nonostante certe considerazioni di sapore addirittura lirico sulle capacità di questi artefatti, nessuno assegnava reale importanza conoscitiva alla loro realizzazione. Il ﬁttone principale della ricerca cibernetica era acriticamente, ma realmente, ispirato a metodologie classiche, e, per dirla apertamente, viveva un fortunato intermezzo tra l’accettazione indiscussa della ipotesi cartesiana, vittoriosamente rivendicata dalla psicologia sperimentale della prima metà del Novecento (si veda il resoconto di questa vittoria ne *La scoperta dell’artiﬁciale* di R. Cordeschi [Cordeschi 1998]), e lo scontro ancora da venire con i veri problemi della mente. L’ambizione cibernetica esplicita è ancora limitata al dominio della percezione, della reattività e delle più modeste forme di apprendimento, ma corona, di conserva con la biologia molecolare sua contemporanea, il processo di rottamazione dello *élan vital*, alla stregua del ﬂogisto e del calorico.

Ora il mancato chiarimento di quei fondamenti epistemologici è in realtà un occultamento. C’è un non-detto originario (locuzione di sapore psicanalitico che va intesa come una lacuna espressiva che non giunge né alla rimozione né alla menzogna, ma che tralascia qualche importante elemento di ciò di cui si sta trattando) che segna la transizione tra cibernetica e intelligenza artiﬁciale.

Anche se la data di nascita ufﬁciale della intelligenza artiﬁciale suole porsi nel 1956, è indubbio che uno dei suoi prodromi più autorevoli fu l’articolo di Turing “Computing Machinery and Intelligence” del 1950 [Turing 1950]. L’altro grande contributo di Turing allo studio – sia pure indiretto – della mente è, naturalmente, l’articolo del 1936 “On computable numbers, …” [Turing 1936]. Quest’ultimo, in sintonia con i risultati gödeliani che in esso vengono formulati in quella maniera “inespressiva” cui si accennava più sopra, può vedersi come una conclusione della visione classica, puramente immateriale, della mente. L’articolo del 1950 apre a una visione nuova, chiaramente algoritmica, della mente. é forte la tentazione di vedere in questi due articoli contrapposti un *avatar* di Giano. In contrasto con questo suggerimento interpretativo, le metodologie poste in essere nei due articoli sono invertite. I metodi dell’articolo del 1936 sono drammaticamente innovativi. Quelli faticosamente e insoddisfacentemente cercati nell’articolo del 1950 cercano, in ultima analisi e senza trovarla, una legittimazione, nei metodi, già accantonati, della psicologia comportamentista.

È nell’articolo del ‘50 che viene introdotto il “test di Turing” fondato sul “gioco dell’imitazione”, cioè su un confronto tra uomo e macchina. Scopo di questo lavoro è la riconduzione del nuovo modo di studiare la mente ai termini, così fecondi e fruttuosi in ﬁsica, della deﬁnizione operativa. Il gioco dell’imitazione, quindi del confronto, è, nelle mani di Turing, uno strumento metodologico per introdurre una deﬁnizione operativa di intelligenza al cui successo – come traspare tra le righe – egli, a differenza dei suoi epigoni che presero molto sul serio il suo “test”, sembra prestare un interesse e una ﬁducia alquanto minori rispetto ai suoi lavori di logica, calcolabilità e anche di morfogenesi. Sarebbe scorretto attribuire a Turing il non-detto originario dell’intelligenza artiﬁciale. Ma si può argomentare che a partire da quel momento prende forma un equivoco i cui diversi elementi rimarranno occultati per decenni nel non-detto, nonostante tentativi di chiariﬁcazione da parte dei più accorti proponenti della dottrina funzionalista.

*Metodologia della macchina*

Il primo elemento di questo non-detto concerne la natura dell’interesse per le macchine. Non si ripeterà mai abbastanza – ma ﬁnora è stato ﬁato sprecato – che questo interesse non è “sportivo”, ma metodologico. Anche se non si può escludere che la pulsione profonda del gesto intellettuale della cibernetica e della intelligenza artiﬁciale possa essere stata un’ancor più rimossa hybris faustiana, né tanto meno si può negare che l’attuale sopravvivenza dell’intelligenza artiﬁciale è possibile solo in quanto sia divenuta la branca “nobile” della ingegneria del software, tuttavia in origine la cibernetica e l’intelligenza artiﬁciale si volevano scienze naturali classiche indirizzate a studiare autonomamente la mente umana, anzi il genus mente, ove se ne fossero scoperte specie distinte.

Con ‘autonomamente’ s’intende: indipendentemente da altre strade già battute, quali la ﬁlosoﬁca, la neurologica, la psicologica, ecc. Che ciò fosse affatto possibile costituiva una scommessa sia fattuale, sia metodologica. Di fatto si preconizzava – e si coltivò – un approccio che includeva nella sua strumentazione teorica almeno la teoria dei sistemi, la nozione di informazione e l’elaborazione simbolica sia discreta (digitale), sia analogica dell’informazione.

*Metodologia della simulazione*

Un secondo elemento consiste nel praticare il metodo imitativo senza aver fatto reale chiarezza sul perché. Si ribadisce quindi che le macchine, e il confronto della mente umana con esse, sono uno strumento metodologico, ma si passa sotto silenzio il perché una imitazione, una duplicazione, una simulazione di qualcosa costituiscano un *explanans* della cosa in questione. Infatti, come acutamente dice Pylyshyn: “... mimare un comportamento è cosa di gran lunga diversa dal darne una spiegazione. Spiegare il comportamento è un compito ben più esigente che non generare questo comportamento.” [Pylyshyn 1984, p.53]. Realizzare un algoritmo che duplica – codici a parte – il comportamento di un ente naturalizza quell’ente nel senso che dimostra l’inutilità di ipotesi vitalistiche o animistiche come accennato sopra. E l’importanza di questa naturalizzazione non può essere sopravvalutata, soprattutto nel clima di possibilismo animista, nascosto sotto le spoglie, per esempio, dell’olismo o dell’intenzionalità, che pervade la cultura contemporanea. Tuttavia il disagio metodologico permane perché la discussione in merito, relativamente modesta rispetto al lavoro duplicativo vero e proprio anche se accesa e agguerrita, si è rivolta principalmente all’analisi del concetto di modello, alle sue varietà, ai vincoli e alle restrizioni alle quali doveva andar soggetto, piuttosto che al perseguimento di una chiariﬁcazione del concetto di spiegazione nell’ambito della scienza cognitiva. Normalmente un modello è uno schema, intermedio tra i dati empirici e la sistemazione teorica, che aiuta l’intuizione. La natura imitativa della metodologia cognitiva ha fatto assegnare al modello, e in particolare al modello computazionale dall’ambigua collocazione ontologica (cfr. la sezione “Teorie e algoritmi”), un valore esplicativo generando incertezza tra il ruolo e la natura della teoria e il ruolo e la natura dei modelli. L’occultamento di questa incertezza costituisce un elemento del non-detto il cui disvelamento sembrerebbe essenziale per la normalizzazione della scienza cognitiva intesa come erede e compendio dei tentativi di studio ﬁsico-algoritmico del processo mentale.

*Criteri di successo*

Un terzo elemento occultato è la non univocità di un criterio di successo o di veriﬁcazione di queste imitazioni, duplicazioni, simulazioni. Si deve ricorrere al senso comune? Certamente il senso comune non basta come si constata, per esempio, anche in un caso semplice come il programma *Deep Blue*, dove l’ovvio successo “mondano” della duplicazione macchinica di una facoltà squisitamente umana non è considerata soddisfacente, anzi suscita discussioni ulteriori proprio sul senso di questa duplicazione. *Serendipity*

Resta un quarto elemento, forse il più occultato di tutti, certo il più importante, che consiste nella speranza che imitando, duplicando, simulando, insorga una qualche nuova intuizione teorica, dalla quale scaturisca un vero, soddisfacente *explanans* del processo mentale, indipendente dal metodo imitativo, duplicativo, simulativo.

Qualche riﬂessione

Quali conseguenze trarre da questi confronti? In merito al primo (le capacità razionali dell’uomo e della macchina), oltre alle osservazioni svolte fatte di seguito alla sua esposizione, va detto che esso costituisce la veste moderna del classico problema dell’anima: la fusione di anima e corpo dell’Aquinate, la dicotomia cartesiana, ecc. ci vengono riproposti nella forma dei dilemmi della calcolabilità effettiva e con una penetrazione analitica nuova che però rimane tutt’ora senza esiti innovativi.

Il secondo (abilità nei giochi) fornisce alcuni elementi incoraggianti nell’ambito di quel non-detto metodologico che consiste nella speranza del sorgere di idee nuove e buone in un ben concimato campo di sperimentazione “imitativo”. Non si può negare che nei cinque campi considerati da Minsky nel suo “Steps Toward Artiﬁcial Intelligence” [Minsky 1961]: tecniche euristiche di ricerca, riconoscimento delle forme (oggi una branca della *Computer Vision*), apprendimento, soluzione di problemi e pianiﬁcazione, induzione, ai quali almeno si deve aggiungere la rappresentazione della “conoscenza”, sono sorte alcune buone idee per la comprensione del processo mentale autentico, quello biologico. Certamente si è empiricamente consolidata la fondamentale omologia tra macchina e organo e si percepisce un’atmosfera di plausibilità nei confronti della ipotesi computazionale. E non va sottovalutata la riapparizione dell’ontologia in vesti ingegneristiche nell’ambito del trattamento e classiﬁcazione della “conoscenza”, che potrebbe riproporre temi di metaﬁsica classica sotto una nuova luce. Questi parziali successi in quel dominio che si potrebbe chiamare lo psicologico non personale, e che più o meno coincide con la fusione non ancora perfezionata dei domini della psicologia cognitivista e della scienza cognitiva, spingono a sostenere che il risultato più rilevante ﬁnora conseguito sia che larga parte del processo mentale cognitivo, quello che più sopra è stato chiamato “pensiero senza soggettività” e che si potrebbe anche chiamare il cognitivo *understanding free*, è, appunto, indipendente dalla soggettività nel senso che può aver luogo in maniera modesta, ma certo non trascurabile, anche altrove che nella mente umana o biologica. E cioè in un algoritmo, per essere precisi. Questo risultato mentre bandisce – alla stregua di risultati analoghi in ﬁsica, chimica, biologia e psicologia sperimentale – ogni *res cogitans* dall’ambito del comportamento razionale, purtroppo non ha generato la benché minima ipotesi sul genuinamente consapevole, come ci si sarebbe potuto aspettare per analogia alle altre discipline menzionate nelle quale l’abbandono di enti più o meno immateriali quali il calorico e il ﬂogisto già menzionati, ma anche l’etere, i miasmi, i ﬂuidi sottili, la vis come principio esplicativo, ecc. è andato di pari passo con una soddisfacente comprensione non solo della fenomenologia rilevante, ma anche del perché quegli enti ﬁttizi in qualche modo svolgevano il loro ruolo esplicativo. Non è così per il processo mentale razionale senza soggetto. Al pari della logica, sua grande e nobile antesignana, esso vive in un mondo di idee apparentemente non pensate da nessuno. E purtroppo ne rimane mal demarcata la frontiera con il processo cognitivo con comprensione consapevole, per quanto banale: per esempio l’elaborazione di testi in linguaggio naturale per ottenere riassunti sensati e traduzioni accettabili, ancora completamente *terra incognita*, oppure la corretta trascrizione del parlato con lessico e parlanti non vincolati.

Il terzo confronto (cervello e calcolatore) è insieme il più incoraggiante e il più deludente. Da una parte i mezzi tecnici generici per una vasta sperimentazione simulativa – qualunque ne sia la validità epistemica – sono in nostro possesso. Dall’altra non conosciamo il progetto, lo schema, l’algoritmo. E sforzi duplicativi anatomicamente fedeli, come quelli di Edelman, sfuggono ancora alla possibilità realizzativa a larga scala.

Dalla quarta occasione di confronto nasce il non-detto. Perché? La domanda rimane altrettanto priva di risposta che quella sulla mancata sperimentazione del cervello macchinico. Constatiamo, e non si può qui far altro che un cenno, che una situazione ben diversa e certo più positiva si riscontra nell’approccio neurobiologico al processo mentale. Ma, perversamente, i cultori di quest’approccio ignorano o disprezzano l’algoritmo ad eccezione dell’oasi, unica vera speranza al momento, costituita dal conﬂuire della neurologia della visione e della *Computer Vision* nello studio del processo visivo.

Teorie e algoritmi

Il fatto è che l’analisi che ha portato sopra al riconoscimento di un non-detto può egualmente applicarsi agli strumenti posti in essere per studiare e trattare computazionalmente la mente. E quindi, in primo luogo, al ruolo epistemico, oltre che strumentale, dei programmi e degli algoritmi. é qui che una delle principali scuole di pensiero in intelligenza artiﬁciale, la *Information Processing Psychology* fondata dai classici lavori di Newell, Shaw e Simon a partire dalla metà degli anni ‘50 [Newell, Shaw, Simon 1960], con i suoi confronti tra protocolli di prestazioni umane e prestazioni di programmi incontra serie difﬁcoltà di natura metodologica. Secondo la ipp non c’è niente da giustiﬁcare nel metodo imitativo perché essa sostiene che l’imitazione – l’algoritmo – è la teoria esplicativa. Peraltro nella ipp il criterio di successo – congruenza tra protocolli umani e dei programmi – è ben esplicito. L’assunzione di base della ipp, che i programmi, o perlomeno certi programmi soggiacenti a determinati vincoli epistemici, sono teorie esplicative è certamente attraente. Come criticare chi propone al tempo stesso come teoria, modello e spiegazione di un ente o processo cognitivo un programma che somigli in maniera impressionante nelle prestazioni, e sotto vincoli metodologicamente signiﬁcativi, all’ente o processo in questione? Acclarare il modo di interconnessione e di interazione delle parti – anche la semplice, convincente individuazione di demarcazioni interne che isolino parti distinte – di un ente o processo non vale come una spiegazione di quell’ente o processo?

Ora se si prende per buona la deﬁnizione aurorale di algoritmo, vale a dire un insieme di istruzioni con una disciplina di esecuzione, allora un algoritmo può essere una spiegazione perché rende palese a chi lo legge come sta succedendo quello che sta succedendo.

Citiamo solo, senza affrontarlo, il problema della complessità: se la complessità di una spiegazione, in questo caso dell’algoritmo, non è parecchio minore di quella dell’ente da spiegare, allora perde la sua vis esplicativa.

Certamente il modo con cui la ipp usa l’algoritmo per scopi epistemici tralascia la considerazione della complessità. Un algoritmo per essere teoria secondo la ipp deve soddisfare a certi vincoli, ma tra questi non compaiono la semplicità o l’intelligibilità. Si tratta di caratteri psicologici e non epistemologici. Ma che cos’è veramente un algoritmo? Il fatto stesso che larga parte della scienza cognitiva, a cui pure appartiene la ipp, ipotizza che il processo mentale sia un algoritmo e non venga solo descritto o teorizzato per mezzo di algoritmi – come invece accade negli approcci fondati sulle reti neuroniche o, più, genericamente, sui sistemi dinamici – fa pensare che la nozione di algoritmo sia assai più vasta di quanto implicato dalla sua deﬁnizione aurorale. é noto che sussiste il problema della individuazione di criteri che permettano di separare un ente ﬁsico generico da un ente algoritmico. Brevemente, il problema epistemologico e ontologico degli algoritmi e dei programmi (perﬁno la distinzione tra algoritmi e programmi è oggetto di discussione) è *vexata quæstio*. Ricordiamo che astrattamente gli algoritmi sono entità individuate solo come elementi di classi di equivalenza per riferimento alla Tesi di Church e dei quali la natura intrinseca e la conoscenza fenomenologica approfondita sono ancora *in ﬁeri*. Il loro statuto, a mezza via tra sistemi formalizzati e strutture innegabilmente intenzionali, rende veramente problematico un loro ruolo di *explanans* per qualcos’altro. In altre parole, non esiste ancora una teoria soddisfacente per i programmi per cui sembra più prudente – soprattutto sotto l’ipotesi che il processo mentale sia un algoritmo – intraprendere la ricerca di una teoria che lo sia al tempo stesso e degli algoritmi e del processo mentale.

Tutto questo campo di attività, a vario titolo connesse all’algoritmo, ha generato una amplissima letteratura che usa e tematizza termini quali simbolo, informazione, rappresentazione, conoscenza, semantica, e, inﬁne, livello, già segnalati più sopra nella discussione sulla nozione di teoria funzionalista. Ebbene l’uso di questi termini fondamentali è quasi sempre equivoco. Con ‘equivoco’ si vuole dire che, a seconda del contesto e dell’opportunità argomentativa, essi vengono usati ora in senso colloquiale, metaforico, allusivo, ora in senso tecnico. Ma in realtà gli usi tecnici sono rari e legati essenzialmente a esigenze applicative. Sembra quasi si tratti di parole ben comprensibili, certamente, ma prive di un signiﬁcato letterale univoco. Il motivo non è certo linguistico. Purtroppo i signiﬁcati letterali di quei termini sono stati a lungo cercati e i relativi concetti avrebbero dovuto essere, nelle speranze degli Ashby e dei MacKay, quelli di una teoria neutrale e intermedia tra il materiale e il mentale, che naturalizzasse intelligibilmente la mente senza ricorrere o postulare un riduzionismo di maniera. Ma i risultati di questa ricerca sono stati molto modesti e quei termini sono rimasti altrettanto oscuri degli enti di natura di cui vorrebbero formulare la teoria. Cordeschi [Cordeschi 1998] svolge una serrata disamina della ricerca di questo “linguaggio neutrale” che ben può chiamarsi il non raggiunto Graal di quella parte della psicologia sperimentale che si salderà nei tardi anni quaranta con la nascente cibernetica.

Queste equivocità e queste oscurità non riconosciute e tanto meno tematizzate generano per la teoria degli algoritmi una situazione non dissimile da quella generata dal non-detto nella intelligenza artiﬁciale e nella scienza cognitiva. In effetti, come accennato sopra, è la nozione stessa di algoritmo che, auspicabilmente, doveva approfondirsi ed evolvere con lo studio algoritmico della mente, in una forse ingenua e nostalgica analogia con il processo di interazione tra ﬁsica e matematica. Di fatto la discussione dell’algoritmo come oggetto naturale – non certo come ente informatico sotto il quale aspetto ﬁorisce con ricchezza esuberante – è rimasta legata, come ricordato sopra, alla discussione sulla Tesi di Church e alle funzioni parziali ricorsive e solo spunti isolati, come il solitario lavoro di B.C. Smith [Smith 1996], ne intraprendono un’analisi che affronta l’ormai vieta, ma non superata, collocazione nel meramente sintattico. Ritornando al disagio iniziale: cosa ci si aspetta da una teoria degli algoritmi? Cosa esattamente ci aspettiamo potrebbe essere un *explanans* della mente? Gli strumenti che usiamo sono inadeguati.

La coscienza

Recentemente c’è stata una transizione agli studi sulla coscienza. Qui, insieme a lavori di tutto rispetto ancorché sempre obliqui rispetto al problema della soggettività e prevalentemente di carattere neurobiologico, si è veriﬁcata una esplosione di riscoperte dell’acqua calda spesso conﬂuenti nella rissa verbale. é forse l’unico dominio scientiﬁco, dove si pubblicano [Baruss 1990] lavori di psicologia o sociologia della conoscenza, seriamente fondati su indagini empiriche di natura statistica, nei quali traspare, velata dallo scopo socio-psicologico dichiarato, la speranza che i questionari, rivolti al pubblico laico a proposito dei termini fondamentali del dominio: coscienza, autocoscienza, consapevolezza, ecc. per condurre l’indagine, possano poi fornire anche reali elementi di conoscenza in merito agli enti denotati da quei termini stessi.

Questi termini godono purtroppo dello stesso regime di equivocità di quelli già segnalati: simbolo, informazione, rappresentazione, conoscenza, semantica, livello. Viene loro dato senso solo da un soggetto consapevole, capace di elaborazione simbolica e che sia in grado di formarsi delle rappresentazioni mentali. Perché si insiste su questa donazione di senso da parte del soggetto? Dopo tutto non accade così in qualunque dominio scientiﬁco, e in qualunque discorso di senso comune? Per esempio in matematica il senso dei numeri e delle loro relazioni non esiste solo perché – a meno di non rivolgersi a soluzioni platonizzanti – un soggetto consapevole ne intrattiene il pensiero? Il punto è che in matematica risultati e fatti relativi ai numeri si accumulano indipendentemente da chi li pensa: a parte la genialità creativa ancora non dominata dall’artefatto, tutto il campo del *theorem proving* indica una chiara appartenenza della matematica – nella misura in cui gli assiomi di Peano al primo ordine ne esprimano un segmento altamente signiﬁcativo – al cognitivo *understanding free*. Viceversa, operazioni simboliche apparentemente semplici a confronto di complesse derivazioni matematiche, come gli esempi riportati più sopra dell’elaborazione di testi in linguaggio naturale per ottenere riassunti sensati e traduzioni accettabili e della corretta trascrizione del parlato corrente, sono intrattabili perché l’esecuzione – non la deﬁnizione o le descrizioni iniziali – di questi compiti richiedono proprio quella comprensione consapevole che dovrebbero imitare. E questo, nonostante tutte le speranze di naturalizzazione e tutte le dichiarazioni di intenti di omettere la considerazione della coscienza nello studio del processo mentale cognitivo, ha fatto migrare – senza che il fatto sia ancora esplicitamente riconosciuto – il trattamento del senso comune, e di tutto ciò che gli è legato, dalla scienza cognitiva agli studi sulla coscienza. Von Weizsäcker, non per celia, deﬁnisce l’informazione come “ciò che viene capito, oppure come ciò che viene prodotto dall’informazione” [von Weizsäcker 1985]. Ritorna vindice la circolarità che cibernetica (anello di retroazione) e logica (autoriferimento gödeliano) avevano fatto erroneamente balenare come una strada nuova verso la spiegazione della mente. Lo studio e la considerazione della circolarità, della riﬂessione, dell’autoriferimento [Cordeschi *et al*. 1999], che hanno segnato il Novecento, sono forse necessari, ma non sono sufﬁcienti. Siamo in attesa del nuovo.

E per concludere sul non-detto: anche senza scomodare Feyerabend, nessuno pretende una purezza metodologica ed epistemologica *ante rem*. La ﬁsica è stata e continua ad essere costruita senza nessuna impalcatura metateorica ed epistemologica realmente tematizzata. Ma era ed è rigurgitante di fatti, di problemi e talvolta anche di spiegazioni. Qui, a fronte del risultato di “sufﬁcienza”, e cioè delle capacità largamente vicarianti dell’algoritmo rispetto al processo mentale cognitivo, abbiamo una sovrabbondanza di enigmi e risistemazioni verbali. Sembra che di fronte alla soggettività non sia concessa gradualità o neutralità. Questo è il dilemma: o si sottoscrive una cosmologia pronta ad accettare senza emozioni una molteplicità di enti materiali e non materiali, variamente strutturati e interagenti tra loro – la Scolastica, la ﬁlosoﬁa fenomenologica e il senso comune – oppure si elimina il soggetto come un epifenomeno. Il primo corno del dilemma (dualismo) comunemente praticato in ogni attività umana, comporta appena si passi dalla prassi alla teoria, tutti i problemi e le contraddizioni, per non dire i meri non sensi, il cui superamento ha generato il pensiero moderno. Il secondo (eliminativismo) nega l’esistenza della consapevolezza soggettiva da cui consegue che tu, ipocrita lettore, in questo momento sei al più una macchina di Turing, priva di esperienza e di vissuto. La scienza bella, fruttuosa, serena, normale, di fronte al problema della soggettività consapevole si trova *au pied du mur*.

Riferimenti bibliograﬁci

[Baruss 1990]

Imants Baruss, *Personal Nature of Notions of Consciousness: A Theoretical and Empirical Examination of the Role of the Personal in the Understanding of Consciousness*, University Press of America 1990;

[Börger 1989]

Egon Börger, *Computabilità, complessità, logica*, Bollati Boringhieri Informatica 1989;

[Cordeschi 1998]

Roberto Cordeschi, *La scoperta dell’artiﬁciale*, Dunod/Masson Italia 1998;

[Cordeschi *et al*. 1999]

Roberto Cordeschi, Guglielmo Tamburrini, Giuseppe Trautteur, “The Notion of Loop in the Study of Consciousness”, in: C. Taddei-Ferretti and C. Musio (Eds.), *Neuronal Bases and Psychological Aspects of Consciousness*, World Scientiﬁc 1999;

[[Damasio 1995]

Antonio Damasio, *L’errore di Cartesio*, Adelphi 1995;

[Hutton 1976]

A. Hutton, “This Gödel is killing me”, Philosophia 6 (1976), 135-144;

[Gödel 1995]

“Some basic theorems on the foundations of mathematics and their implications” in: K. Gödel (S. Feferman *et a*l. Eds.), *Collected Papers III*, Oxford University Press 1995, pp. 304-323;

[Lucas 1961]

J.R. Lucas, “Minds, Machines, and Gödel.”, Philosophy XXXVI (1961), 112-127;

[Minsky 1961]

M. Minsky, “Steps Toward Artiﬁcial Intelligence”, Proc IRE 49 (1961), 1-30;

[von Neumann 1957]

John von Neumann, *Il calcolatore e il cervello*, in V. Somenzi e R. Cordeschi (a cura di), *La ﬁlosoﬁa degli automi*, Bollati Boringhieri , pp. 130-134;

[Newell 1982]

A. Newell, “The knowledge level”, Artiﬁcial Intelligence 18 (1982), 87-127;

[Newell, Shaw, Simon 1960]

Newell, C. Shaw and H.A. Simon (1960): “Report on a general problem-solving program for a computer”, Proc. Int. Conf. on Information Processing, UNESCO, Paris 1960, pp. 256-264;

[Newell and Simon 1976]

A. Newell and H.A. Simon “Computer Science as empirical enquiry: symbols and search”, CACM 19 (1976), 113-126;

[Odifreddi 1989]

Piergiorgio Odifreddi, *Classical Recursion Theory*, Elsevier 1989;

[Penrose 1990]

“Précis of *The emperor’s new mind: Concerning computers, minds and the laws of physics*”, Behav. and Brain Sciences 13 (1990), 643; e “Continuing Commentary”, Behav. and Brain Sciences 16 (1993), 611;

[Pylyshyn 1984]

Z.W. Pylyshyn, *Computation and Cognition*, MIT Bradford 1984;

[Smith 1996]

Brian Cantwell Smith, *On the Origin of Objects*, Oxford Univ. Press 1996;

[Somenzi 1967]

V. Somenzi, “Uomini e macchine”, in *L’uomo e la macchina*, Atti del XXI Congresso Nazionale di ﬁlosoﬁa, Pisa 1967, Edizioni di “Filosoﬁa” 1967;

[Tamburrini 1997]

G. Tamburrini, “Mechanistic theories in Cognitive Science: The import of Turing’s Thesis”, in: M.L. Dalla Chiara (Ed.), *Logic and Scientific Method*, Synthèse Library, Kluwer 1997;

[Turing 1936]

A.M. Turing, “On computable numbers, with an application to the Entscheidungsproblem”, Proc. Lon. Math. Soc. 42 (1936-7), pp. 230-265 e pp. 544-546;

[Turing 1950]

A.M. Turing, “Macchine calcolatrici e intelligenza” in V. Somenzi e R. Cordeschi (a cura di), *La ﬁlosoﬁa degli automi*, Bollati Boringhieri e in D. Dennet e D.R. Hofstadter (a cura di), *L’io della mente*, Adelphi 1986; cfr. anche: D.R. Hofstadter, *Gödel, Escher, Bach:* Adelphi 1984 ristampe, pp. 641-648;

[von Weizsäcker 1985]

C.F. von Weizsäcker, *Aufbau der Physik*, Carl Hanser Verlag 1985, pp. 200-201.

1. \* Capitolo 11 di: Umberto Bottazzini ed Elena Di Bella (a cura di), *Le costruzioni della mente*, Dynamie, McGraw-Hill Italia 2002, pp.157-176+312-313. Questo contributo è stato sviluppato in un intervento al convegno “L’uomo e la macchina: trent’anni dopo” della Società ﬁlosoﬁca Italiana (Bari 1997). Il lavoro si avvale, parzialmente, dei ﬁnanziamenti murst ex-40% 1995 e infm. Esprimo profonda gratitudine a Roberto Cordeschi, Vittorio Somenzi e Guglielmo Tamburrini per le discussioni, i generosi suggerimenti e le attente letture di questo testo. [↑](#footnote-ref-1)