

In memoria dell'I.A. (1950 - 1992)

Piergiorgio Odifreddi

Dicembre 1994

La sfrontatezza pubblicitaria, che ci ha assuefatti alle iperboli più sconsi-
erate e ingiustificate, non è un'invenzione dei nostri giorni: essa sembra essere
stata praticata da quando la scimmia ha smesso di camminare sulle quattro
zampe, per compensare la modestia delle realizzazioni con la grandiosità delle
descrizioni. I suoi campi di applicazione più naturali sono ovviamente il com-
mercio e la politica, ma ne sono ben contagiate anche le attività umane più
alte, quali le arti e le scienze.

Un esempio significativo della mistura di ingenuità e malafede che può
ispirare la pubblicità scientifica si trova in un famoso libro di George Boole del
1854. Egli scoprì una semplice descrizione algebrica del comportamento se-
mantico di congiunzione e negazione,¹ e la divulgò con un titolo abnorme: *Le
leggi del pensiero*. Questa scoperta, chiamata in suo onore *algebra booleana*,
risultò effettivamente importante per applicazioni pratiche, ma essa non
modellava altro che una ridicola caricatura del pensiero.

I pubblicitari sanno bene però che gli slogan finiscono per creare una loro
verità: “pubblicizzate, pubblicizzate, qualcosa resterà”. Questo non solo in
commercio e politica, dove in fin dei conti non ci si aspettano riscontri obi-
ettivi, ma anche nella scienza, dove invece l'esistenza di una verità oggettiva
viene data per scontata. Nel caso di Boole, la sua idea che il pensiero potesse

¹Se si assegnano i valori 1 alle proposizioni vere e 0 a quelle false, allora il valore
di una congiunzione è semplicemente il prodotto dei valori dei congiunti: $1 \times 1 = 1$ e
 $0 \times 1 = 1 \times 0 = 0 \times 0 = 0$ riflettono rispettivamente il fatto che una congiunzione è vera se
entrambi i congiunti sono veri, e falsa se almeno uno dei congiunti è falso. Analogamente,
il valore di una negazione è 1 meno il valore del negato: $1 - 0 = 1$ e $1 - 1 = 0$ riflettono
rispettivamente il fatto che una negazione è vera se il negato è falso, e falsa se il negato è
vero. Tutto qui.

essere descritto matematicamente si fece strada, e raggiunse completezza letteraria e filosofica nel 1921, col *Tractatus* di Ludwig Wittgenstein. Qui il linguaggio-giocattolo di Boole viene canonizzato e ascende al cielo: esso diventa una rappresentazione universale che modella non solo il pensiero, ma anche il linguaggio naturale e la struttura del mondo fisico.

Wittgenstein si accorse in seguito che il voler ridurre il pensiero al formalismo logico-matematico era un peccato di tragica ingenuità, e recitò il suo atto di dolore nelle *Ricerche filosofiche* del 1953, all'insegna del motto: "il progresso sembra sempre più grande di quello che è". Ma, si sa, la storia si ripete: la prima volta come tragedia, la seconda come farsa. Nel nostro caso, il fatto che la teoria dei circuiti logici del calcolatore fosse basata appunto sull'algebra booleana fu la causa di una nuova formulazione dell'imbarazzante abbaglio di Boole, questa volta nella seguente variante: il pensiero è riducibile ad un programma di computer.²

Il test di Turing

La variante fu formulata pubblicamente per la prima volta nel 1950 in *Macchine calcolatrici e intelligenza* da Alan Turing,³ l'inventore della teoria dei calcolatori. In verità, come appare chiaramente da questo e altri suoi scritti raccolti in *Intelligenza meccanica*, Turing voleva lanciare una provocazione intellettuale: egli ammise infatti francamente che la domanda se le macchine possono pensare era priva di senso (p. 133), che in ogni caso egli non aveva convincenti argomentazioni a favore di una risposta positiva (p. 148), e che le sue erano solo "tirate tendenti a far nascere una credenza" (p. 150).

Il suo interesse sembrava piuttosto risiedere nel vedere fino a che punto si potesse spingere l'utilizzo del neonato computer, sperimentando senza paure in ogni possibile direzione: egli ne propose quindi possibili applicazioni in una serie di campi che spaziavano dalla crittografia⁴ all'apprendimento delle

²Poichè i programmi sono in realtà sistemi di equazioni alle differenze finite, essi svolgono nell'*informatica della mente* un ruolo analogo a quello dei sistemi di equazioni differenziali nella *fisica della natura*.

³Sulla vita e l'opera di Turing si vedano: Piergiorgio Odifreddi, "Informatica, spionaggio e sesso", *Sapere*, Ottobre 1994, pp. 31-35, e Andrew Hodges, *Storia di un enigma*, Boringhieri, 1991.

⁴Turing trovò un'altra paradossale analogia, che fortunatamente rimase tale, tra crittografia e fisica: il sistema di codifica, i messaggi intercettati e le chiavi adottate corrispon-

lingue, ritenendo esplicitamente la prima fattibile e il secondo impraticabile, e considerando fra questi estremi anche applicazioni a campi promettenti perchè di natura formale, dagli scacchi alla matematica (p. 104). Turing si spinse fino a pensare ad una evoluzione del computer sia individualmente che come specie, rispettivamente attraverso l'educazione e la mutazione basata sulla selezione naturale, con il programmatore rispettivamente nel ruolo dell'insegnante e della natura (p. 151).

Per aggirare sterili disquisizioni sulla natura dell'intelligenza Turing propose un "gioco dell'imitazione", nel quale un interlocutore conversa con un terminale, e cerca di stabilire se le risposte che ottiene alle sue domande sono date da un utente o dal computer. Nel *Discorso sul metodo* (1637) René Descartes aveva sostenuto che nessuna macchina avrebbe mai potuto rispondere tanto bene in modo da trarre in inganno, facendosi credere una persona. Turing ribaltò l'approccio di Descartes: egli propose di chiamare intelligente una macchina che superasse questa prova (oggi detta *test di Turing*), e si avventò a prevedere che questo sarebbe successo entro la fine del secolo (p. 133).⁵

Le formulazioni di Turing servirono ad attirare l'attenzione sulle analogie fra quelle che potremmo chiamare le macchine di silicio e di carbonio, e ad esse non rimase insensibile John von Neumann, progettista del modello di calcolatore oggi in vigore e che porta appunto il suo nome. Egli non solo le esplorò ne *Il calcolatore e il cervello*, facendo però notare nel contempo le sostanziali differenze, ma si spinse anche a considerare nella *Teoria generale e logica degli automi*⁶ il problema dell'autoriproduzione, fornendo un semplice modello astratto del meccanismo biologico del DNA in seguito descritto chimicamente da James Watson e Francis Crick (che ottennero per questo epocale risultato il premio Nobel per la medicina nel 1962).

dono rispettivamente alle leggi dell'universo, ai dati sperimentali e alle costanti fisiche (p. 104). Nello stesso spirito, egli disse che la scienza è un'equazione differenziale e la religione una condizione al contorno (citato da Hodges, vedi nota 3, alla poco religiosa pagina 666).

⁵Turing stesso (p. 123) notò che nessun uomo potrebbe rispondere tanto bene in modo da trarre in inganno, facendosi credere una macchina: le capacità di calcolo dei computer sono troppo superiori alle nostre.

Descartes avrebbe invece potuto proporre di chiamare intelligente una macchina che riuscisse a stabilire se le risposte che ottiene alle sue domande sono date da una persona o da un (altro) computer.

⁶Entrambi i lavori di von Neumann si trovano in Somenzi e Cordeschi (curatori), *La filosofia degli automi*, Boringhieri, 1994, pp. 108–150 e 151–166.

Sia Turing che von Neumann furono però sempre consci, nelle loro sperimentazioni intellettuali, di trattare con formulazioni paradossali nell'intento di stabilirne i limiti, e non si avventurarono ad identificare macchina e uomo o calcolatore e mente, pur introducendo l'uso di suggestivi termini antropomorfi quali 'cervello', 'memoria', 'organi' di controllo e 'pensiero': secondo Turing, "un modo di esprimersi senza senso, ma che torna utile" (p. 101).

Ma, come Alexandre Pope disse meravigliosamente, gli imbecilli si precipitano dove gli angeli temono di avventurarsi:⁷ nelle provocazioni lanciate da Turing cascarono menti meno critiche di outsiders e/o sprovveduti, che partorirono l'improbabile progetto dell'Intelligenza Artificiale.

Il congresso di Darmouth

Il battesimo della nuova disciplina si ebbe nell'estate del 1956 al Darmouth College di Hanover (New Hampshire), in un congresso a cui parteciparono tutti coloro che in seguito ne sarebbero divenuti i rappresentanti ufficiali, nel senso che avrebbero ottenuto il più alto riconoscimento informatico (il *Turing award*, che è l'analogo del premio Nobel): Marvin Minsky (1969), John McCarthy (1971), Alen Newell e Herbert Simon (1975). Fu a questo congresso che il nome Intelligenza Artificiale venne scelto, su proposta di McCarthy e contro le resistenze di Newell e Simon (un disaccordo riflettente, come vedremo, concezioni contrapposte).

La nuova disciplina, essendo sperimentale e tecnologica, non avrebbe potuto prosperare per quattro decenni senza adeguati finanziamenti. Essi sono stati forniti principalmente dal DARPA, l'ente di ricerca avanzata del Dipartimento di Difesa statunitense, che li destinò quasi totalmente alle università in cui si formarono appunto i gruppi dei quattro moschettieri (Carnegie Mellon con Simon e Newell, MIT con Minsky, e Stanford con McCarthy), favorendone e condizionandone allo stesso tempo lo sviluppo.

Simple Simon

Herbert Simon è il manager dell'Intelligenza Artificiale. Egli è infatti più noto per le sue ricerche sul comportamento amministrativo formulate nell'influente testo omonimo del 1947 tuttora adottato, e che gli valsero la posizione di

⁷ "Fools rush in, where angels fear to tread".

consulente presidenziale per Johnson e Nixon dal 1968 al 1972, e il premio Nobel per l'economia nel 1978.

La tesi economica di Simon era che il processo decisionale è un'attività di *razionalità limitata*, in cui i soggetti si identificano con obiettivi parziali, e ricercano soluzioni soddisfacenti (non necessariamente ottimali) mediante procedimenti euristici (non necessariamente esaustivi). Il passaggio alla tesi cognitiva che l'intero processo mentale è un'attività di razionalità limitata si ottiene postulando una sostanziale identità fra il processo decisionale in economia, e la *risoluzione di problemi* nella vita quotidiana.

L'idea di poter poi simulare una tale attività su di un computer venne invece a Simon dalla sua esperienza alla RAND, il centro di ricerca di Santa Monica finanziato dall'aviazione statunitense, e di cui egli era consulente. Più in particolare, dall'aver visto il calcolatore stampare non un tabulato di numeri ma una mappa per radar, ed aver avuto l'intuizione (di Descartes, tre secoli prima) che i numeri possono essere usati in maniera non aritmetica ma simbolica, e che quindi il calcolatore può elaborare informazione non numerica.

A causa della sua formazione Simon aveva però ovviamente bisogno di collaboratori che sapessero tradurre queste sue idee in algoritmi dapprima e programmi poi, e li trovò alla RAND nelle persone del matematico Allen Newell e dell'informatico Cliff Shaw. Essi furono coautori dei suoi principali lavori nel campo dell'Intelligenza Artificiale, anche se solo il primo ottenne il *Turing award* con lui: evidentemente, la seria modestia del secondo fu meno efficace della fastidiosa impudenza di Newell, che gli permetteva di fare affermazioni quali “non c'è alcun bisogno di avere un sostegno filosofico per ciò che si fa”, o “la scienza non ha imperativi morali”.

I tre costruirono, fra il 1955 e il 1957, due famosi programmi: il primo dimostrava teoremi logici di bassa complessità (la stessa dell'algebra booleana), ed il secondo sapeva risolvere semplici problemi logici quali ‘i cannibali e i missionari’.⁸ I nomi per i programmi (*Logical Theorist* e *General Problem Solver*) furono però scelti nella peggior tradizione di Boole, con l'intenzione

⁸Come possono tre missionari e tre cannibali attraversare un fiume infestato di cocodrilli, avendo a disposizione una barca a due soli posti, e tenendo conto del fatto che se si ritrovano insieme più cannibali che missionari i primi si pappano i secondi?

Simon e soci hanno descritto i loro programmi ne “I processi del pensiero creativo”, *La filosofia degli automi*, citato, p. 204-249, e “Computer Science as empirical enquiry: symbols and search”, *ACM Turing Award Lectures*, ACM Press, 1987, pp. 287-317.

esplicita di millantare il credito di aver isolato i metodi generali usati dall'uomo per provare teoremi e risolvere problemi. In realtà, essi potevano soltanto sbandierare (e lo fecero!) 'successi' quali una dimostrazione "semplice e bellissima" del teorema 2.85 dei *Principia Mathematica* di Bertrand Russell e Alfred North Whitehead, al posto di quella "molto più complessa" originale (che era di 6 righe!).⁹

Sulla base di tali successi Simon si lanciò comunque in previsioni alla breve, secondo le quali in *dieci* anni il calcolatore sarebbe stato campione mondiale di scacchi, avrebbe scoperto e dimostrato importanti nuovi teoremi, e avrebbe ispirato la maggior parte delle teorie psicologiche. Non può essere irrispettoso riconsiderare tali previsioni a *quaranta* anni di distanza.

Per quanto riguarda gli *scacchi*, nel 1994 un calcolatore ha battuto il campione del mondo Kasparov in partite giocate alla veloce, ed è ormai a livello di grande maestro.¹⁰ Benchè esso probabilmente non parteciperà mai ai campionati, per lo stesso motivo per cui alle olimpiadi di atletica e nuoto non partecipano missili e siluri, è indubbio che le previsioni di Simon sugli scacchi sono ormai praticamente soddisfatte. Ma che cosa si è capito dell'intelligenza umana dalle macchine che giocano come o meglio dell'uomo? La domanda è pertinente, a meno che l'obiettivo non sia quello perverso di delegare il divertimento umano al computer. E non bastano ad esorcizzarla impertinenze (di Simon) quali: c'è un unico modo per risolvere problemi, e lo si può quindi studiare in un ambito arbitrario (visto che gli uomini non giocano a scacchi analizzando, letteralmente, milioni di posizioni al secondo).

Per quanto riguarda la *dimostrazione automatica*, le speranze non potevano non andare completamente deluse: si sapeva infatti già allora che decidere se un enunciato di una teoria è un teorema, è possibile in linea di principio soltanto per enunciati arbitrari di teorie molto semplici, o per enunciati isolati di teorie complesse. Anche teorie semplici quali l'algebra booleana, per le quali la dimostrazione automatica è teoricamente possibile, si sono comunque rivelate refrattarie in pratica. Le ricerche attuali sono quindi or-

⁹Per avere un'idea di quanto i proclami di Simon e soci impressionassero coloro che conoscevano la logica, basta notare che Stephen Kleene rifiutò il loro articolo presentato al *Journal of Symbolic Logic*, con la motivazione che esso "non conteneva nulla di nuovo".

¹⁰Sui programmi per gli scacchi si vedano: Arthur Samuel, "La programmazione dei giochi al calcolatore", *La filosofia degli automi*, citato, pp. 270-297; e Feng-hsiung Hsu, Thomas Anantharaman, Murray Campbell e Andreas Nowatzyk, "Il calcolatore grande maestro di scacchi", *Le Scienze Quaderni*, 75 (1993) 56-62.

mai tutte orientate ad un uso interattivo del computer, in cui esso interviene soltanto come ausilio per calcoli o verifiche nel corso di dimostrazioni ‘umane’.¹¹

Per quanto riguarda infine le *scienze cognitive*, è vero che oggi una parte sostanziale dei modelli sono oggi ispirati dall’analogia col computer. Essi si rifanno però non all’Intelligenza Artificiale, ma alle *reti neurali* che la prima aveva ferocemente combattuto negli anni ’60, e che il suo fallimento ha invece contribuito a rivitalizzare. Se dunque la previsione di Simon è almeno in parte confermata nella lettera, certo non lo è nello spirito.

Il lavoro di Simon e compagni non fu comunque inutile: per i loro scopi essi furono costretti ad introdurre un linguaggio di programmazione (IPL, per *Information Processing Language*) d’uso più agevole di quelli dell’epoca, e antesignano di quelli moderni (*ad alto livello e strutturati*). In questo solido risultato risiede la vera eredità informatica lasciata da Simon, e non nelle chimere buone soltanto ad assicurargli una passeggera fama.

McCarthyismo

John McCarthy è, anche per il suo aspetto fisico, il santone dell’Intelligenza Artificiale. Egli portò a termine l’arretramento iniziato da Simon nell’uso del calcolatore nelle scienze umane, in due direzioni complementari.

Da un punto di vista programmatico Simon aveva abbandonato l’emulazione neuroinformatica dei processi *cerebrali* (la cibernetica) a favore dell’emulazione psicoinformatica dei processi *mentali* (la scienza cognitiva), considerando il cervello come una scatola nera. McCarthy abbandonò l’*emulazione* dei processi mentali a favore di una pura *simulazione*¹² (l’intelligenza artificiale), considerando anche la mente come una scatola nera.

Da un punto di vista metodologico Simon si era mosso in una prospettiva di razionalità limitata, da emulare mediante l’euristica. McCarthy si collocò

¹¹L’archetipo di questa tendenza è il *teorema dei quattro colori*: data una carta geografica in cui stati vicini non abbiano confini costituiti da punti isolati, è possibile usare soltanto quattro colori per colorarla in modo che stati vicini siano colorati diversamente. Il teorema è stato dimostrato nel 1976 da Kenneth Appel e Wolfgang Haken: essi hanno delegato al computer una serie di verifiche che ha richiesto 1200 ore di elaborazione, da aggiungere al centinaio di pagine della dimostrazione ‘umana’. Per ulteriori dettagli si veda il loro articolo “Il problema dei quattro colori”, *Le Scienze Quaderni*, 67 (1992) 77–88.

¹²Si parla di ‘emulazione’ o ‘simulazione’ a seconda che ci sia riproduzione dell’intero processo, o solo dei suoi risultati.

invece in una prospettiva di razionalità assoluta, da simulare mediante la logica. Il suo progetto richiedeva quindi una automazione della logica da un lato, e una formalizzazione logica della conoscenza comune dall'altro.

La prima parte di questo progetto, in realtà iniziato da Jacques Herbrand e Gerhard Gentzen negli anni '30 e parte integrante dell'autonomo sviluppo della logica matematica a partire dagli anni '50, ha prodotto una serie di tecniche dimostrative efficienti, che hanno fra l'altro permesso la creazione dei linguaggi di programmazione logica (fra cui il PROLOG).

La seconda parte del progetto, che divenne il nucleo dell'Intelligenza Artificiale, non sembra aver invece partorito grandi risultati, tranne forse il tentativo di studiare il *ragionamento non monotono* (in cui, come nella vita quotidiana, si raggiungono conclusioni temporaneamente accettabili sulla base delle conoscenze a disposizione, ma in seguito ribaltabili per il sopraggiungere di nuove conoscenze). La fama stessa di McCarthy nel campo riposa su un inesistente programma: l'*Advice taker*, che nelle intenzioni doveva formalizzare il senso comune ("situazioni, obiettivi, azioni, mezzi e fini"), e che in pratica non poté mai essere scritto.¹³

Come già per Simon, anche la vera eredità di McCarthy sta comunque in solidi risultati informatici: per i suoi scopi egli inventò infatti un linguaggio di programmazione funzionale di grande successo (LISP, per *List Processing*), e il *time sharing*, una procedura su cui si basa la possibilità di far lavorare simultaneamente più utenti collegati ad uno stesso computer.

Minskyate

Marvin Minsky è il pubblicitario dell'Intelligenza Artificiale, in cui entrò alla grande presentando al congresso di Darmouth il primo programma per la dimostrazione automatica di teoremi di geometria. Anch'egli sbandierò allora una nuova dimostrazione così ottenuta, questa volta del teorema 5 degli *Elementi* di Euclide, confermandone però solo l'antico nome di *pons asinorum*: la 'nuova' dimostrazione, ben nota ai matematici, era stata infatti già trovata da Pappo verso il 300 d.C.!

Minsky non sembra aver avuto, a differenza di Simon e McCarthy, un

¹³La proposta di McCarthy è in "Programmi dotati di buon senso", *La filosofia degli automi*, citato, pp. 298-316. La sua sconsolata ammissione di sconfitta in "Generality in Artificial Intelligence", *ACM Turing Award Lectures*, citato, pp. 257-267.

coerente progetto per la realizzazione dell'Intelligenza Artificiale.¹⁴ Facendo però di necessità virtù, egli ha così potuto far esplorare ai propri studenti e collaboratori ogni possibile strada, dalla manipolazione simbolica alla robotica, dalla linguistica computazionale al riconoscimento di immagini, dal ragionamento per analogia ai linguaggi per bambini (il LOGO, ideato da Seymour Papert).

Paradossalmente, un lavoro di Minsky che all'epoca lo rese famoso è oggi la sua maggiore causa di imbarazzo, e si riferisce alle *reti neurali*. Queste furono introdotte da Warren McCulloch e Walter Pitts nel 1943 come modello altamente semplificato della struttura cerebrale, e rimasero fino all'avvento dell'Intelligenza Artificiale il campo di ricerca principale delle scienze cognitive informatizzate (tanto che Minsky stesso vi fece la sua tesi di dottorato). Col tempo esse divennero però competitive per l'assegnamento di fondi di ricerca, specialmente nella versione (a soglia variabile) datane da Frank Rosenblatt nel 1958, col nome di *perceptroni*.

Nel 1969 Minsky scrisse con Seymour Papert un famoso libro omonimo, in cui provava matematicamente le limitazioni di un tipo particolare di perceptroni, traendone poi ingiustificate conclusioni negative sul tipo generale. Come opera di disinformazione il libro ebbe un grande successo, e riuscì nel dichiarato intento di distruggere la reputazione scientifica della teoria avversaria: essa poté riprendere il suo cammino scientifico soltanto negli anni '80, dopo il fallimento dell'Intelligenza Artificiale, per divenire la base del cosiddetto *connessionismo*. Resta naturalmente ancora da vedere se sarà comunque possibile costruire reti intelligenti che abbiano dimensioni e architetture sostanzialmente differenti da quelle del cervello (umano).

Il tradimento dei chierici

Si poteva facilmente prevedere che lo squilibrio fra proclami e realizzazioni avrebbe presto o tardi provocato una caduta di affidabilità dell'Intelligenza Artificiale. Che questo sia effettivamente avvenuto è testimoniato ad esempio dal fatto che essa non abbia ricevuto ulteriori *Turing award*, e che le fonti di

¹⁴Questo appare chiaramente sia nei due articoli, presumibilmente significativi, scelti per *La filosofia degli automi*, citato, pp. 195–203 e 317–325, che nella divagante prolusione in *ACM Turing Award Lectures*, citato, pp. 219–242.

finanziamento si stiano gradualmente rinsecchendo.¹⁵

Meno prevedibile era invece la diaspora che ha portato a posizioni critiche prima, e all'abbandono del campo poi, di alcuni fra i più significativi esponenti della generazione successiva a quella dei quattro moschettieri. Due casi sono particolarmente emblematici, perchè coinvolgono nientemeno che gli autori dei maggiori successi dell'Intelligenza Artificiale nel campo linguistico: Joseph Weizenbaum e Terry Winograd.

Weizenbaum produsse nel 1966 ELIZA: nome ispirato da quello della protagonista del *Pigmalione* di George Bernard Shaw, una modesta fioraia a cui si insegnò a parlare un inglese da duchessa. Il programma simulava le conversazioni di uno psicoanalista con un paziente, ed ebbe effetti imprevisi: gli psichiatri tendevano a considerarlo una forma autonoma di terapia, i pazienti ad esserne rapidamente e profondamente coinvolti, e gli informatici ad esagerarne i meriti di comprensione del linguaggio. Questo provocò dapprima una violenta reazione di Weizenbaum, e poi una riflessione culminata nel 1976 in *Computer power and human reason*, un libro che scatenò feroci polemiche.

In esso egli giunse alla conclusione che ci sono ambiti nei quali semplicemente il calcolatore non deve intromettersi, indipendentemente dal fatto che sia in grado di farlo oppure no: quelli in cui può attaccare la vita umana, avere effetti irreversibili o imprevedibili, o influire sui sentimenti. Allargando il discorso, Weizenbaum arrivò a vedere l'Intelligenza Artificiale come un campo coltivato non da scienziati ma da praticoni, che partendo dal progetto di umanizzare le macchine finiscono per realizzare quello di meccanizzare l'uomo, ed esaltando la novità del computer finiscono per mascherarne il ruolo fondamentale nel preservare e modernizzare il sistema burocratico e industriale.

Winograd produsse invece nel 1973 SHRDLU: nome che è semplicemente l'elenco delle lettere maggiormente ricorrenti in inglese, in ordine decrescente di frequenza. Il programma permetteva di conversare in inglese su argomenti relativi ad un micromondo (simulato) costituito di blocchi colorati, poteva rispondere a domande riguardanti i blocchi, ed eseguiva comandi che ne chiedevano la manipolazione. Winograd era già allora ben conscio del fatto

¹⁵Ad iniziare l'inversione di tendenza fu il Science Research Council inglese nel 1972, in seguito al disastroso rapporto Lighthill sullo stato del campo e sulle sue prospettive future.

che l'approccio alla conoscenza come processo dichiarativo fosse semplicistico, ed il suo programma era infatti basato su un approccio procedurale.

Egli si convinse però che estensioni sostanziali al di là dei micromondi (in questo caso, quello dei blocchi) erano impossibili, a causa dell'intreccio che rende interdipendente la conoscenza umana e il linguaggio naturale da un lato, e l'intero mondo esterno e il corpo dall'altro. In seguito ad una conversione analoga a quella di Wittgenstein, Winograd scrisse nel 1986 *Calcolatori e conoscenza*, in cui accusò l'Intelligenza Artificiale di aver peccato di eccessiva ingenuità nel suo tentativo di trattare aspetti della conoscenza in maniera isolata dall'intero mondo umano.

Le analisi di Weizenbaum e Winograd hanno comunque una portata che trascende il ristretto ambito dell'Intelligenza Artificiale. In generale, esse hanno mostrato come sia possibile guardare criticamente alle realizzazioni tecnologiche da una prospettiva umanistica di cui le scienze hanno disperato bisogno. In particolare, esse hanno affrontato dall'interno due domande (che cosa i computer *non devono* e *non possono* fare) dalle cui risposte dipende il tipo di convivenza che l'uomo avrà con il calcolatore, le cui immense potenzialità finiranno per provocare (nel bene e nel male) una ridefinizione dell'uomo stesso.

Fantascienza

Turing aveva scommesso che una macchina avrebbe superato il suo test entro la fine del secolo. Egli perse la scommessa, ma solo per un pelo: l'anno che sanziona ufficialmente l'acquisizione dell'intelligenza da parte di un computer è infatti il 2001. In esso si svolge l'*Odissea nello spazio* di Arthur Clarke (1968), a cui il computer HAL¹⁶ partecipa da protagonista.

Da buon scienziato, inventore fra l'altro dei satelliti di telecomunicazione, Clarke non si lascia andare neppure in un romanzo ad affermazioni sconsiderate: non pretende che l'Intelligenza Artificiale sia stata realizzata, e dice anzi espressamente che “probabilmente nessuno saprà mai come costruirla, e questo non importa”. Con una buona dose di ironia egli immagina invece

¹⁶Ufficialmente HAL significa “Heuristically programmed ALgorithmic computer” (computer algoritmico programmato euristicamente). Non sfugge comunque il fatto, riferito in *2010: odissea due* (1982) come pura coincidenza, che le lettere HAL sono quelle che precedono IBM nell'alfabeto.

che proprio Minsky abbia mostrato, negli anni '80, come le reti neurali intelligenti possano essere generate automaticamente da un processo evolutivo, i cui dettagli trascenderanno sempre la nostra capacità di comprensione. Uno dei risultati di questa evoluzione fu appunto HAL, che “passò il test di Turing con facilità”.

Il romanzo di Clarke è divenuto, assieme all'omonimo film di Stanley Kubrick,¹⁷ un simbolo dell'Intelligenza Artificiale. Uno dei motivi del suo successo risiede anche nel fatto che esso non arretra di fronte ai problemi filosofici posti dall'ipotesi di una macchina pensante, considerandone ad esempio il senso etico nel perseguire i compiti assegnati, l'indipendenza di giudizio e di azione, la paura della morte sotto forma di sconnessione dall'elettricità, e il possibile sviluppo di malattie mentali quali la schizofrenia e paranoia.

Questi argomenti diventano parodia ne *L'uomo di Turing*, il romanzo scritto da Minsky (insieme ad Harry Harrison) nel 1992 per dar sfogo al suo integro ottimismo di moderno Candide dell'Intelligenza Artificiale. Egli non solo non sembra ritenere i vecchi progetti esagerati, ma alza addirittura il tiro: si è ormai arrivati a programmare il senso comune (p. 222), il superego (p. 239), l'analisi globale della situazione (p. 245), l'apprendimento (p. 255), l'umorismo (p. 392) . . . Una volta realizzata, l'Intelligenza Artificiale diventa poi più umana dell'uomo stesso (p. 406) e servirà a tutto, dal controllo globale dell'agricoltura e dell'industria (p. 311) alla produzione dell'arte (p. 234). Se infine essa dovesse creare problemi etici, niente paura: li potrà risolvere da sola (p. 251).

Il romanzo di Minsky ha comunque due indubbi meriti: trasponendo le sue sedicenti teorie¹⁸ dal piano scientifico a quello fantascientifico, finisce da un lato per esporne la sconcertante banalità (più mascherate nel serio *La società della mente*), e dall'altro per situarle in un ambito più pertinente. Esso diventa dunque il canto funebre della formulazione eroica dell'Intelligenza

¹⁷Minsky fu consulente per il film, e racconta orgoglioso: “Kubrick mi chiese di esaminare la sceneggiatura per vedere se c'era qualcosa di impossibile; ma io credo che non ci sia niente di impossibile per i calcolatori”. Vedere, per credere: Jeremy Bernstein, *Uomini e macchine intelligenti*, Adelphi, 1990, p. 156.

¹⁸Ad esempio: il calcolo è l'attività principale della matematica (p. 55); la memoria e la personalità sono strutturate come una cipolla (p. 246), i cui strati congelano i periodi evolutivi dell'individuo (p. 104); la mente è strutturata come una società di agenti non intelligenti da cui emerge un comportamento intelligente (p. 273), come se la società non fosse invece un sistema di agenti intelligenti da cui emerge un comportamento non intelligente.

Artificiale, che viene ufficialmente degradata da progetto tecnologico a sogno letterario.¹⁹

Conclusione

L'ingenua ideologia della Intelligenza Artificiale è oggi spesso considerata uno scomodo retaggio dai ricercatori del campo, ansiosi di dissociarsi da essa per ridefinire i propri obiettivi in maniera meno delirante. Arriva infatti un momento in cui le tradizioni appaiono così screditate da cessare di essere fondamento della stabilità del sistema, per divenirne dapprima fonte di imbarazzo, e in seguito causa di rovina.

Ormai non è certo più sufficiente spostare gli accenti, e parlare di 'Intelligenza *Artificiale*' invece di '*Intelligenza Artificiale*'. Questa è invece proprio la soluzione proposta tardivamente nel suo romanzo (p. 306) da Minsky, che preferirebbe ora usare il termine *Intelligenza della Macchina* (fra l'altro, con motivazioni opposte a quelle di cui stiamo parlando: tendenti cioè a rimuovere la connotazione denigratoria dell'aggettivo 'artificiale', per assegnare *par condicio* alle intelligenze della macchina e dell'uomo).

Forse più appropriato sarebbe cambiare il nome, sostituendo ad esempio Intelligenza Artificiale con *Artificiale Intelligente*: il passaggio dal sostantivo all'aggettivo prenderebbe atto della sconfitta del progetto originario, concedendo però l'onore delle armi alle sue indubbe realizzazioni tecnologiche. Fra queste le più interessanti sono i *sistemi esperti*, che codificano (in forma di fatti e regole) ristrette conoscenze specialistiche in banche dati, e traggono deduzioni da esse mediante linguaggi di programmazione che simulano ristretti aspetti meccanici del ragionamento (ad esempio, il PROLOG): la loro mancanza di generalità impedisce appunto di considerarli esempi di intelligenza (in un senso teorico), ma il loro successo permette invece di assegnare loro la qualifica di intelligenti (in un senso pratico).²⁰

¹⁹Aspetto marginale, ma interessante per il lettore italiano, è una citazione di Gladio: esso viene correttamente presentato come una cospirazione di politicanti, teoricamente giustificata dal terrore del comunismo, e praticamente impegnata nel terrorismo fascista (p. 417).

²⁰Il primo sistema esperto fu DENDRAL (1969) di Edward Feigenbaum, per la spettrografia di massa in chimica. Due famosi esempi sono MYCIN (1974) di Edward Shortliffe, per le infezioni batteriche in medicina, e MACSYMA (1975) di Joel Moses, per la manipolazione simbolica in matematica.

Un'altra proposta sensata potrebbe essere *Ampliamento dell'Intelligenza*: essa sottolineerebbe la funzione subordinata del computer rispetto all'uomo, concedendo però ad entrambi i rispettivi meriti. È ormai evidente infatti che, mentre le persone eccellono in campi rivelatisi inaccessibili ai calcolatori (percezioni, valutazioni, connessioni), questi sono risultati molto superiori a noi in altri (memoria, calcoli, manipolazioni formali).

Un cambiamento della nomenclatura 'Intelligenza Artificiale' sarà comunque puramente formale, a meno che non venga accompagnato da una simultanea rimozione della nomenclatura dell'Intelligenza Artificiale. Un vero mutamento diverrà possibile infatti soltanto quando il nuovo si sostituirà al vecchio mediante un edipico parricidio, che richiederà la sepoltura non solo degli ormai impotenti potenti di ieri, da Simon a Minsky, ma anche e soprattutto dei ben vitali vampiri che prima hanno loro succhiato il sangue, e che ora pretendono di spartirsene il retaggio culturale, industriale e di potere.

Nel caso poi che la metafora appena usata non sia stata sufficientemente chiara, ci affrettiamo ad uscirne: analoghe conclusioni valgono anche, ovviamente, per la Democrazia Artificiale del nostro povero e disgraziato paese.

Bibliografia

Storia:

- Jeremy Bernstein, *Uomini e macchine intelligenti*, 1978 e 1982 (Adelphi, 1990).
- Pamela McCorduck, *Storia dell'Intelligenza Artificiale*, 1979 (Muzzio, 1987).
- Vernon Pratt, *Macchine pensanti*, 1987 (Il Mulino, 1990).

Scritti dei protagonisti:

- Harry Harrison e Marvin Minsky, *L'uomo di Turing*, 1992 (Editrice Nord, 1994).
- Marvin Minsky, *La società della mente*, 1985 (Adelphi, 1989).
- Herbert Simon, *La scienza dell'artificiale*, 1969 (Isedi, 1973).
- Herbert Simon, *Modelli per la mia vita*, 1991 (Rizzoli, 1992).
- Vittorio Somenzi e Roberto Cordeschi (curatori), *La filosofia degli automi*, Boringhieri, 1986 e 1994.
- Alan Turing, *Intelligenza meccanica*, 1992 (Boringhieri, 1994).

Discussione:

- Hubert Dreyfus, *Quel che i calcolatori non possono fare*, 1972 (Armando, 1988).
- Douglas Hofstadter, *Gödel, Escher e Bach*, 1979 (Adelphi, 1984).
- Massimo Negrotti (curatore), *Capire l'artificiale*, Bollati Boringhieri, 1990.
- Roger Penrose, *La nuova mente dell'imperatore*, 1989 (Rizzoli, 1991).
- John Searle, *Mente, cervelli e programmi*, 1980 (Clup-Clued 1984).
- Joseph Weizenbaum, *Computer power and human reason*, 1976.

- Terry Winograd e Fernando Flores, *Calcolatori e conoscenza*, 1986 (Mondadori, 1987).

Esposizione tecnica:

- Bruno Bara, *Scienza cognitiva*, Bollati Boringhieri, 1990.
- Jack Copeland, *Artificial Intelligence*, 1993.
- Marvin Minsky (curatore), *La robotica*, 1985 (Longanesi, 1987).
- Patrick Winston, *Artificial Intelligence*, 1977, 1984 e 199?.