



Università degli Studi di Padova

Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata

Corso di Laurea in Filosofia L-5

Vita, computer, fisica quantistica.

La nuova scienza della coscienza
di Federico Faggin nel dibattito
tra uomo e intelligenza artificiale.

Relatore

Prof. Fabio Grigenti

Laureando

Stefano Zanin

Matricola 1150390

Anno Accademico 2022-2023

Indice

Premessa	p. 5
1. Introduzione	p. 7
2. Fisica “classica”, fisica quantistica e natura della vita	p. 10
3. La natura irriducibile della coscienza e i computer	p. 19
4. Il dibattito contemporaneo	p. 27
Note	p. 31
Bibliografia	p. 37

Premessa

La presente tesi riguarda un tema di mio particolare interesse, che ritengo di estrema attualità: il dibattito riguardante la mente, la coscienza e l'intelligenza artificiale. Tra tutti i corsi di studio del ciclo di laurea triennale, questo ambito ha catturato fin da subito la mia attenzione per trasversalità e numero di discipline interessate: mente, linguaggio, scienza ed etica sono solo alcuni degli ambiti filosofici di questo confronto cominciato nella seconda metà del secolo scorso con la nascita dei primi calcolatori elettronici. Le posizioni che hanno composto tale dibattito a loro volta provengono da studiosi di diversa natura: filosofi, neuroscienziati, biologi e matematici i quali, partendo dal loro specifico ambito di studi, hanno proposto tesi e portato contributi rilevanti che hanno fatto progredire la discussione.

Dopo aver visto per caso lo scorso anno un'intervista a Federico Faggin, del quale in passato avevo letto solo qualche articolo di giornale, sono rimasto particolarmente affascinato dalle teorie esposte in quella occasione dove parlava della sua ultima pubblicazione dal titolo "Irriducibile". Da lì ho voluto approfondire il pensiero di questo fisico che, dopo aver passato gran parte della vita a progettare invenzioni che hanno rivoluzionato l'esistenza di tutti noi o quasi, è arrivato a delle conclusioni in apparenza opposte a tutto ciò su cui ha lavorato per anni. In questo testo cercherò di inserire il punto di vista di Faggin nel dibattito sopra citato, di motivare il fatto che lui ritenga la coscienza un fenomeno irriducibile e che essa sia inserita in un contesto di concezione olistica dell'universo alla luce delle scoperte della fisica quantistica. L'intento sarà quindi spiegare perché non solo sia lecita questa visione, ma perché possa essere un contributo rilevante per gli studi attuali e futuri.

Mi permetto, al termine di questa premessa, di scrivere dei ringraziamenti rivolti a chi mi ha accompagnato in questo percorso accademico intrapreso ormai qualche anno fa. Era da tempo che pensavo di frequentare un corso di laurea per puro arricchimento personale, ma il lavoro, il timore di non farcela e la famosa "comfort

zone” mi avevano sempre fatto desistere. Ora che mi accingo a concludere il ciclo di laurea triennale, posso solo confermare la bontà della scelta intrapresa. Se sono arrivato al termine di questo percorso di studi, pur con il tempo che ci è voluto, lo è stato in larga misura per merito delle molte persone che mi hanno incoraggiato, stimolato e supportato. Non farò nomi, se non quello di mia moglie Claudia che mi è sempre stata vicina, ma ringrazio in generale tutti coloro che hanno contribuito a far sì che questo pezzo di vita andasse nella direzione giusta. Un altro ringraziamento lo rivolgo ai miei genitori che fin dal primo giorno di scuola, in prima elementare, mi hanno sempre incoraggiato e insegnato a fare del mio meglio.

1. Introduzione

Fin dall'antichità i primi filosofi greci hanno riflettuto sull'intelletto, sulla mente e sull'anima, contribuendo a formulare il linguaggio e i concetti a riguardo. È solo però nel XVII secolo con René Descartes che il rapporto mente-corpo ha avuto la sua piena formulazione, portando ancora oggi una questione tutt'altro che chiusa. Sempre nello stesso secolo Blaise Pascal, riprendendo l'idea aristotelica di una macchina progettata per svolgere svariati compiti, arrivò a costruire la macchina calcolatrice denominata *Pascalina*. Anche il desiderio di trasformare il ragionamento in calcolo condusse Goffried Leibniz alla realizzazione di una calcolatrice meccanica capace di eseguire le quattro operazioni matematiche. Da secoli, quindi, l'uomo si è interrogato sul tema della mente e di pari passo ha cercato di ricreare l'intelligenza umana per mezzo di strumenti artificiali, ma fu Alan Turing nel 1936 a concepire con un esperimento mentale una "macchina universale" (*la macchina di Turing*), in grado cioè di eseguire qualsiasi algoritmo generale. Da qui l'invenzione e la costruzione nel Regno Unito, rimasta segreta per anni, di *Colossus*, uno dei primi computer elettronici programmabili che riuscì durante la Seconda Guerra Mondiale a decifrare i codici sviluppati dalla cifratrice *Lorenz SZ 40/42*, usata dai nazisti per comunicare segretamente.

Furono sicuramente la spinta di Turing e il suo celeberrimo *test*¹ a segnare dalla metà del secolo scorso la partenza della discussione racchiusa nella domanda: "*Le macchine possono pensare?*". Il lavoro di Turing ha posto quindi le basi per la scienza informatica con ricadute anche in altri ambiti quali la filosofia, le neuroscienze e le scienze cognitive. Inoltre, è sempre grazie a lui se oggi consideriamo un passaggio decisivo la Conferenza di Dartmouth del 1956, negli Stati Uniti, l'evento che segnò ufficialmente l'inizio del campo di ricerca dell'intelligenza artificiale (IA). Per John McCarthy, Marvin L. Minsky e Herbert Simon il problema dell'intelligenza della macchina prendeva avvio dal test di Turing e affermava che l'iniziativa era promossa "*sulla base della congettura che ogni aspetto*

dell'apprendimento, o ogni altro carattere dell'intelligenza, possa essere descritto in maniera sufficientemente precisa, da poter essere simulato da una macchina" (Carli, Grigenti 2019, p. 109).

Su questo tema si sono poi interessati studiosi di vario genere: filosofi, psicologi, neuroscienziati, biologi, i quali hanno dato frutto a un intenso dibattito. Le ricerche sulla mente e la coscienza dell'uomo in questi ultimi decenni hanno avuto uno sviluppo impressionante e la riflessione sulle origini dell'IA ha posto l'attenzione su nuove questioni come il *machine learning*² e sui risvolti etici riguardanti i possibili scenari futuri.

Le tesi portate avanti dai più rilevanti esponenti in questa materia presentano a mio avviso, seppur con significative sfumature, dei punti di convergenza, ma non è raro trovare idee diametralmente opposte l'una all'altra. Se da un lato il superamento del dualismo cartesiano mente-corpo sembra essersi fatto strada nella gran parte degli studiosi, da un altro lato invece, c'è una profonda frattura tra chi ritiene che le macchine prima o poi arriveranno a eguagliare l'uomo in termini di consapevolezza e intenzionalità e chi invece nega categoricamente questa possibilità. La visione riduzionista descrive il pensiero e la coscienza come dei fenomeni "emergenti" di sistemi particolarmente complessi, come lo è il nostro cervello, alla stregua del rapporto che c'è tra software e hardware. Tale posizione si scontra con chi per diverse ragioni considera la coscienza un fenomeno non riproducibile dall'IA.

Proprio in questo dibattito trovo interessante la posizione del fisico e imprenditore Federico Faggin, "padre" del microprocessore in silicio, il quale, dopo aver passato la vita a progettare nuove tecnologie (tra queste anche il touchpad e il touchscreen) ha deciso di studiare, grazie anche alle reti neurali artificiali, come la coscienza potesse sorgere da segnali elettrici o biochimici. La conclusione, arrivata anche a seguito di un evento personale che lo ha segnato nel profondo, si fonda sulle scoperte della fisica quantistica e va ad affermare come la coscienza sia un fenomeno irriducibile e quindi non riproducibile da nessuna macchina artificiale. Più che sul dualismo cartesiano, nei suoi libri egli pone l'accento sulla dicotomia che nota tra la fisica "classica" e la fisica quantistica; in questi termini la sua argomentazione sulla coscienza si snoda sulla contrapposizione che definisco una sorta di "*dualismo fisico*". Lo studio della fisica quantistica e la collaborazione con il

prof. Giacomo Mauro D'Ariano lo hanno portato a formulare una teoria denominata QIP (*Quantum Information-based Panpsychism*).

In questo scritto mi concentrerò principalmente a trattare la posizione di Faggin sulla natura della coscienza e sul rapporto coscienza, libero arbitrio - intelligenza artificiale, o in senso più ampio tra esseri viventi e macchine fisiche "classiche", avventurandomi parzialmente sulle teorie della meccanica quantistica e sulla teoria QIP di Faggin - D'Ariano, vuoi per la preparazione scientifica richiesta a comprendere fino in fondo alcuni assunti, vuoi perché non è questo il focus della presente tesi di laurea. Per tutte le nozioni tecniche a riguardo, riporterò nella sezione delle note le definizioni presenti nel glossario del libro "*Irriducibile*" di Faggin. Per concludere vorrei condividere una celebre frase del premio Nobel per la fisica, Richard Phillips Feynman: "*Credo di poter dire con sicurezza che nessuno... comprende la fisica quantistica*".

2. Fisica "classica", fisica quantistica e natura della vita

Per comprendere fino in fondo la natura e la portata della posizione di Federico Faggin è bene avere in mente alcuni cenni sulla sua vita o sulle sue "vite", come scrive nel suo primo libro autobiografico: *"Sono nato a una nuova vita ogni volta che, osservando il mondo da insospettati punti di vista, la mia mente si è allargata a nuove comprensioni"* (Faggin 2019, p. 11). Nato a Vicenza nel 1941 da mamma casalinga e papà filosofo, divenuto famoso per aver tradotto le *Enneaidi* di Plotino, si diploma come perito radiotecnico nel 1960. Accetta poco dopo un lavoro come tecnico presso la Olivetti e si trasferisce a Milano. Qui entra in contatto per la prima volta con i circuiti elettronici e si occupa della costruzione di un piccolo computer elettronico sperimentale a transistori, ma nel 1961, l'opportunità dei diplomati degli istituti tecnici di poter avere finalmente accesso all'università, gli fa abbandonare il posto di lavoro per iscriversi alla facoltà di fisica di Padova. Al terzo anno conosce Elvia, la ragazza che diventerà poi sua moglie. Dopo la laurea con 110 e lode, si presenta dapprima la possibilità di lavorare in una piccola società che sviluppa circuiti a film sottile e successivamente il posto alla SGS-Fairchild, l'unica società italiana a produrre semiconduttori al silicio. Arriva il matrimonio, ma l'inizio della "seconda vita" è segnato dall'occasione di andare in California grazie a un programma di scambio di ingegneri tra la sede italiana e quella statunitense. Le nuove esperienze lavorative in realtà faranno diventare gli Stati Uniti la nuova residenza definitiva di Federico e la moglie Elvia. Faggin si dedica ai transistori MOS e poi allo sviluppo della prima tecnologia di gate in silicio, denominata SGT. Da lì a poco arriva il passaggio alla neonata Intel, prendendo il comando del progetto che porterà alla nascita del microchip *Intel 4004*, il primo microprocessore al mondo. In seguito, la rottura dei rapporti con Intel è drastica e Faggin subisce una sorta di *damnatio memoriae*; solo la tenacia della moglie Elvia, con una battaglia che la vedrà impegnata per anni, riuscirà nel tempo a far riconoscere i giusti meriti del marito. Nel 1974 comincia la sua "terza vita", entra nell'imprenditoria e fonda la Zilog,

la prima azienda al mondo interamente dedicata ai microprocessori; lo *Z80-CPU* diventa rapidamente campione di vendite ed è tuttora in produzione. Nella seconda metà degli anni Ottanta si interessa alle reti neurali artificiali e fonda la Synaptics, dove sviluppa chip analogici in grado di emulare le reti neurali. Dopo una prima diffidenza da parte degli esperti di intelligenza artificiale, questa tecnologia diventerà l'unica soluzione ai problemi di riconoscimento di pattern complessi. Dalle reti neurali artificiali passa allo sviluppo della tecnologia del touchpad e successivamente del touchscreen, ma è qui che arriva il passaggio alla "quarta vita", dedicata allo studio della coscienza. Questo cambio in parte è causato da una particolare esperienza che Faggin definisce *"il risveglio"*.

In una notte del 1990 durante le vacanze di Natale, mentre aspetta di riaddormentarsi, sente improvvisamente una forte energia mai provata prima: *"Questa energia viva era amore, ma un amore così intenso e così incredibilmente appagante che superava ogni sentimento e nozione che avevo sulla natura dell'amore. Ancora più incredibile era il fatto che la sorgente di questo amore fossi io"* (Faggin 2022, p. 20). Si sente contemporaneamente l'osservatore del mondo e il mondo stesso. Per lui quest'esperienza contiene un senso di verità senza precedenti. L'incontro con quella che definisce "natura spirituale" dà inizio a un percorso di indagine personale sulla natura della *sua* coscienza, l'unica che può sperimentare e conoscere. Da questo momento Faggin comincia a mettere in discussione la teoria che vede l'uomo come una macchina biologica analoga al computer, sostenendo che *"se l'idea corrente che la coscienza emerge dal cervello come sistema informatico classico fosse valida, anche un computer potrebbe essere cosciente, e quindi la coscienza sarebbe copiabile come lo è l'informazione classica di cui sono fatti tutti i programmi. Ma noi sappiamo perfettamente che la nostra esperienza è un tutt'uno privato e in continua evoluzione, impossibile da descrivere completamente con simboli classici (le nostre parole), perché essa va oltre qualsiasi descrizione"* (Faggin 2022, p. 14). Sulla base di questo pensiero, nel 2010 sviluppa il *quadro concettuale CIF'* (in inglese CIP), acronimo di: Coscienza, Informazione e Fisico. Nella sua concezione, la natura della realtà è costituita da due aspetti complementari e irriducibili: lo spazio-C, cioè lo spazio semantico della coscienza, e lo spazio-I, lo spazio informatico. Lo spazio-F, lo spazio fisico, è per lui uno spazio

virtuale sperimentato dalla coscienza che controlla l'organismo vivente. Secondo tale concezione, la coscienza esiste soltanto nello spazio-C.

È nell'evoluzione del suo pensiero che Faggin inizia a guardare al mondo della fisica con occhio diverso, vedendola non più come unica certezza e forma di verità e osservando con interesse la rivelazione di un nuovo mondo dato dalla teoria della relatività e dalla meccanica quantistica. Per questo voglio porre l'attenzione su quello che nel capitolo introduttivo ho definito "dualismo fisico", perché ritengo sia importante ai fini di una migliore comprensione dell'idea di coscienza e della sua *teoria QIP* che tratterò nel prossimo capitolo.

Nel XVII secolo, le scoperte di Galilei unite al successo della meccanica di Newton fecero progredire di molto la fisica. Il metodo scientifico messo a punto da Galilei cambiò il modo di pensare e presto divenne la nuova metodologia per indagare sulla natura. Per Faggin è qui che la razionalità è passata al primo posto, prendendo il posto del pensiero speculativo basato sull'intuizione, una visione basata sul riduzionismo e sulla fede nella logica matematica che ha permesso di studiare matematicamente e con successo molti sistemi complessi riducendoli alla somma delle loro parti. Nel XIX secolo ci fu lo sviluppo della termodinamica, della meccanica statistica e dell'elettromagnetismo e i progressi scientifici portarono Pierre-Simon Laplace alla famosa affermazione: *"Possiamo considerare lo stato presente dell'universo come l'effetto del suo passato e la causa del suo futuro. Un intelletto che a un certo punto conoscesse tutte le forze che mettono in moto la natura (...) potrebbe abbracciare in un'unica formula i movimenti dei più grandi corpi dell'universo e quelli del più piccolo atomo; per un tale intelletto nulla sarebbe incerto e il futuro come il passato sarebbe presente davanti ai suoi occhi"* (Faggin 2022, p. 33). Alla fine del XIX secolo la visione dominante era una posizione di determinismo e riduzionismo che Faggin nel suo libro elenca punto per punto e che qui riassumo:

1. *"Il mondo fisico è tutto ciò che esiste (naturalismo) e la sua esistenza è indipendente dall'osservatore (realismo).*
2. *La realtà fisica è creata dalle interazioni di "particelle" di materia nello spazio e nel tempo (atomismo).(...)*
3. *Lo spazio e il tempo sono indipendenti e assoluti. Sono il palco su cui particelle di massa e di elettricità interagiscono.(...)*

4. *Il comportamento di un sistema complesso può essere completamente descritto come la somma dei comportamenti di tutte le sue parti (riduzionismo).*
5. *Se conosciamo le condizioni iniziali di tutte le particelle di un sistema, possiamo, almeno in linea di principio, predire tutta l'evoluzione passata e futura del sistema (determinismo).*
6. *Viviamo in un universo statico e chiuso in cui l'entropia (disordine) è in costante aumento.*
7. *Si può osservare il mondo senza disturbarlo.*
8. *La matematica può darci una descrizione vera e completa della realtà.*
9. *L'evoluzione di tutte le specie viventi segue il principio darwiniano di variazione casuale e selezione naturale.*
10. *La natura della mente e della materia è dibattuta. (...) Per il materialista, la mente è semplicemente una funzione del cervello" (Faggin 2022, pp. 34-35).*

Sulla base di questi punti, sostiene Faggin, la mentalità materialista era convinta che la fisica classica sarebbe riuscita a spiegare la natura della coscienza usando lo stesso metodo che aveva prodotto fino a quel momento l'enorme quantità di scoperte e relativa conoscenza. Quando ormai, agli inizi del XX secolo, alcuni fisici sostenevano che non ci fosse altro da scoprire in fisica, si fecero strada tre fenomeni inspiegabili: la *radiazione del corpo nero*², l'*effetto fotoelettrico*³ e le *trasformazioni di Lorentz*⁴. Nel primo caso, il comportamento anomalo di questa radiazione portò Planck nel 1900 alla scoperta del *quanto* per risolvere l'enigma, quando i quanti frazionari non erano ammessi. Nel secondo caso Einstein nel 1905 riuscì a spiegare l'effetto fotoelettrico contraddicendo i principi consolidati della fisica classica e questo gli valse il Nobel. Più tardi il quanto di luce fu chiamato *fotone*. Nel terzo caso, sempre nel 1905, Einstein spiegò la misteriosa trasformazione di Lorentz, ipotizzando che la velocità della luce sia la stessa in tutti i sistemi di riferimento inerziali, cioè quei sistemi che si muovono a velocità costante. La conseguenza di questa scoperta fu fondamentale, perché portò a definire che tempo e spazio dipendevano dalla velocità relativa tra osservatore e osservato, quindi non più indipendenti e assoluti come considerati fino ad allora. Questo studio fu chiamato

relatività speciale e nel 1915 Einstein completò la sua teoria della *relatività generale* (RG)⁵. Il colpo di grazia alla fisica classica arrivò tra il 1925 e il 1926 con le postulazioni riguardanti la meccanica quantistica di Erwin Schrödinger, Max Born e Werner Heisenberg. Nel 1927 Heisenberg postulò il principio di indeterminazione, una relazione matematica che mostra l'impossibilità di misurare con precisione due variabili coniugate, come la posizione e il moto di una particella o l'energia e il tempo della stessa. Nel 1931, invece, il matematico Kurt Gödel completò il lavoro che conteneva i due teoremi sull'incompletezza della matematica.

In pratica, nei primi tre decenni del XX secolo vengono scalfite quelle certezze che si erano consolidate nei secoli precedenti: da un lato il determinismo della fisica smontato dalla RG e dal principio di indeterminazione e dall'altro la matematica "scopre" che la verità dei suoi enunciati fonda su un'insieme di assiomi non dimostrabili su cui però c'è accordo. A mio avviso queste scoperte non hanno avuto nella società l'eco che un uomo come Federico Faggin ha colto nel profondo. Certamente con questa esposizione all'interno del libro "Irriducibile" egli non intende sminuire secoli di scoperte importantissime, ma propone una nuova visione e le sue ultime righe del primo capitolo parlano chiaro: *"La conclusione è che qualsiasi teoria matematica della realtà fisica è solo un modello valido entro certi limiti, e quindi non è totalmente affidabile. Il mondo reale surclassa qualunque tentativo di una sua completa descrizione"* (Faggin 2022, p. 43).

Una delle cose a mio avviso più curiose che Faggin scrive per spiegare la natura della realtà quantistica è l'*entanglement*⁶, un fenomeno che consiste nell'esistenza di particelle che hanno proprietà congiunte, per esempio, due elettroni possono avere il loro spin *entangled* in modo che la somma dei due spin sia 0. Se si misura lo spin dell'elettrone A e si trova che è +1/2, simultaneamente lo spin dell'elettrone B sarà -1/2 e viceversa. Questo "dialogo simultaneo" delle particelle porta all'ipotesi che il risultato della misurazione può propagarsi solo alla velocità della luce, mentre la conoscenza da parte del sistema entangled (costituito dalle due particelle) di ciò che si manifesterà è istantanea. Nessuno può spiegare come mai lo spin di una particella sia sempre correlato a quello della particella misurata, a prescindere dalla loro distanza; questa proprietà dei sistemi quantistici è chiamata *non-località*⁷.

Alla luce di queste importanti scoperte, Faggin propone nel corso della trattazione alcuni assunti sulla visione del mondo della fisica quantistica che si possono riassumere così:

1. *"La natura non è continua, ma discreta: c'è un limite alla divisibilità di spazio, tempo e materia. Tutte le proprietà, sia dei campi sia degli stati (particelle), sono discrete. La natura fondamentale della realtà fisica è quantistica.*
2. *Le particelle elementari appaiono nello spazio-tempo solo quando i campi sono eccitati o interagiscono con altri campi. Le particelle non esistono come oggetti, ma unicamente come stati eccitati del campo omonimo.*
3. *Quando si verifica un'interazione tra due sistemi quantistici, gli stati di entrambi cambieranno con la creazione di uno "stato entangled", avente nuove proprietà che non sono la somma degli stati componenti. Per esempio, quando il campo di elettroni interagisce con il campo di protoni, il risultato potrebbe essere un atomo di idrogeno (...).*
4. *L'indeterminismo è una proprietà irriducibile della natura. (...) Non possiamo (nemmeno) predire dove si troverà una particella, neanche in linea di principio. Possiamo solo predire le probabilità di trovarla in una certa regione dello spazio.*
5. *Un apparato di misura che interagisce con un sistema quantistico è un sistema quantistico-classico che amplifica il risultato di un'interazione quantistica fino a produrre un segnale classico che può essere condiviso. I molti stati che sono possibili nel mondo quantistico "collasseranno" nel nostro mondo fisico in un'unica realtà. E lo stato specifico che si manifesterà non può essere predetto prima della misurazione.*
6. *Le particelle non sono soggette a "forze" così come ci dice la fisica classica. Le particelle interagiscono scambiandosi altre particelle. In questo modo si comportano come se ci fosse una forza che agisce tra di loro.*
7. *La realtà fisica è descritta da un insieme di campi soggetti alle leggi della teoria quantistica dei campi (TQC) e della relatività generale (RG). (...)*

8. *L'unione della TQC e della RG in un'unica teoria coerente promette di ridurre la realtà fisica a un singolo campo quantistico unificato. Questa teoria però non esiste ancora.*
9. *La non-località rivela l'esistenza di una totalità che non può essere divisa in parti separabili. (...) Un sistema quantistico composto ha proprietà incompatibili con ogni proprietà delle sue parti. Questo tipo di olistico non esiste nei sistemi classici, dove le proprietà del sistema sono compatibili con le proprietà delle sue parti" (Faggin 2022, p. 49-50).*

L'effetto di questi assunti è a dir poco enorme e secondo Faggin fanno sparire il determinismo e il riduzionismo, anche in linea di principio; l'interazione tra l'"osservatore" e la realtà crea un legame che va a scardinare il realismo, dove la realtà esiste sia che la osserviamo o meno. Essa quindi, è creata dalle interazioni che avvengono tra le sue parti.

Ho voluto esporre questa sorta di passaggio dalla fisica "classica" a quella quantistica, perché lo trovo indispensabile per capire le idee di Faggin sulla concezione olistica di Universo e sulla natura quantistica della coscienza. Nel superamento sia del dualismo cartesiano che del materialismo monistico, egli espone una tesi in cui mente e materia sono due aspetti irriducibili della stessa realtà, due facce della stessa medaglia. Un concetto che si rifà al panspichismo, dal quale egli prende spunto per la sua *teoria QIP*, volendolo però argomentare con una base scientifica e quindi con possibilità di falsificazione. A suo dire, è una sorta di monismo che ricorda la particella-onda, ossia i due aspetti complementari della fisica quantistica. Per Faggin l'universo deve contenere le basi sia della mente che della materia, al tutto della fisica si deve quindi aggiungere l'interiorità della realtà che chiamiamo mente, includendo cioè la coscienza, il libero arbitrio e la vita.

È proprio parlando della vita che Faggin entra nella questione che riguarda il confronto con i computer, l'intelligenza artificiale e i robot. Essendo tutti gli esseri viventi, dall'organismo monocellulare fino a interi ecosistemi, dei sistemi aperti e dinamici che scambiano materia, energia e informazione con l'ambiente, la domanda che si pone è se saremo in grado con la nostra abilità scientifica di costruire un computer capace di fare lo stesso. Pur ipotizzando che fra qualche decennio o secolo possa accadere che dei robot riescano a collegare il loro "genoma artificiale"

a una stampante 3D e da questa esca un robot figlio, identico a loro, la risposta è no. No, perché i robot continuerebbero comunque ad essere costruiti dall'esterno, poiché sono dei sistemi chiusi, dove il tutto è la somma delle loro parti, mentre negli esseri viventi la vita avviene "da dentro a fuori". La vita è fatta di cellule vive e non di parti inerti; solo con "componenti quantistici", sostiene Faggin, è possibile costruire organismi vivi. I computer e i robot non possono quindi essere né coscienti né avere libero arbitrio.

Altra caratteristica importante degli esseri viventi per Faggin è l'*informazione viva*⁸, un termine da lui coniato. In pratica, in una cellula vivente non esiste la separazione tra hardware, software e alimentazione elettrica che c'è nei computer; una particolare molecola all'interno della cellula invece può essere materiale con cui è costruita (hardware), informazione (software) o energia (alimentazione). Materia, informazione, energia costituiscono l'*informazione viva*, che differisce dall'informazione di Shannon, tipicamente espressa in *bit*. Nel computer "*...il tutto è semplicemente la somma delle sue parti. Infatti il comportamento del computer è interamente determinato da quello delle sue parti, e quindi la causalità procede solo dalle parti al tutto (feedforward)*". In un organismo vivente, invece, esiste anche un ente cosciente che è un tutto indipendente dalle parti, che può quindi influenzarne il comportamento coerente attraverso meccanismi di feedback (dal tutto alle parti), di cui per il momento non si sa quasi nulla" (Faggin 2022, p. 101). Per Faggin quindi l'informazione viva è l'aspetto degli organismi viventi che può essere spiegato solo con l'informazione quantistica, poiché la materia, l'energia e l'informazione di una particella elementare non sono separabili. Le cellule elaborano l'informazione combinando principi computazionali analogici, digitali, classici e quantistici: per capire la vita si dovrebbe dunque studiare l'attività ai tre livelli dell'informazione viva (strutturale, informatico, energetico), aspetti non separabili come nei computer. La conclusione a cui arriva Faggin è che non si possono studiare gli organismi viventi come se fossero solamente oggetti fisici classici, perché la coscienza è fondamentale e va accettata per comprendere la vita. "*La vita è una espressione della coscienza e non della materia inanimata. L'informazione viva e il significato rappresentano rispettivamente la faccia simbolica e semantica di una realtà dinamica indivisibile*" (Faggin 2022, p. 103).

La vita è quindi un sistema olistico fatto di connessioni e non può essere spiegata da un sistema riduzionista, ogni organismo è in simbiosi con l'ambiente e interagisce continuamente con gli altri esseri viventi che lo compongono; proprio per questa interdipendenza un organismo non dovrebbe essere studiato da solo, poiché verrebbero perse delle informazioni importanti. Per capire il comportamento intenzionale e intelligente di ogni organismo vivente, dal più piccolo battere arrivando all'uomo, secondo Faggin, dobbiamo presumere che ciascuno di essi sia "collegato" con un ente cosciente dotato di libero arbitrio. In questa relazione la coscienza è la capacità di conoscersi, di scoprirsi, mentre il libero arbitrio è la capacità di scegliere come agire nel mondo. Un libero arbitrio che può avere sia un aspetto oggettivo che porta l'essere alla creazione di determinati comportamenti stereotipati, ma anche un lato soggettivo che unito alla coscienza consentirà all'organismo vivente di cambiare comportamento in maniera imprevedibile, cosa alquanto impossibile per i robot. Per Faggin la vita è sia quantistica che classica, non c'è demarcazione netta tra materia animata e inanimata, ma la differenza più evidente tra un computer e una cellula vivente è che il primo è fatto di materia classica *permanente*, mentre la seconda è fatta di materia classica *dinamica*. Gli atomi presenti in un chip sono praticamente gli stessi fin da quando viene creato, nella cellula invece gli atomi fluiscono e si trasformano, quelli che escono non sono più gli stessi che vi sono entrati proprio per permettere di adempiere le funzioni vitali. Le considerazioni finali di Faggin sul comportamento di un organismo unicellulare, come il paramecio, rapportate al comportamento di robot dotato di microchip per cervello, lo portano alla seguente conclusione: *"In sintesi, le proprietà cruciali che differenziano gli organismi viventi dai robot dotati di intelligenza artificiale derivano dalla coscienza e dal libero arbitrio, che possono comunicare con la parte quantistica della cellula, determinandone il comportamento complessivo mediante l'informazione viva"* (Faggin 2022, p. 127).

3. La natura irriducibile della coscienza e i computer

Il tema della vita, dell'universo e della coscienza in Faggin si intrecciano per formulare la sua *teoria QIP*, sviluppata con il fisico Giacomo Mauro D'Ariano. Ed è proprio la coscienza il fulcro della trattazione del suo libro "Irriducibile", come parte dell'universo insieme alla materia, capace di dare all'autore quella certezza tale da non vedere la possibilità che computer e intelligenza artificiale possano arrivare ad avere consapevolezza e libero arbitrio. Nelle pagine del libro troviamo la domanda *"a cosa serve la coscienza se un robot può guidare un'auto meglio di noi?"*. Le considerazioni che matura Faggin in risposta alla domanda mi vedono in completo accordo. Nella prospettiva del quesito i nostri sentimenti, le emozioni, i sogni, l'arte e l'amore per la vita diventerebbero un "optional" o ancor peggio illusioni. *"Siete disposti a buttare via tutto ciò che dà senso alla vostra vita per un'idea basata sul desiderio di persone che credono che, quando la complessità diventerà abbastanza grande, la coscienza ne scaturirà "in qualche modo"?. La Complessità è tutt'al più una condizione necessaria, ma non sufficiente, per creare un organismo cosciente. (...) È la coscienza che ci permette di conoscere il significato dei simboli che i nostri sensi ci rivelano e che informano le nostre libere scelte. È la coscienza che ci dà quella libertà di scelta che molti scienziati e filosofi ritengono illusoria perché esclusa dagli assiomi della fisica classica che hanno scelto di accettare"* (Faggin 2022, p. 56).

Il primo approccio scientifico di Faggin con la coscienza avvenne nel 1987, mentre lavorava alle reti neurali artificiali alla Synaptics. All'epoca, in linea con il parere chiesto a un professore di neuroscienze, egli sposava l'assunto materialista secondo cui tutto ciò che esiste è frutto di interazioni di atomi e molecole, di conseguenza cominciò a considerare l'idea di come poter realizzare un computer consapevole. Subito si presentò un ostacolo evidente: la totale mancanza di comprensione della natura delle sensazioni e dei sentimenti, ossia dei *qualia*¹. Per Faggin, la nostra coscienza è lo spazio semantico interiore dove i segnali provenienti dal mondo fisico,

sia esterno che interno al corpo ed elaborati tramite il cervello, vanno a trasformarsi in sensazioni, emozioni o sentimenti; nei computer invece questa trasformazione non è di tipo qualitativo (i qualia), ma puramente riconoscitivo. Egli porta l'esempio di una rosa che viene riconosciuta dal suo profumo: le sue particolari molecole emesse vengono "captate" da alcune molecole recettoriali nelle membrane delle cellule olfattive presenti nel naso. Quando ciò accade, tali cellule producono segnali elettrici diretti alle reti neurali della corteccia olfattiva, di conseguenza i segnali in uscita corrispondono al nome dell'oggetto identificato come *rosa*. Anche un robot o un computer possono riconoscere una rosa emulando il processo naturale appena descritto, il segnale ricevuto in ingresso produce un altro segnale che corrisponde al nome dell'oggetto. Tuttavia tale riconoscimento è molto diverso dal nostro: nel caso dell'uomo il profumo della rosa è un qualia ed è correlato ai segnali elettrici generati dalle reti neurali, ma non è identico. Sentire questo profumo può evocare in noi ricordi, emozioni o produrre nuove sensazioni, in pratica è un'esperienza che avviene nella nostra coscienza che può segnare profondamente. A mio avviso l'esempio di Faggin sposa in pieno l'idea di John Searle sull'incapacità dell'intelligenza artificiale di attribuire un significato ai simboli. Pur possedendo la sintassi, cioè la struttura della frase, la semantica è tutt'altra cosa e non è presente nel computer. Trovo che l'esperimento mentale della *stanza cinese* di Serle sia un valido esempio per comprendere questa posizione. Riporto di seguito la spiegazione presente nel libro *"Mente, cervello, intelligenza artificiale"* di Eddy Carli e Fabio Grigenti: *"Supponiamo che uno di noi sia rinchiuso in una stanza, e gli vengano passati dei simboli cinesi, insieme a regole, in italiano, che fanno corrispondere a certi insiemi di simboli altri insiemi di simboli. La persona chiusa nella stanza non sa che i primi sono domande in cinese, e i secondi le risposte, in cinese, a quelle domande; tuttavia, se segue le regole, il suo comportamento linguistico sarà indistinguibile da quello di un parlante cinese. Come in questo caso non diremmo che l'uomo comprende il cinese, nonostante la sua capacità di manipolare simboli, così non lo diremmo nel caso di una macchina che disponesse dello stesso software per "comprendere" il cinese"* (Carli, Grigenti 2019, p. 42).

L'aver cercato di capire inutilmente come la coscienza potesse emergere da segnali elettrici o biochimici, porta Faggin a constatare che i segnali elettrici possono

produrre solo altri segnali elettrici o altre conseguenze fisiche come forza e movimento, ma mai sensazioni e sentimenti a causa della natura qualitativamente diversa di quest'ultimi. La conclusione che ne deriva è chiara: la coscienza dev'essere una proprietà irriducibile delle "particelle elementari", al pari della carica elettrica, un qualcosa di fondamentale che non deriva da proprietà più semplici. Un'idea questa che lo porta a ritenere che tutto l'universo debba essere cosciente. *"Questa idea è vecchia di millenni e si chiama panpsichismo. Il panpsichismo, però, non è mai stato preso sul serio dalla scienza, perché è considerato un'ipotesi che offre ben poche opportunità di falsificazioni. Non sembra infatti che ci sia alcuna connessione tra ciò che sentiamo e il mondo esterno. In poche parole, se per ogni azione fisica c'è una spiegazione che non richiede la coscienza, a che cosa serve la coscienza? Ecco perché essa è considerata epifenomenale, cioè un fenomeno che si accompagna a un altro fenomeno, ma che non è la vera causa di ciò che osserviamo. L'alternativa è quella di considerare che le leggi fisiche possano essere proprietà emergenti della coscienza, un assunto che per molti scienziati è difficile da accettare. Vorrebbe dire che il mondo oggettivo deriva dal mondo soggettivo! E questo è chiedere loro troppo"* (Faggin 2022, p. 132).

La conclusione è ovviamente in piena controtendenza con la posizione largamente condivisa dalla comunità scientifica, ma a mio avviso è ben argomentata e presenta degli spunti di riflessione che ritengo molto interessanti. Per Faggin l'elaborazione dell'informazione sensoriale diventa un fenomeno di esperienza interiore solo grazie alla coscienza, quindi è necessaria per conoscere anche le cose più banali. Una misurazione dei segnali cerebrali può rivelare solo alcuni aspetti dell'esperienza cosciente correlati al mondo fisico, ma non i *qualia* che una persona prova veramente.

"Percezione", "riconoscimento" e "comprensione" sono termini che Faggin collega alla teoria della coscienza partendo dall'apprendimento del bambino. Se sofisticate reti neurali sono attive ancora prima della nascita, l'autocoscienza che percepisce, conosce ed esplora il mondo fisico mediante il corpo nasce quando la coscienza comincia a distinguere i fenomeni a cui deve prestare attenzione. Egli ritiene che in un bambino le reti neurali organizzino "automaticamente" le informazioni sensoriali in dati-oggetti che da inconsciamente con il tempo possono poi essere riconosciuti

consapevolmente pur senza la conoscenza dei nomi. Secondo lui, il riconoscimento di pattern di base dev'essere una proprietà automatica del cervello che prescinde dalla coscienza, una funzione simile al *machine learning*. Una volta completamente formato, il bambino capisce di essere un corpo fisico che costituisce un mondo fatto di altre persone e oggetti. Questo processo di consapevolezza fa capire che la coscienza è coinvolta nell'apprendimento e permette di raggiungere livelli di competenza e comprensione via via più complessi che superano di gran lunga le capacità di una macchina.

Credo anch'io, come Faggin, che la consapevolezza di sé giochi un ruolo importante nella comprensione e di conseguenza nell'apprendimento fin dai primi anni di vita; l'apprendimento dei computer tramite le reti neurali artificiali invece non necessita della comprensione, in un certo senso è la bravura del programmatore a fare la differenza.

Vedo il computer come un validissimo amplificatore delle capacità umane, di conseguenza sono scettico verso un certo tipo di comunicazione che tende a esaltare l'intelligenza artificiale ponendola in un'ottica di sostituzione delle capacità umane a volte fallaci e imperfette. La vera intelligenza non sta in un algoritmo, è la comprensione cosciente a giocare un ruolo importante nell'uomo: *"la percezione cosciente è quindi il processo che converte le attività elettromagnetiche ed elettrochimiche autonome del sistema nervoso e del corpo in qualia. Il processo che "estrae" il significato dai qualia è chiamato comprensione. Sappiamo di comprendere perché "capiamo" il significato contenuto nella percezione "fatta" di qualia"* (Faggin 2022, p. 137). Questo riconoscimento cosciente va oltre le capacità del computer, il quale si limita al riconoscimento di pattern simbolici che non richiedono comprensione.

Proprio nella distinzione tra apprendimento automatico e apprendimento cosciente, Faggin pone un altro tassello sulla questione uomo - intelligenza artificiale, prendendo come esempio di apprendimento automatico quello di quando si insegna a un computer a riconoscere una tazza a partire dalla sua immagine. Per ottenere questo riconoscimento è necessario avere un set di immagini di tazze e un programma come la rete neurale artificiale capace di trovare automaticamente i tratti comuni nelle varie foto. Una volta che il programma ha imparato le correlazioni tra le

immagini a cui viene assegnato il nome "tazza", sembra che il computer capisca cos'è un tazza, ma per il programma si tratta in realtà di un nuovo simbolo. Come scritto in precedenza, questa capacità di riconoscimento di pattern di base, per Faggin deve esistere anche in un bambino prima che la percezione e la comprensione cosciente possano svilupparsi per contribuire all'apprendimento. Nel bambino, però, gradualmente si passa dalla distinzione dell'*oggetto tazza* rispetto allo sfondo, all'associazione del *suono tazza* ai qualia visivi e tattili di prima, per arrivare alla *funzione tazza* osservando la madre mentre versa il latte. Laddove le reti neurali artificiali richiedono molti esempi prima di effettuare la generalizzazione, l'uomo può comprendere consapevolmente con uno o pochi esempi. Quindi, la comprensione nell'uomo è un processo che inizia con la percezione, ma poi subentrano il ragionamento logico e il desiderio di capire. Concordo perciò con l'autore quando afferma che la vera intelligenza non è algoritmica, ma riguarda capacità quali la comprensione, il capire in profondità e trovare connessioni inaspettate. *"La nostra intelligenza va ben oltre le limitazioni del sistema nervoso, (...), ha origine in una realtà più vasta della realtà fisica che conosciamo. La vera intelligenza è intuizione, immaginazione, creatività, ingegno e inventiva"* (Faggin 2022, p. 143). Da queste parole si avverte chiaramente che ci stiamo spostando in una realtà diversa da quella descritta dalla fisica classica, una nuova visione che guarda alle scoperte della meccanica quantistica per la formulazione di una teoria tutta nuova.

Nella seconda parte di "Irriducibile" Faggin entra quindi nel vivo della sua teoria, premettendo già dalle prime righe che si può spiegare l'esistenza dell'interiorità senza invalidare l'evidenza sperimentale della fisica, ma che la difficoltà di capire il messaggio della fisica quantistica è dovuta al fatto che essa descrive caratteristiche del mondo interiore e non esteriore. L'aver cercato di spiegare la vita e la coscienza con le leggi della fisica classica, per Faggin ci ha condotti in un vicolo cieco, dobbiamo perciò fare marcia indietro e considerare invece che la realtà sia molto più vasta e profonda e che la realtà fisica che noi conosciamo sia solo uno degli stati, quello emergente. Non è stato facile per l'autore cercare collaboratori interessati e disposti a sviluppare questo nuovo campo, finché non si è imbattuto nella *teoria OPT* del professor D'Ariano e della sua équipe, la quale dimostra che la fisica quantistica

deriva dall'informazione quantistica. Tale informazione, per Faggin, non potrebbe aver senso se fosse priva di significato, ecco che subentra il ruolo della coscienza. I qualia possono di conseguenza essere interpretati come "l'esperienza di un sistema quantistico che si trova allo stato puro", la connessione che mancava tra la teoria di D'Ariano e quella di Faggin e che porta alla formulazione della *teoria QIP*: Quantum Information-based Panpsychism, da poco pubblicata. *"Uno stato puro è uno stato ben definito, ma non clonabile. Ha quindi tutte le caratteristiche straordinarie di un'esperienza cosciente, che è privata e perciò conoscibile soltanto dal sistema che è in quello stato. "Da fuori", cioè nello spazio-tempo della fisica, si può osservare solo un evento classico quando il sistema quantistico interagisce con un apparato di misura, mentre il suo è lo spazio della conoscenza, uno spazio fatto di qualia che si può descrivere in maniera astratta come uno spazio di Hilbert complesso e multidimensionale"* (Faggin 2022, p. 157). Tale teoria ha l'intento di dimostrare che la coscienza e il libero arbitrio richiedono le proprietà dell'entanglement quantistico per essere descritti propriamente.

Nel presente testo non è mia intenzione addentrarmi nei dettagli di questa teoria scientifica, ma esporre e analizzare gli assunti che stanno alla base della sua posizione sul rapporto uomo - intelligenza artificiale. Tuttavia, per una maggiore comprensione del pensiero di Faggin, penso sia importante avere presente altri due concetti della sua teoria, quello di *Seity*² e di *Uno*³. Abbiamo visto che lo stato puro quantistico non è clonabile, ogni tentativo di copiarlo lo cambia e lo riduce all'informazione classica, un po' come quando proviamo a descrivere a parole un sentimento che proviamo: la descrizione a parole, per quanto possa essere dettagliata a mio avviso non potrà mai sostituire l'essenza intima di quello stato d'animo così personale e soggettivo. Similarmente, anche un computer potrebbe ripetere le stesse parole con cui una persona descrive l'amore, ma l'esperienza interiore dell'individuo è unica, non clonabile; il computer si limiterebbe a descrivere la sensazione in maniera accurata dando l'impressione che stia provando quel sentimento. Ecco che l'ente cosciente può essere descritto come un sistema quantistico che è in uno stato puro, tale sistema è chiamato *seity*: *"Nel contesto QIP, una seity è definita come un ente cosciente che sa di esserlo, che può agire con libero arbitrio e che ha un'identità permanente come l'abbiamo noi. (...) Sapere di*

essere cosciente è una marcia in più rispetto agli enti che sono coscienti ma non sanno di esserlo, e che quindi non hanno un'identità propria. Definisco "thoughtform" questi enti quantistici che, non avendo senso di sé, non possono guidare la loro esperienza con il libero arbitrio che non sanno di avere. Una seity è un "campo" in uno stato puro che esiste in una realtà più vasta del mondo fisico che contiene il nostro corpo. Una seity esiste anche senza il corpo fisico, e questa è un'affermazione cruciale, perché implica che la nostra esistenza non dipende dal corpo" (Faggin 2022, p. 159).

Una tesi senza dubbio singolare, che giustifica l'esistenza della coscienza e del libero arbitrio collocandoli sul livello degli stati puri quantistici, arrivando a postulare l'esistenza di un tutto olistico che contiene tutte le proprietà fondamentali. Questo tutto viene chiamato *Uno* per distinguerlo dal campo della fisica. In questa concezione, l'universo fisico è una delle realtà ed è il risultato del *collasso della funzione d'onda*⁴, dove esiste il nostro corpo e tutto ciò che è misurabile nello spazio-tempo. Ma se queste seity esistono, per Faggin dev'esserci una ragione che le giustifichi ed è qui che *Uno* assume una connotazione teleologica, un principio di auto-conoscenza e di auto-realizzazione. Mi rendo conto che una simile teoria, ipotizzata per di più dal fisico inventore del microprocessore, possa in un primo momento disorientare se non addirittura sembrare inaccettabile; personalmente ritengo invece che un "pensiero laterale", cioè la capacità di problem solving che guarda al problema da un'altra prospettiva, possa essere produttivo e possa portare per lo meno un contributo allo studio della fisica quantistica e alla risoluzione di molti quesiti ancora irrisolti.

Penso altresì che in questa interpretazione completamente nuova la fisica non venga accantonata, poiché le leggi fisiche continuano ad avere il loro valore sul piano della nostra realtà. Da Faggin queste leggi infatti vengono concepite come *le leggi sintattiche dei linguaggi* che le seity hanno sviluppato spontaneamente nelle loro comunicazioni. Resta chiaro però che da un computer non potrà mai "emergere" la coscienza e il libero arbitrio poiché esso non è una seity, ma una macchina classica descritta da una teoria in cui non esiste l'entanglement; tutt'al più il computer potrà manifestare il libero arbitrio della seity che lo controlla.

4. Il dibattito contemporaneo

Nel dibattito contemporaneo su mente, coscienza e intelligenza artificiale, la posizione di Faggin è senz'altro singolare e chiaramente differisce da molti illustri pareri, tuttavia non mancano dei punti in comune con alcuni famosi esperti.

Già nel precedente capitolo avevo riportato l'esperimento mentale della *stanza cinese* pensato intorno agli anni ottanta da John Searle come risposta alla teoria di quella che lui definisce *intelligenza artificiale "forte"* (strong AI), cioè quella corrente di pensiero che considera il computer equivalente alla mente umana. Per Searle, invece, le macchine sono sì in grado di manipolare sintatticamente simboli, ma non hanno la capacità di comprenderli, manca dunque il livello semantico che solo la mente umana possiede. Il modello computazionale della mente che sta alla base dell'intelligenza artificiale non considera quindi la coscienza e l'intenzionalità. Sempre secondo Searle, sia a livello cerebrale che mentale, l'uomo si rivela un animale razionale complesso, impossibile da paragonare all'hardware e al software di un computer e alla logica binaria che governa la programmazione delle macchine. In Faggin la posizione a riguardo è molto simile pur partendo però da un terreno completamente diverso, cioè quello della fisica quantistica e dell'universo ontologico, egli vede la coscienza e il libero arbitrio collocati in uno stato puro quantistico diverso dal piano fisico. Le ricerche di Searle partono invece dal linguaggio e arrivano agli stati cognitivi ed emotivi, trattano l'intenzionalità e la *fuzzy logic*, cioè la logica del vago che domina le nostre emozioni e sensazioni.

Un'altra posizione fortemente critica del funzionalismo meccanico è stata quella del famoso biologo Gerald Edelman, morto nel 2014, teorizzatore del *darwinismo neurale*¹. Egli portò avanti la critica a chi proponeva la coscienza come il risultato di un processo funzionale, una linea teorizzata anche da Putnam in "Mente, linguaggio e realtà" (Adelphi 1987). Per Edelman, al contrario, gli stati mentali non potevano essere definiti nei termini degli stati di un computer, ma la sua base di partenza, a differenza di Searle e ancor più di Faggin, era una teoria della coscienza

biologicamente fondata che estendeva le informazioni provenienti dalla teoria della selezione dei gruppi neuronali (TSGN) del cervello. Un punto di vista molto particolare che criticava l'intelligenza artificiale come modello analogo del cervello, ma che non escludeva la prospettiva di produrre macchine coscienti; non a caso riuscì a progettare alcuni "automi percettivi" basati sul concetto di selezione e non su comandi e istruzioni, tra questi il famoso *NOMAD* il cui sistema nervoso era simulato su un computer. La posizione nei riguardi dell'intelligenza artificiale era però chiara: *"Il nostro cervello non è costruito seguendo una serie di istruzioni e di regole fisse. Il mondo non è organizzato come un nastro magnetico portatore di un unico messaggio. Se consideriamo il nostro cervello secondo la teoria evuzionista di Darwin, lo possiamo vedere come l'evoluzione di una popolazione di neuroni, nella quale sopravvivono quelli che sono in grado di adattarsi ai cambiamenti dell'ambiente esterno"* (Carli, Grigenti 2019, pp. 17-18). Personalmente non solo trovo molto interessante la teoria del darwinismo neuronale di Edelman, ma mi vedo in completo accordo con lui quando affermava che sulla materia della mente è necessaria una visione globale che colleghi le varie discipline come fisica, filosofia, biologia e psicologia.

Una posizione alquanto singolare, nettamente in contrasto con Faggin e con quello che è stato detto finora, è quella del filosofo Daniel Dennett. Tra tutti i filosofi egli è il più schierato a favore dell'intelligenza artificiale e dello studio della mente dell'uomo attraverso i modelli di calcolo. Dennett, tramite esperimenti mentali e sofisticate congetture scientifiche, le sue famose "pompe intuitive", mostra come una mente dotata di comprensione possa emergere dai processi non intenzionali della selezione naturale. Negli anni non sono mancate accese dispute e dure reazioni in particolare da parte di Searle. L'obiettivo di Dennett è proprio quello di cancellare definitivamente il dualismo cartesiano tra *res cogitans* e *res extensa*, arrivando a negare l'esistenza di "anima" o coscienza. *"Credo che alla gente non piaccia molto sentir dire che noi siamo soltanto delle macchine, ma questo dipende dal fatto che ci si riferisce a un'idea di macchina troppo semplicistica. Io non sto sostenendo che gli esseri umani siano riconducibili a dei tostapane, a lavatrici o a televisori. Sto dicendo che essi sono macchine estremamente complesse e sofisticate..."* (Carli, Grigenti 2019, p. 119). Nulla di più lontano dal pensiero di Faggin, dove il piano fisico è

considerato solo uno dei piani di una realtà ben più ampia e la materia è l'aspetto simbolico di enti coscienti che comunicano. Se poi aggiungiamo i risvolti teleologici fatti di concetti come "scopo della vita" e termini quali "creazione", allora le distanze tra le due posizioni diventano enormi.

Dopo aver esaminato solo alcuni dei tanti punti di vista che si sono susseguiti in questi anni, in generale possiamo dire che dal clamore del dibattito scaturito negli anni cinquanta il tema dell'intelligenza artificiale nei decenni a seguire ha avuto una grande influenza nella nostra società, entrando indirettamente anche nelle case dei non "addetti ai lavori". Mi vengono in mente alcuni esempi, tra i quali il computer HAL 9000 del film *"2001: Odissea nello Spazio"* di Kubrick del 1968, gli albori della musica elettronica dei primi anni settanta con il quartetto dei Kraftwerk e il loro tipico sound "robotico", per non parlare delle serie TV che negli anni ottanta avevano per protagonisti dei robot, su tutti la celebre auto *"Kitt"* del telefilm *"Supercar"*.

Ora il tema dell'intelligenza artificiale è tornato al centro dell'attenzione, probabilmente grazie agli enormi passi avanti della tecnologia nell'ambito di applicazioni di uso comune; *"ChatGPT"*, *"Siri"*, *"Alexa"* e i sistemi di guida autonoma hanno suscitato nuovo interesse ponendo inoltre l'accento su questioni legate ad aspetti etici e normativi. Alla sfida uomo - computer in termini di intelligenza e consapevolezza si è aggiunta la discussione tra le persone a favore di un pieno progresso nell'ambito dell'intelligenza artificiale e chi invece vede i rischi di uno sviluppo senza regole e chiede un codice morale che possa indirizzare la strada futura.

La questione etica è estremamente complessa a causa dei diversi ambiti coinvolti: economia, lavoro, sicurezza e ambiente sono solo alcuni di questi aspetti. Il famoso filosofo e scienziato informatico Jerry Kaplan, nel volume dal titolo *"Le persone non servono. Lavoro e ricchezza nell'epoca dell'intelligenza artificiale"* (LUISS, 2016) parla del problema di come l'IA abbia creato un'enorme ricchezza concentrata nelle mani di un gruppo ristretto di persone e di come tutto questo abbia favorito squilibri nella società. La continua e veloce evoluzione delle competenze nell'ambito lavorativo è già una delle difficoltà che si notano e il rischio molto probabile è che possa aumentare sempre più nel breve termine. L'analisi di Kaplan non solo evidenzia le varie criticità nei diversi ambiti sopra citati, ma propone soluzioni

positive che consentono di guardare al futuro dell'IA nella prospettiva di un aiuto per l'uomo.

Altri rischi sono stati individuati dal filosofo svedese Nick Bostrom, in particolare la minaccia legata alla *superintelligenza*², *"un intelletto qualunque, capace di superare di molto le prestazioni cognitive degli esseri umani, in quasi tutti i domini di interesse"* (Carli, Grigenti 2019, p. 148). Uno scenario a tratti inquietante tanto da non aver lasciato indifferente il famoso imprenditore Elon Musk, il quale ha donato dieci milioni di dollari a una fondazione che si occupa di impedire la realizzazione delle previsioni descritte nel libro.

In questo panorama di riflessioni a maggior ragione penso sia utile tenere in considerazione il punto di vista di Federico Faggin, la sua prospettiva di sicuro sposta l'attenzione unica verso la macchina e riporta non solo l'uomo al centro, ma la vita in generale, cercando di ridare significato alla coscienza e all'esistenza. Anche lui all'inizio di "Irriducibile" evidenzia alcune preoccupazioni: *"Sono convinto che una razionalità informata unicamente dai principi del materialismo, del riduzionismo e della sopravvivenza del più adatto possa portare soltanto a una competizione sfrenata, al razzismo e alla guerra"* (Faggin 2022, p. 57).

Credo fortemente che lo sforzo comune dovrebbe proprio essere quello di porre l'attenzione verso l'uomo, in una sorta di nuovo umanesimo che non guardi al singolo, ma all'essere umano nel contesto di una comunità fatta di relazioni vive. Non condivido le campagne contro il progresso a prescindere, ma il progredire dell'intelligenza artificiale dev'essere fondato su valori condivisi e non guidato dagli interessi economici di un ristretto gruppo di persone. Per portare avanti tutto questo serve sicuramente un'informazione e una consapevolezza maggiori nelle persone in generale. A tal riguardo ritengo che l'educazione giochi un ruolo importante, a patto che coinvolga tutti i soggetti: le famiglie, la società civile e le istituzioni. La scuola avrà un compito fondamentale da qui in avanti: tenere insieme l'ambito umanistico con quello del digitale e dell'intelligenza artificiale, senza contrapposizioni, nell'ottica di un apprendimento consapevole della tecnologia e del ruolo che dovrà avere.

Note

1. Introduzione

¹*Test di Turing. "Nell'articolo Computing Machinery and intelligence (1950), Alan Turing proponeva di sostituire la domanda "Le macchine possono pensare?" con il seguente problema: è possibile progettare una macchina che sia estremamente difficile identificare come tale nel corso di una conversazione "cieca", in cui cioè l'interlocutore umano non vede il suo partner, e comunica con lui solo attraverso messaggi scritti? Questa situazione sperimentale è nota come "test di Turing": le macchine che si dimostrano capaci di superare il test possono essere definite "macchine pensanti". (Carli, Grigenti 2019, p. 62)*

²*Machine learning. "È un insieme di procedure che insegna ai computer e ai robot a compiere azioni e attività come le farebbe un sistema cognitivo "naturale" non semplicemente eseguendo un programma, ma imparando dall'esperienza. Per questo, tali algoritmi mostrano capacità adattiva, ovvero sono in grado di migliorare le loro prestazioni via via che gli "esempi" da cui apprendere aumentano." (Carli, Grigenti 2019, p. 151)*

2. Fisica "classica", fisica quantistica e natura della vita

¹*Quadro concettuale CIF. "Le iniziali delle parole Coscienza, Informazione e Fisico formano l'acronimo CIF, in cui la C indica lo spazio della coscienza (spazio-C), la I quello dell'informazione (spazio-I) e la F sta per lo spazio fisico. Il quadro CIF postula che la natura fondamentale della realtà sia costituita da un campo che ha un aspetto semantico (spazio-C) e un aspetto simbolico (spazio-I) irriducibili. In questo quadro:*

1. Lo spazio-C è lo spazio semantico soggettivo della coscienza. È uno spazio e un tempo esperienziale extra-fisico - non lo spazio della fisica - che contiene il mondo interiore di tutti gli enti coscienti. Lo spazio-C è inseparabile dallo spazio simbolico (spazio-I). Insieme formano l'ontologia di Uno. (...) Lo stato puro di un sistema quantistico rappresenta l'esperienza di un ente cosciente e si riferisce allo spazio-C.

2. Lo spazio-I è lo spazio simbolico oggettivo che contiene l'informazione viva in cui una parte del significato è stata tradotta dalle seity per essere comunicata simbolicamente alle altre seity. (...)

3. Lo spazio-F è il mondo fisico sperimentato nella coscienza della seity che crede di essere l'organismo vivente in cui è "incarnata". La natura dello spazio-F dipende quindi in maniera cruciale dalla specie di organismo vivente e dai qualia della seity. (...)" (Faggin 2022, pp. 258-259-260)

²Radiazione del corpo nero. "È lo spettro di frequenza della luce emessa da un oggetto rovente in funzione della sua temperatura. Il comportamento anomalo di questa radiazione fu spiegata da Max Planck nell'anno 1900 usando un espediente matematico che all'inizio sembrò ingiustificato. Planck trovò che, se lo scambio di energia tra materia e radiazione fosse avvenuto soltanto in multipli interi di un valore discreto di energia - chiamato quanto di energia -, allora il mistero sarebbe stato risolto." (Faggin 2022, pp.36-37)

³Effetto fotoelettrico. "Nel 1905, un giovane Albert Einstein riuscì a spiegare l'effetto fotoelettrico assumendo che l'interazione tra luce e materia che produce elettroni sia causata da "particelle di luce" aventi quanti di energia forniti dalla relazione di Planck. Einstein ipotizzò che la luce, quando interagisce con gli atomi della materia, si comporti come tante particelle, e non come un'onda, come invece aveva dimostrato Thomas Young cent'anni prima." (Faggin 2022, p. 37)

⁴Trasformazioni di Lorentz. "Nel 1905 Albert Einstein spiegò anche la misteriosa trasformazione di Lorentz, ossia il fatto che gli oggetti descritti dalle equazioni di Maxwell violano la semplice trasformazione inerziale galileiana quando si muovono con una velocità vicina a quella della luce. Quest'anomalia fu risolta elegantemente ipotizzando che la velocità della luce sia la stessa in tutti i sistemi di riferimento inerziali, cioè i sistemi che si muovono a velocità costante." (Faggin 2022, p. 37)

⁵Relatività generale. "È la teoria della gravitazione pubblicata da A. Einstein nel 1915. Tale teoria generalizza la relatività speciale e sostituisce la legge di gravitazione universale di Newton, che richiede un'azione a distanza, con una teoria locale. Si tratta di una teoria matematica che lega lo spazio-tempo a quattro dimensioni con la distribuzione e il flusso di massa, energia e impulso, usando un sistema di equazioni differenziali alle derivate parziali. In questo modo, la forza peso diventa una forza apparente osservata nei riferimenti non inerziali (...)." (Faggin 2022, pp. 260-261)

⁶Entanglement quantistico. "L'entanglement quantistico, o correlazione quantistica, è un fenomeno che non esiste nella fisica classica. Quando avviene un'interazione tra due sistemi quantistici, si crea un sistema più ampio il cui stato non è più descrivibile come la semplice somma degli stati dei sistemi che hanno interagito. Questo significa che i sistemi non sono più indipendenti, e quindi la misurazione di un'osservabile di uno dei due

determina simultaneamente il valore della stessa osservabile anche per l'altro sistema, indipendentemente da quella che è la loro distanza fisica. (...) L'unico modo per spiegare questo fenomeno è che lo stato che si misurerà non può esistere prima della misurazione, ma dev'essere creato simultaneamente per i due sistemi durante il processo di misurazione (...)." (Faggin 2022, p. 253)

⁷Non-località. Si riferisce alla violazione del principio di realismo locale, che afferma che un oggetto può essere influenzato direttamente soltanto da ciò che esiste nelle sue vicinanze immediate. La statistica delle misurazioni fatte su particelle quantistiche entangled dimostra la non-località, perché viola il limite classico stabilito dal teorema di Bell. Quindi la realtà fisica è non-locale, benché non sia possibile comunicare con velocità superiore a quella della luce (dato che il risultato di una misura è probabilistico). Pertanto, la non-località quantistica è compatibile con la relatività speciale. L'unica interpretazione compatibile con l'entanglement è che lo stato che si misura non sia preesistente, ma sia invece creato nel momento stesso della sua misurazione. Nel contesto della teoria QIP, la creazione dello stato che si misura corrisponde a un atto di libero arbitrio del sistema quantistico cosciente che trasforma l'informazione quantistica in informazione classica." (Faggin 2022, p. 257)

⁸Informazione viva. "Con il termine "informazione viva" mi riferisco alle interazioni quantistiche che avvengono tra particelle elementari, atomi e molecole, come succede per esempio nella formazione di una proteina o nel funzionamento di una cellula vivente. Sistemi fatti d'informazione viva sono sistemi quantistici e classici ordinati, descrivibili come organizzazioni di qubit entangled che interagiscono quantisticamente come nel caso delle cellule viventi. Questo tipo di interazioni sono per il momento poco conosciute, visto che gli organismi viventi sono stati studiati come se fossero sistemi biochimici classici e non invece come sistemi informatici sia quantistici sia classici." (Faggin 2022, pp. 255, 256)

3. La natura irriducibile della coscienza e i computer

¹Qualia. "Il termine qualia si riferisce alle sensazioni e ai sentimenti che emergono nella coscienza di un ente in seguito alla percezione di una particolare informazione viva o di una nuova comprensione endogena (...)." (Faggin 2022, p. 260).

²Seity. "La seity è un ente quantistico con tre proprietà fondamentali irriducibili e indivisibili: coscienza, agentività e identità. Le seity elementari, che ho chiamato unità di coscienza (UC), emanano direttamente da Uno. La combinazione di seity comunicanti, ciascuna con la propria coscienza, identità e libero arbitrio, dà vita a gerarchie di seity, significati, simboli, regole sintattiche e linguaggi. Comunicando tra di loro, le seity creano, strato dopo strato,

varie organizzazioni in cui sperimentare se stesse e aumentare la propria conoscenza di sé." (Faggin 2022, p. 261)

³Uno. "Uno è la totalità di ciò che esiste, sia in potenza sia in atto. Uno è irriducibilmente olistico e dinamico. Uno è l'interiorità che connette "da dentro" tutte le sue creazioni. Uno desidera conoscere se stesso e vuole autorealizzarsi. Da Uno emergono unità di coscienza che comunicano tra di loro e si combinano, creando così tutta la realtà sia interiore che esteriore. Ciò che chiamiamo materia è l'aspetto esteriore della realtà, fatto di forme simboliche che rappresentano il significato che le seity si scambiano per conoscersi di più." (Faggin 2022, p. 265)

⁴Collasso della funzione d'onda. "Nella fisica quantistica, l'evoluzione dello stato di un sistema isolato è rappresentata dall'evoluzione di un vettore unitario in uno spazio di Hilbert complesso a N dimensioni la cui direzione cambia nel tempo. Quando si fa una misurazione, il sistema quantistico interagisce con il sistema di misura e produce un'osservabile classica nello spazio-tempo; per esempio, attiva un sensore in una certa regione dello spazio-tempo che noi attribuiamo alla presenza di una particella (nessuno, però, ha mai visto una particella). Questo processo si chiama "collasso della funzione d'onda", o collasso del vettore di stato, e costituisce il problema della misurazione, che è ancora dibattuto al giorno d'oggi. La teoria matematica della fisica quantistica predice soltanto la probabilità dei possibili stati che si potranno misurare, ma non predice lo stato che si manifesterà." (Faggin 2022, p. 252)

4. Il dibattito contemporaneo

¹Darwinismo neurale. Nel volume "Neural Darwinism" (1987) Gerald Edelman espone la teoria dello sviluppo e dell'organizzazione delle "funzioni cerebrali superiori", nei termini del processo della "selezione neuronale di gruppo", una teoria del cervello molto ampia che spazia in varie discipline, dalla biologia alla psicologia. Tale teoria viene applicata a ogni livello neurale, dai collegamenti tra neuroni fino alla coscienza, e si basa sull'idea del darwinismo neurale: "con il darwinismo neurale Edelman ha detronizzato i geni, e in questi ultimi anni le neuroscienze teoriche e sperimentali si sono evolute lungo tracce delineate e argomentate in questo libro". (Carli, Grigenti 2019, p. 6)

²Superintelligenza. Nick Bostrom con il suo libro "Superintelligence. Paths, Dangers, Strategies" (2014) ha dato avvio a un importante dibattito sui rischi legati allo sviluppo senza limiti dell'intelligenza artificiale. L'idea di fondo è che l'intelligenza artificiale completamente

dispiegata in tutte le potenzialità possa diventare a noi incomprensibile e lesiva nei confronti dell'uomo.

Bibliografia

- Carli Eddy, Grigenti Fabio, *Mente, cervello, intelligenza artificiale*, Pearson, 2019.
- Damasio, Davidson, Dennett, Dreyfus, Edelman, Fodor, Rorty, Searle, Stich, *Cervelli che parlano. Il dibattito su mente, coscienza e intelligenza artificiale*, Introduzione e cura di E. Carli, Paravia Bruno Mondadori, 2003.
- Faggin Federico, *Irriducibile. La coscienza, la vita, i computer e la nostra natura*, Mondadori, 2022.
- Faggin Federico, *Silicio. Dall'invenzione del microprocessore alla nuova scienza della consapevolezza*, Mondadori, 2019.
- Okasha Samir, *Il primo libro di filosofia della scienza*, Piccola Biblioteca Einaudi, 2006.
- Searle R. John, *Il mistero della realtà*, Raffaello Cortina Editore, 2019.