

Bruno Cancellieri, ChatGPT e Claude.ai

A che serve la coscienza?

Ipotesi sull'architettura della mente

22 luglio 2026

Copyright © Bruno Cancellieri

Indice

Nota	4
1 Introduzione	5
2 I programmi inconsci: gli agenti della mente	7
3 Lo spazio cosciente come simulatore della realtà	9
4 Le emozioni: il sistema di controllo della simulazione	11
5 Apprendimento: come la mente modifica se stessa	13
6 I limiti della simulazione	15
7 Rapporti con le teorie esistenti	17
8 Una teoria architettuale della mente	18

Nota

Questo articolo nasce da un dialogo con ChatGPT e Claude.ai, secondo un metodo che ho adottato anche per altri lavori: propongo un tema, discuto con questi strumenti l'impostazione e i contenuti, e faccio scrivere e poi correggere i testi attraverso un confronto articolato, in cui le modifiche vengono proposte, argomentate e discusse insieme.

Il quadro teorico specifico si sviluppa in questi dialoghi, ma si inserisce in una formazione psicologica e filosofica che ho maturato in autonomia ben prima di ricorrere a questi strumenti.

La responsabilità delle tesi sostenute, e delle loro eventuali debolezze, resta comunque mia.

Bruno Cancellieri

1 Introduzione

Ogni caratteristica biologica sufficientemente complessa solleva una domanda fondamentale: quale vantaggio evolutivo ne ha favorito la comparsa e giustificato il mantenimento? Per le ali, gli occhi o il sistema immunitario la risposta è relativamente evidente. Molto meno evidente è la funzione della coscienza.

La difficoltà nasce da un fatto ormai ben documentato: una parte enorme dell'attività mentale si svolge senza alcuna esperienza cosciente. Riconosciamo volti, comprendiamo il linguaggio, interpretiamo situazioni, impariamo dall'esperienza, coordiniamo movimenti complessi e prendiamo molte decisioni senza esserne pienamente consapevoli. Quanto questo valga anche per le decisioni vere e proprie è ancora oggetto di discussione: gli esperimenti di Benjamin Libet e i loro sviluppi successivi hanno ricevuto interpretazioni divergenti, e non esiste un consenso sull'idea che la coscienza venga semplicemente informata di decisioni già prese. Rimane però un dato difficilmente contestabile: l'inconscio svolge una quantità enorme di elaborazioni cognitive.

Da qui il problema affrontato in questo saggio. **Se l'inconscio è già capace di percepire, apprendere, valutare, pianificare e guidare il comportamento, quale funzione rende utile l'esistenza della coscienza?**

La risposta proposta in queste pagine è che la coscienza non è né il centro di comando della mente né uno spettatore passivo dei processi inconsci: è un componente specifico dell'architettura cognitiva, evolutosi per svolgere un tipo particolare di elaborazione che aumenta le capacità adattive dell'intero sistema. Comprendere la coscienza significa allora comprendere l'architettura complessiva della mente.

Che cosa questa teoria spiega, e che cosa no

Nel dibattito contemporaneo si distinguono almeno due sensi del termine; Ned Block li ha chiamati coscienza d'accesso e coscienza fenomenica. La coscienza d'accesso indica il fatto che certe informazioni sono globalmente disponibili al sistema cognitivo: possono essere riferite verbalmente, usate nel ragionamento, integrate con altre informazioni. La coscienza fenomenica indica invece il fatto che esiste "qualcosa che si prova" a vedere il rosso, sentire dolore o provare nostalgia: l'esperienza soggettiva in quanto tale.

Questo saggio propone una teoria funzionale della coscienza d'accesso: cerca di spiegare a che cosa serva rendere globalmente disponibili certe informazioni, non perché tale disponibilità sia accompagnata da un'esperienza qualitativa. Il cosiddetto "problema difficile" formulato da David Chalmers — perché l'elaborazione dell'informazione sia accompagnata da esperienza soggettiva — resta fuori dalla portata di questa teoria, come del resto di tutte le teorie funzionali oggi disponibili. Chi cerca in queste pagine una spiegazione dei qualia non la troverà; chi cerca una risposta alla domanda "quale vantaggio adattativo offre il tipo di elaborazione che chiamiamo cosciente" troverà una proposta articolata.

Una mente organizzata su livelli funzionali

Quando si parla di coscienza e inconscio è facile immaginare due "menti": una razionale, consapevole e deliberativa, l'altra automatica, impulsiva e nascosta. Questa immagine è intuitiva ma fuorviante. È più utile interpretare la mente come un unico sistema composto da numerosi processi specializzati che cooperano continuamente tra loro.

Alla base dell'architettura operano moltitudini di programmi inconsci, che analizzano gli stimoli provenienti dall'ambiente e dal corpo, riconoscono configurazioni percettive, richiamano ricordi, formulano previsioni, generano emozioni e preparano possibili risposte. Questi processi raggiungono livelli di complessità straordinari senza diventare oggetto di consapevolezza: riconoscere un volto, comprendere una frase, mantenere l'equilibrio, intuire l'intenzione di un interlocutore. L'inconscio non è una parte primitiva o rudimentale della mente: è il principale motore dell'attività cognitiva.

Su questa base intervengono tre componenti che i capitoli seguenti analizzeranno in dettaglio. Le emozioni attribuiscono importanza: valutano continuamente quanto una situazione sia rilevante per i bisogni dell'organismo, trasformando la complessità del mondo in una gerarchia dinamica di priorità. L'attenzione seleziona: tra le moltissime informazioni disponibili, stabilisce quali riceveranno ulteriori risorse cognitive. Lo spazio cosciente, infine, integra.

Con l'espressione **spazio cosciente** non intendo un luogo fisico del cervello, ma una modalità di funzionamento dell'intero sistema cognitivo: l'insieme delle informazioni che, in un dato momento, vengono mantenute temporaneamente disponibili per essere integrate, confrontate e manipolate. La coscienza non riceve direttamente il mondo esterno: riceve il risultato di una lunga catena di elaborazioni inconscie già interpretate, valutate e selezionate. Per questo non è né il centro di comando né un semplice spettatore, ma uno spazio di integrazione in cui il sistema elabora con particolare flessibilità una piccola parte delle informazioni che esso stesso ha prodotto.

Un avvertimento è necessario. Presenterò questi componenti come se fossero disposti in serie: prima la valutazione emotiva, poi la selezione attentiva, poi l'elaborazione cosciente. È una semplificazione espositiva. La ricerca sperimentale mostra, ad esempio, che attenzione e coscienza sono parzialmente dissociabili: è possibile dirigere l'attenzione verso stimoli che non diventano coscienti, e alcuni contenuti sembrano raggiungere la coscienza con un coinvolgimento attentivo minimo. Il rapporto reale tra questi processi è circolare e in parte ancora da chiarire; lo schema seriale va inteso come una prima approssimazione.

2 I programmi inconsci: gli agenti della mente

L'espressione "programmi inconsci" non indica un luogo del cervello né una struttura anatomica: indica un modo di descrivere il funzionamento del sistema cognitivo. L'idea proposta in questo saggio è che la mente non operi come un unico elaboratore centrale, ma come una popolazione di processi relativamente autonomi, ciascuno specializzato nel riconoscere determinate situazioni e nel produrre determinate risposte.

In termini funzionali, ogni programma:

- riconosce una determinata configurazione di eventi;
- le attribuisce un significato rispetto ai bisogni dell'organismo;
- valuta le possibili conseguenze;
- propone una risposta comportamentale o cognitiva.

L'intera vita mentale nasce dall'interazione continua tra migliaia di questi programmi, la maggior parte dei quali lavora senza mai entrare nello spazio cosciente: quando riconosciamo un volto non osserviamo il processo di riconoscimento, quando comprendiamo una frase non assistiamo alle operazioni linguistiche. I programmi producono direttamente risultati; lo spazio cosciente vede soltanto una piccola parte dei loro prodotti.

Evitare l'homunculus

Questa descrizione richiede subito una precisazione, perché contiene un rischio noto in filosofia della mente come fallacia dell'homunculus: spiegare le capacità della mente attribuendo le stesse capacità ai suoi componenti. Dire che un programma "riconosce", "valuta" e "propone" sembra presupporre, dentro ogni programma, una piccola mente che riconosce, valuta e propone — e allora non si è spiegato nulla.

Questi verbi vanno quindi intesi come abbreviazioni funzionali. Un programma "riconosce" una configurazione nello stesso senso in cui un termostato "riconosce" una temperatura: risponde selettivamente a certe condizioni, senza comprenderle. E ogni programma complesso va pensato come scomponibile in sottoprocessi più semplici, questi in altri ancora più semplici, fino a livelli in cui le operazioni sono così elementari da non richiedere alcuna intelligenza. È la strategia con cui Marvin Minsky, nella sua *Society of Mind*, dissolveva il problema: l'intelligenza del sistema non risiede in nessun componente, ma emerge dalla loro organizzazione.

Programmi innati e programmi appresi

Alcuni programmi derivano dall'evoluzione biologica: il riflesso di orientamento, la risposta al dolore, l'attaccamento, il riconoscimento di alcune espressioni emotive. Altri vengono costruiti durante la vita: comprendere una lingua, guidare un'automobile, interpretare un'ironia, risolvere un'equazione. Con il tempo, molti comportamenti inizialmente coscienti si trasformano in nuovi automatismi: l'apprendimento consiste in larga misura proprio nella costruzione di nuovi programmi inconsci.

Competizione e accesso allo spazio cosciente

Ogni momento della vita presenta molte possibilità: parlare o tacere, fuggire o restare, continuare un'attività o interromperla. Programmi diversi propongono contemporaneamente risposte diverse, e la mente deve risolvere una competizione continua. Emozioni, attenzione e spazio cosciente partecipano alla regolazione di questa competizione.

Ne risulta un'immagine della mente non come un sovrano che impartisce ordini, ma come una società di agenti specializzati, nessuno dei quali conosce l'intero sistema. Rispetto alla formulazione di Minsky, la prospettiva qui adottata accentua un punto: gli agenti non competono soltanto per produrre direttamente un comportamento; competono anche per l'accesso allo spazio cosciente, dove le loro proposte possono essere integrate, simulate e confrontate con quelle di altri programmi.

Una conseguenza importante riguarda il modo di interpretare la decisione. Molte teorie cercano il luogo in cui una decisione "nasce"; ma la domanda è probabilmente mal posta. Una decisione non nasce in un punto: emerge dall'interazione di molti programmi che cooperano e competono continuamente. Lo spazio cosciente non li crea; li mette temporaneamente in relazione.

3 Lo spazio cosciente come simulatore della realtà

Rimane la domanda decisiva. **Che cosa fa, esattamente, lo spazio cosciente con le informazioni che vi accedono?** La risposta proposta in questo saggio è che la sua funzione principale consiste nel costruire, mantenere e manipolare modelli della realtà.

La mente non opera direttamente sulla realtà

Spesso parliamo come se la mente entrasse in contatto diretto con il mondo. In realtà ciò che manipoliamo non è mai la realtà stessa, ma una sua rappresentazione: quando immaginiamo la nostra casa, pianifichiamo un viaggio o riflettiamo su una scelta, lavoriamo su modelli mentali costruiti dal sistema cognitivo. Questo vale per ogni organismo dotato di sistema nervoso: nessuno può operare direttamente sulla realtà esterna; deve elaborarne un modello sufficientemente utile a guidare il comportamento.

La simulazione non basta a distinguere la coscienza

Occorre però evitare una conclusione affrettata. Costruire modelli e produrre previsioni non è una prerogativa della coscienza. Un'intera corrente delle neuroscienze contemporanee — il *predictive processing* di Karl Friston, Andy Clark e altri — sostiene che il cervello sia, a tutti i livelli, una macchina predittiva: ogni sistema percettivo e motorio genera continuamente aspettative e le confronta con i segnali in arrivo, quasi sempre in modo del tutto inconscio. I modelli interni del sistema motorio anticipano le conseguenze sensoriali dei movimenti; i sogni sono simulazioni prodotte in assenza di controllo deliberato. Se "simulare" significa generare previsioni, tutto il cervello simula, sempre.

Dire che la coscienza serve a simulare non basta dunque a caratterizzarla. Occorre precisare che cosa la simulazione cosciente abbia di diverso. Propongo tre caratteristiche.

La prima è il **distacco dal presente**. Le previsioni inconscie sono agganciate alla situazione in corso: anticipano la traiettoria della palla, la parola successiva, l'effetto immediato di un gesto. La simulazione cosciente può invece staccarsi dal presente e rappresentare situazioni remote, ipotetiche o inesistenti: un colloquio di domani, una città mai costruita, le conseguenze di una scelta morale.

La seconda è l'**integrazione globale**. La simulazione cosciente può combinare in un unico modello informazioni provenienti da sistemi specializzati diversi — percezione, memoria, linguaggio, conoscenza fisica, conoscenza sociale — che normalmente non comunicano direttamente tra loro.

La terza, forse la più importante, è la **ricorsività**. Il risultato di una simulazione può diventare a sua volta oggetto di un'ulteriore simulazione. Possiamo immaginare una decisione; poi immaginare le sue conseguenze; poi la nostra reazione a quelle conseguenze; poi la reazione degli altri alle nostre reazioni; e infine rivedere la decisione iniziale alla luce dell'intera catena. Perché questa ricorsività richieda uno spazio di accesso globale si può argomentare così: ogni passo della catena usa come input il prodotto del passo precedente, e questo prodotto deve essere disponibile a processi diversi da quello che lo ha generato — la simulazione delle reazioni altrui richiede i sistemi di cognizione sociale, quella delle conseguenze materiali richiede la conoscenza fisica, e così via. Un risultato confinato nel modulo che lo ha prodotto non potrebbe alimentare questa circolazione. La ricorsività sembra dunque presupporre proprio ciò che lo spazio cosciente offre: un luogo funzionale in cui i prodotti di ciascun processo diventano accessibili a tutti gli altri.

Il vantaggio evolutivo

La simulazione staccata dal presente costituisce un enorme vantaggio adattativo: permette di imparare senza dover sperimentare direttamente ogni possibilità. L'organismo può esplorare numerosi futuri possibili pagando soltanto il costo computazionale dell'immaginazione invece del costo biologico dell'errore. L'immaginazione non è un lusso cognitivo: è un laboratorio interno nel quale la mente sperimenta continuamente futuri possibili.

Linguaggio e simulazione

In questa prospettiva anche il linguaggio acquista una collocazione teorica precisa. Le parole non servono soltanto a comunicare: servono a costruire simulazioni. Attraverso il linguaggio possiamo rappresentare oggetti assenti, eventi futuri, possibilità ipotetiche, relazioni causali e — soprattutto — simulazioni costruite da altre persone. Leggere un romanzo significa eseguire una simulazione progettata dall'autore; studiare una teoria scientifica significa incorporare un modello della realtà elaborato da altri. Il linguaggio moltiplica così la capacità dello spazio cosciente di costruire e modificare modelli.

La decisione come arresto della simulazione

Normalmente immaginiamo che prima si pensi e poi si decida. Il processo sembra piuttosto questo: la mente costruisce simulazioni alternative, le confronta, le modifica, le combina; la decisione emerge quando una delle simulazioni viene giudicata sufficientemente soddisfacente rispetto agli obiettivi, ai vincoli e alle risorse disponibili. Herbert Simon ha chiamato *satisficing* questo criterio: non cercare l'opzione ottimale, ma fermarsi alla prima abbastanza buona. La funzione principale dello spazio cosciente non consiste dunque nel decidere: consiste nel costruire e confrontare futuri possibili. La decisione è il risultato provvisorio di questo processo — e il momento dell'arresto, come vedremo nel prossimo capitolo, è regolato in larga parte dalle emozioni.

4 Le emozioni: il sistema di controllo della simulazione

Che cosa determina il contenuto delle simulazioni? Perché la mente dedica ore a un problema e ne ignora completamente un altro? Perché in alcune situazioni immaginiamo soprattutto gli scenari peggiori, mentre in altre siamo portati a vedere opportunità e soluzioni?

La risposta proposta in questo capitolo è che le emozioni svolgono una funzione molto più ampia di quanto comunemente si pensi: non sono semplicemente una componente affettiva della vita mentale, ma il principale sistema di controllo della simulazione cosciente.

Il posto delle emozioni nell'architettura

Serve prima una precisazione, per evitare un'ambiguità. Nel primo capitolo le emozioni comparivano tra i prodotti dei programmi inconsci; qui compaiono come regolatori della simulazione cosciente. Non c'è contraddizione: sono l'uno e l'altro. Le emozioni sono generate da programmi inconsci di valutazione — ciò che la psicologia chiama *appraisal*, nella tradizione di Magda Arnold, Richard Lazarus e Klaus Scherer — e funzionano come segnali globali che modificano il funzionamento degli altri processi, inclusa la simulazione cosciente. Nella terminologia di questo saggio: sono l'output di alcuni programmi e, insieme, parametri di controllo per tutti gli altri.

Le emozioni come valutazioni

Nella concezione tradizionale emozione e ragione vengono contrapposte: da una parte il pensiero razionale, dall'altra le emozioni come fattori di disturbo. La contrapposizione è fuorviante, perché ogni emozione contiene una valutazione implicita. La paura segnala una possibile minaccia; la curiosità un'opportunità di apprendimento; la rabbia una violazione percepita; la vergogna un rischio per la propria immagine sociale. Le emozioni non descrivono il mondo: ne valutano il significato rispetto ai bisogni dell'organismo.

Le ricerche di Antonio Damasio sui pazienti con lesioni prefrontali mostrano che cosa accade quando questi segnali mancano: persone con intelligenza e memoria intatte diventano incapaci di decidere efficacemente, perse in deliberazioni interminabili su questioni banali. La ragione privata delle emozioni non diventa più pura: diventa inconcludente.

Che cosa simulare, e come

Lo spazio cosciente potrebbe costruire un numero praticamente infinito di simulazioni, ma le risorse cognitive sono limitate. Le emozioni stabiliscono le priorità: chi teme per la propria salute simulerà soprattutto scenari di malattia; chi è innamorato costruirà spontaneamente simulazioni riguardanti la persona amata. Esse rispondono implicitamente a una domanda fondamentale: **quale simulazione ha, in questo momento, il maggiore valore adattativo?**

La loro funzione, però, non termina qui: le emozioni non controllano soltanto l'argomento della simulazione, ma i parametri stessi del processo. La paura restringe il numero delle alternative considerate; l'ansia prolunga l'esplorazione degli scenari negativi; la curiosità amplia il campo delle possibilità; la rabbia rende improbabile l'esplorazione di soluzioni concilianti; l'entusiasmo alza il livello di rischio ritenuto accettabile; la tristezza favorisce valutazioni prudenti. Ampiezza della ricerca, profondità della simulazione, orizzonte temporale, rischio accettabile, peso dei diversi

obiettivi, momento dell'arresto: tutto questo è regolato emotivamente. Da questo punto di vista il ragionamento non è un processo distinto dall'emozione: è una modalità di simulazione continuamente regolata dalle emozioni.

Piacere e dolore come segnali fondamentali

È possibile chiedersi se, al di sotto della varietà delle emozioni, esista un principio più semplice. Un'ipotesi è che gran parte della vita emotiva derivi dall'evoluzione di due segnali fondamentali: il piacere, che segnala una situazione favorevole alla soddisfazione di bisogni importanti, e il dolore, che segnala il contrario. Molte emozioni sembrano interpretabili come forme specializzate e anticipatorie di questi segnali: la curiosità come piacere anticipato della scoperta, la paura come dolore anticipato di un possibile danno, l'attrazione come piacere anticipato dell'avvicinamento. Questa riduzione è naturalmente una semplificazione: le emozioni sociali, morali ed estetiche mostrano una ricchezza che richiede modelli più articolati. L'idea di fondo, però, resta plausibile: le emozioni sono il linguaggio con cui il sistema cognitivo comunica allo spazio cosciente ciò che merita investimento.

Quando termina una simulazione?

In teoria una simulazione potrebbe continuare indefinitamente: ogni scenario può generarne un altro, ogni previsione può essere raffinata, ogni decisione può essere rimessa in discussione. Perché allora, a un certo punto, smettiamo di riflettere e agiamo?

Probabilmente perché il sistema valuta che il beneficio atteso di ulteriori simulazioni sia diventato inferiore al loro costo. Anche questa valutazione è emotiva: l'urgenza, l'ansia, la fiducia, la stanchezza, la pressione sociale e la paura di sbagliare contribuiscono a determinare il momento in cui la simulazione si interrompe e una linea d'azione viene selezionata.

Decidere, in questa prospettiva, non significa scegliere tra alternative già date: significa interrompere un processo di simulazione quando ulteriori elaborazioni non sembrano più convenienti. La decisione non è un atto isolato della coscienza: è il momento in cui una simulazione viene giudicata abbastanza buona da diventare comportamento.

5 Apprendimento: come la mente modifica se stessa

Se la funzione della simulazione fosse soltanto quella di produrre la decisione presente, il suo vantaggio evolutivo sarebbe limitato: ogni situazione nuova richiederebbe di ripartire quasi da zero. La mente usa invece la simulazione anche per modificare progressivamente il proprio modo di funzionare. Ogni esperienza lascia una traccia; ogni decisione cambia, almeno in parte, il sistema che produrrà le decisioni successive.

Il ciclo dell'apprendimento

I programmi inconsci non sono immutabili: vengono continuamente modificati dall'esperienza. La simulazione produce aspettative; l'azione le mette alla prova; le differenze tra ciò che era atteso e ciò che è accaduto producono apprendimento. Quando una previsione si dimostra corretta, i processi che l'hanno prodotta si rafforzano; quando fallisce, il sistema è costretto a riorganizzarsi.

Dal punto di vista dell'apprendimento, l'errore non è qualcosa da evitare: è una delle fonti di informazione più preziose. Se il mondo si comportasse sempre esattamente come previsto, il sistema non avrebbe alcun motivo di modificare i propri modelli. Ogni sorpresa contiene un'informazione; ogni fallimento segnala che almeno una parte del modello era incompleta o inesatta.

Va notato che questo principio non è esclusivo della mente biologica: anche i moderni sistemi di intelligenza artificiale apprendono correggendo errori di previsione. La differenza non sta nel principio, ma nel modo in cui la mente lo applica: continuamente, durante il funzionamento, integrando percezione, emozione, azione e riflessione in un unico ciclo che modifica il sistema stesso che lo esegue.

Apprendere senza agire

La simulazione introduce una possibilità straordinaria: apprendere senza sperimentare direttamente le conseguenze delle proprie azioni. Un bambino impara osservando gli adulti; uno studente impara leggendo; uno scienziato impara analizzando esperimenti eseguiti da altri. Perfino immaginare una situazione modifica il sistema cognitivo, creando nuove connessioni tra informazioni già presenti.

L'apprendimento sociale

La maggior parte di ciò che sappiamo non deriva dalla nostra esperienza diretta, ma dall'esperienza degli altri. Ogni cultura può essere interpretata come un immenso archivio di simulazioni accumulate nel corso delle generazioni: i racconti trasmettono esperienze, le teorie scientifiche trasmettono modelli, le tradizioni trasmettono strategie. L'educazione permette a ogni individuo di incorporare una quantità enorme di simulazioni che non ha mai vissuto personalmente. La mente non modifica se stessa soltanto attraverso ciò che sperimenta: si modifica anche attraverso ciò che riesce a condividere.

Un sistema che evolve

Considerata nel suo insieme, l'architettura descritta forma un ciclo continuo: i programmi inconsci interpretano la situazione, le emozioni ne valutano l'importanza, l'attenzione seleziona le informazioni rilevanti, lo spazio cosciente costruisce simulazioni, la simulazione orienta l'azione,

l'esperienza modifica i programmi. La coscienza appare così come uno dei principali strumenti con cui il sistema cognitivo accelera il proprio apprendimento: la simulazione esplora molte possibilità, l'esperienza seleziona le più efficaci, l'apprendimento ne incorpora stabilmente i risultati. La mente diventa capace non soltanto di adattarsi al mondo, ma di aumentare progressivamente la propria capacità di adattamento.

6 I limiti della simulazione

La capacità di simulare rappresenta uno straordinario vantaggio evolutivo, ma non è priva di costi. La simulazione non produce conoscenza certa: produce ipotesi. Ogni modello mentale è una ricostruzione della realtà, non la realtà stessa. Per questo la stessa capacità che rende possibili la pianificazione, la creatività e il ragionamento astratto rende possibili anche l'errore, l'illusione e l'autoinganno.

Approssimazione e semplificazione

Ogni simulazione nasce da informazioni necessariamente incomplete ed elimina una quantità enorme di dettagli: quando immaginiamo una persona non rappresentiamo ogni capello, quando pianifichiamo una conversazione non simuliamo ogni possibile reazione dell'interlocutore. Questa approssimazione non è un difetto: attendere informazioni complete significherebbe, nella maggior parte dei casi, non agire mai. Un modello imperfetto disponibile subito è spesso più utile di un modello perfetto disponibile troppo tardi. Ma la semplificazione comporta il rischio di trascurare proprio l'informazione che, in quella particolare situazione, avrebbe fatto la differenza. E poiché il futuro non è mai completamente prevedibile, la simulazione non serve a prevederlo con precisione: serve ad aumentare la probabilità di affrontarlo con successo.

Quando il modello diventa più importante della realtà

La simulazione introduce anche un rischio più sottile. Poiché la mente opera prevalentemente sui propri modelli, questi possono acquistare un'importanza psicologica maggiore della realtà che dovrebbero rappresentare. Una persona convinta di essere destinata al fallimento interpreterà ogni episodio come conferma, ignorerà le informazioni contrarie, costruirà simulazioni sempre più coerenti con il modello iniziale. Il sistema tende ad autoalimentarsi: la rappresentazione diventa progressivamente più stabile, ma non necessariamente più accurata. Le emozioni contribuiscono a entrambi i lati del fenomeno: la paura concentra l'attenzione sulle minacce, la rabbia privilegia le interpretazioni conflittuali, l'entusiasmo sottovaluta i rischi. Se le priorità emotive diventano troppo rigide, la simulazione perde flessibilità.

Le convinzioni come strumenti adattativi

Spesso consideriamo le convinzioni come semplici descrizioni della realtà. Dal punto di vista funzionale svolgono un ruolo diverso: una convinzione è un modello sufficientemente stabile da poter guidare rapidamente il comportamento. La sua funzione principale non è essere perfettamente vera: è ridurre il costo cognitivo delle decisioni. Sarebbe impossibile riesaminare da zero ogni convinzione prima di ogni scelta; per questo la mente stabilizza molti dei propri modelli. Il prezzo è che i modelli molto stabili diventano difficili da correggere anche quando l'esperienza li smentisce.

L'autoinganno

L'autoinganno viene tradizionalmente interpretato come un'anomalia. In questa prospettiva appare piuttosto un effetto collaterale dell'architettura. Quando una rappresentazione protegge obiettivi considerati particolarmente importanti, il sistema può attribuirle un valore emotivo tale da renderne difficile la revisione: la simulazione continua a produrre scenari coerenti con quel modello, mentre le informazioni incompatibili vengono progressivamente trascurate. Non serve alcun desiderio cosciente di mentire a se stessi: basta il modo in cui emozioni, attenzione e simulazione interagiscono.

Errori inevitabili

Non esiste una mente capace di simulare senza commettere errori. Ridurre gli errori richiede modelli più sofisticati; modelli più sofisticati richiedono maggiori risorse; ogni organismo deve trovare un compromesso tra accuratezza e velocità. L'evoluzione non seleziona il sistema più preciso: seleziona quello sufficientemente efficace da garantire sopravvivenza e riproduzione. La forza della mente non risiede nell'infallibilità, ma nella possibilità di correggere progressivamente i propri modelli attraverso l'esperienza — la stessa capacità che rende possibili la conoscenza scientifica, il progresso culturale e lo sviluppo personale.

7 Rapporti con le teorie esistenti

Una teoria che non dichiara i propri debiti si espone a due obiezioni: ignorare la letteratura, oppure rivendicare un'originalità che non ha. Questo capitolo dichiara che cosa la proposta qui avanzata riprende da teorie esistenti e che cosa vi aggiunge.

Il workspace globale. L'idea dello spazio cosciente come luogo funzionale in cui informazioni selezionate diventano globalmente disponibili corrisponde in larga parte alla *Global Workspace Theory* di Bernard Baars, sviluppata sul piano neurobiologico da Stanislas Dehaene e Jean-Pierre Changeux. Da questa tradizione il saggio riprende l'architettura di fondo: molti processi specializzati inconsci, una competizione per l'accesso, uno spazio di lavoro a capacità limitata i cui contenuti vengono diffusi all'intero sistema.

La società della mente. L'immagine della mente come popolazione di agenti specializzati riprende la *Society of Mind* di Marvin Minsky, inclusa la strategia contro l'homunculus: la scomposizione in agenti sempre più semplici.

Il cervello predittivo. L'idea che la mente costruisca e corregga continuamente modelli del mondo, apprendendo dagli errori di previsione, è il nucleo del *predictive processing* di Karl Friston, Andy Clark e Jakob Hohwy. Da qui il saggio riprende il ciclo aspettativa-errore-aggiornamento descritto nel capitolo 5.

Emozioni e decisione. La lettura delle emozioni come valutazioni riprende la tradizione dell'*appraisal* (Arnold, Lazarus, Scherer); il loro ruolo indispensabile nella decisione riprende l'ipotesi del marcatore somatico di Antonio Damasio; il criterio di arresto della deliberazione riprende il *satisficing* di Herbert Simon e la letteratura sulla razionalità limitata.

Accesso e fenomenologia. La distinzione che delimita l'oggetto della teoria è quella di Ned Block tra coscienza d'accesso e coscienza fenomenica.

Che cosa aggiunge, allora, questo saggio? Principalmente tre cose. Primo, una tesi sul rapporto tra emozioni e spazio cosciente che nelle formulazioni classiche della Global Workspace Theory resta marginale: le emozioni non sono soltanto contenuti che accedono allo spazio cosciente, né soltanto criteri di selezione, ma regolatori dei parametri stessi della simulazione — ampiezza, profondità, orizzonte temporale, soglia di arresto. Secondo, una caratterizzazione della funzione dello spazio cosciente centrata sulla simulazione ricorsiva, con un argomento sul perché la ricorsione richieda accesso globale. Terzo, una sintesi: la proposta di leggere workspace, appraisal, predizione e razionalità limitata come aspetti di un'unica architettura, presentata in una forma accessibile anche a chi non frequenta la letteratura specialistica.

Si tratta di contributi di integrazione più che di scoperta, e mi sembra corretto dichiararlo. Ma nelle scienze della mente, dove le teorie tendono a svilupparsi in tradizioni separate che spesso si ignorano a vicenda, l'integrazione non è un contributo trascurabile.

8 Una teoria architetture della mente

Abbiamo affrontato una domanda apparentemente semplice: **perché esiste la coscienza?** Per rispondere è stato necessario allargare progressivamente il campo di osservazione, perché la coscienza non può essere compresa isolatamente: può essere compresa soltanto all'interno dell'architettura complessiva della mente. L'idea centrale di queste pagine è che la mente sia un sistema dinamico di processi che operano continuamente gli uni sugli altri, e che la coscienza sia uno di questi processi. Non il più importante, non il centro di comando: uno dei processi.

Il ciclo della mente

L'intero funzionamento del sistema può essere descritto come un ciclo. L'ambiente produce stimoli; i programmi inconsci li elaborano; le emozioni ne valutano il significato; l'attenzione seleziona ciò che merita ulteriori risorse; lo spazio cosciente costruisce modelli e simulazioni; la simulazione orienta il comportamento; l'azione modifica l'ambiente; le conseguenze, confrontate con le aspettative, aggiornano i programmi. Il ciclo ricomincia. La mente non è un sistema statico: usa ogni interazione con il mondo per trasformare se stessa. Ciò che l'evoluzione biologica realizza sul patrimonio genetico in tempi lunghissimi, la mente lo realizza su se stessa nel corso di una vita: non evolve soltanto la specie, evolve continuamente anche il singolo individuo.

La risposta alla domanda iniziale

Perché esiste, dunque, la coscienza — nel senso, delimitato nel primo capitolo, di coscienza d'accesso? Perché permette alla mente di costruire, confrontare e modificare modelli della realtà con una flessibilità che i soli automatismi inconsci difficilmente potrebbero raggiungere. La simulazione staccata dal presente consente di esplorare possibilità senza viverle direttamente; la ricorsività consente di elaborare catene di conseguenze; il linguaggio consente di incorporare simulazioni costruite da altri; l'apprendimento incorpora stabilmente i risultati. Il vantaggio evolutivo della coscienza non consiste semplicemente nel prendere decisioni migliori: consiste nell'accelerare l'evoluzione cognitiva dell'individuo.

Se invece la domanda riguarda la coscienza fenomenica — perché tutto questo sia accompagnato da un'esperienza soggettiva — questa teoria non ha una risposta, e non pretende di averla.

Alcune conseguenze

Se questa teoria è corretta, molte questioni tradizionali della psicologia possono essere reinterpretate. Le emozioni non sono un ostacolo alla razionalità, ma il sistema che assegna priorità ai processi cognitivi. La memoria non è un archivio, ma un repertorio di modelli costruiti nel corso dell'esperienza. L'immaginazione non è evasione dalla realtà, ma esplorazione della realtà possibile. Il ragionamento non è una facoltà autonoma, ma una forma di simulazione guidata da modelli e regolata dalle emozioni.

Che cosa potrebbe smentire questa teoria

Una teoria che spiega tutto e non vieta nulla non spiega niente. Vale quindi la pena indicare almeno alcune previsioni che la teoria rende plausibili e che l'esperienza potrebbe smentire.

Primo: se le emozioni regolano i parametri della simulazione, indurre sperimentalmente stati emotivi diversi dovrebbe modificare in modo misurabile l'ampiezza e la profondità della deliberazione. Ad esempio, il numero di alternative generate in un compito di pianificazione

dovrebbe ridursi sotto ansia indotta e aumentare sotto curiosità indotta. Risultati sistematicamente contrari — parametri deliberativi indifferenti allo stato emotivo — costituirebbero un problema serio per il capitolo 4.

Secondo: se la simulazione ricorsiva richiede accesso globale, i compiti che richiedono catene di inferenze — usare il risultato di una previsione come input di una successiva — non dovrebbero essere eseguibili su informazioni presentate in modo subliminale, mentre singole previsioni di primo livello possono esserlo, come la letteratura sul priming suggerisce. Una dimostrazione robusta di ragionamento ricorsivo a più passi interamente inconscio smentirebbe il capitolo 3.

Terzo: se la decisione è un arresto regolato emotivamente e non il completamento di un'analisi, il momento in cui le persone smettono di deliberare dovrebbe dipendere dall'urgenza e dai costi percepiti più che dalla completezza delle informazioni raccolte; e individui con segnalazione emotiva anomala — come nei quadri clinici descritti da Damasio, o nell'alessitimia — dovrebbero mostrare anomalie specifiche dell'arresto: deliberazioni interminabili oppure troncate. Se la deliberazione si arrestasse tipicamente al completamento dell'evidenza, indipendentemente dallo stato emotivo, la teoria della decisione qui proposta risulterebbe falsa.

Sono previsioni formulate a grana grossa, e trasformarle in esperimenti rigorosi richiederebbe un lavoro che va oltre i limiti di questo saggio. Ma indicano la direzione in cui la teoria accetta di essere giudicata.

Conclusione

Abbiamo iniziato chiedendoci quale sia la funzione evolutiva della coscienza. La risposta proposta è diversa da quella tradizionale: la coscienza non è il luogo in cui nasce la mente; è uno degli strumenti attraverso i quali la mente costruisce il proprio futuro. Attraverso lo spazio cosciente il sistema cognitivo integra informazioni provenienti da processi diversi, costruisce modelli della realtà, simula possibilità alternative, confronta le conseguenze delle proprie azioni e usa l'esperienza per modificare progressivamente se stesso. Forse è proprio questa la caratteristica più notevole della mente umana: non il fatto di essere cosciente, ma il fatto di poter usare la coscienza per cambiare se stessa.