

ARTICOLI

Cognizione logica e modelli mentali

Alberto Binazzi

1. *Introduzione*

I “modelli mentali” sono stati introdotti dallo psicologo inglese Philip Johnson Laird come strutture cognitive semi-analogiche capaci di rappresentare specifici stati di cose e relativi processi inferenziali. La teoria dei modelli mentali¹ sviluppata da Johnson-Laird è una teoria della mente che si pone l’ambizioso progetto di offrire una visione unitaria dei processi mentali pur rimanendo, in gran parte, all’interno del quadro del cognitivismo classico. Nello studio della competenza deduttiva la teoria ha dimostrato capacità predittiva grazie allo sviluppo di diversi programmi per calcolatore in grado di simulare le effettive prestazioni di soggetti alle prese con problemi di logica.

Johnson-Laird e i successivi sostenitori della teoria dei modelli mentali attribuiscono alle procedure costruttive nell’immaginazione e nell’intelligenza spaziale una notevole importanza per lo sviluppo e la piena acquisizione della competenza deduttiva. Questo può far pensare a una convergenza con l’approccio intuizionistico in matematica.

Brouwer, infatti, concepiva la matematica come un’attività mentale costruttiva ed essenzialmente a-linguistica, il linguaggio mantenendo soltanto una funzione di comunicazione delle esperienze matematiche. In opposizione alle teorie formaliste, tale separazione fra costruzioni mentali e linguaggio poteva avere dei vantaggi per l’attività matematica stessa: se questa si caratterizzava come essenzialmente a-linguistica, i problemi incontrati nella ricerca volta a identificare un linguaggio formale canonico avevano minore importanza (l’assenza di contraddizioni in un sistema formale privo di contraddizioni non potendo essere considerata come criterio di esistenza matematica). La fiducia dei formalisti nei confronti dei ‘poteri magici’ del linguaggio era collegata da Brouwer alla loro fede nella validità della logica classica, e, più

¹ Si veda Johnson-Laird (1996). Per un’aggiornata introduzione alla teoria è possibile consultare il sito: <http://webscript.princeton.edu/~mentmod/models.php/> dove sono presentati modelli computazionali (in LISP) per la logica proposizionale, il ragionamento modale, il ragionamento temporale, controfattuale e sillogistico.

in particolare, nel principio del terzo escluso. Ma questa fede non poteva essere condivisa, il principio del terzo escluso essendo affidabile solo per domini finiti o decidibili². La matematica doveva, kantianamente, essere fondata sull'intuizione, e, in tal senso, il matematico olandese poteva richiamarsi alle riflessioni psicologico-filosofiche di Poincaré.

La costruibilità degli enti logico-matematici e la concezione della matematica come attività della mente di natura a-linguistica sono state solo in parte recuperate da Johnson-Laird: nel quadro di una filosofia cognitiva della logica, la costruzione delle strutture denominate "modelli" mentali" diventa oggetto di studio da parte dello psicologo e la competenza deduttiva da analizzare è quella in ambiente classico.

Mettendo per ora da parte questa limitazione, possiamo notare che l'impulso maggiore allo sviluppo della teoria è derivato dalla necessità di spiegare il tipo di processi mentali implicati nel ragionamento deduttivo evitando di cadere nella circolarità in cui cadono le proposte basate su teorie sintatticistiche, formali o semiformali, della mente umana:

Unlike theories that stress that reasoning is verbal process (Polk and Newell, 1992) or is governed by formal rules (Rips, 1983) mental models theory assumes that a major component of reasoning is nonverbal, that is the construction of models with a structure corresponding to the structure of situations³.

Tale assunto ha condotto a indagare sperimentalmente l'esistenza di uno stretto rapporto tra competenza inferenziale e processi immaginativi (visual imagery) con riferimento al ruolo della corteccia pre-frontale nella ricerca dei controesempi per invalidare un ragionamento. Sebbene siano stati condotti diversi esperimenti che suffragherebbero l'esistenza del rapporto, la questione è ancora aperta: i dati sperimentali offerti dai neuroscienziati sono discordanti e negli attuali manuali di neuropsicologia si fa ancora fatica a rintracciare una sintesi soddisfacente.

L'aspetto diagrammatico-spaziale del ragionamento, che Johnson-Laird ha il merito di incorporare in una teoria della mente, era già stato evidenziato nei diagrammi di Eulero-Venn e si ritrova nei recenti sviluppi della logica categoriale, ma insieme all'aspetto diagrammatico-spaziale, il contributo forse più significativo offerto dalla teoria dei modelli mentali – intesa come teoria cognitiva della logica (classica e per domini finiti) – è che il sistema non accede a tutti i modelli disponibili, ovvero, vi è un processo di "forcing" che seleziona, tra i possibili modelli, un sottoinsieme caratteristico di questi. Più i nostri modelli sono ricchi e numerosi, più il linguaggio rifletterà questa varietà semantica. L'idea è che per argomentare su qualcosa si deve avere come sfondo un modello mentale di mondo, dinamico e rappresentabile in strutture di alto livello, rese possibili da ben determinate architetture neurali.

² Cfr. Bozzi, Mangione (2001).

³ Johnson-Laird (1996) p.1005.

2. Architetture e processi mentali

Pei Wang (2004) ha proposto di considerare tre tipologie di sistemi di ragionamento: 1) sistemi assiomatici “puri” di cui un tipico esemplare è rappresentato dal sistema formale hilbertiano; 2) sistemi come quelli della logica controfattuale o della logica fuzzy; 3) sistemi non-assiomatici che fanno dell’insufficiente conoscenza di un dato dominio e delle limitate risorse computazionali il loro punto di forza. Wang afferma che, anche se i sistemi assiomatici puri sono stati utilizzati con successo in matematica, si sono rivelati inadeguati per spiegare “molti aspetti dell’intelligenza e della cognizione”. I sistemi di tipo 3) cui Wang è interessato hanno una limitata potenza computazionale, evolvono nel tempo, fanno riferimento a un dominio aperto e non impongono nessun vincolo al tipo di esperienze/input aggiuntivi, a condizione che siano rappresentabili per mezzo di un’adeguata interfaccia linguistica. Si assume, infatti, che esista una traducibilità in forma proposizionale del formato (e del contenuto) di esperienza acquisita per mezzo di varie interfacce (motorie, visive, tattili, ecc.). Ma quali devono essere i formati delle esperienze per essere traducibili in forma linguistica? E come sfruttare tale traducibilità nel ragionamento?

Johnson-Laird ha il merito di aver messo in luce come la nostra capacità inferenziale sia soggetta a restrizioni dovute a: struttura della memoria (essendo rilevante il tipo di formato in cui è immagazzinata un’informazione), maggiore o minore capacità immaginativa, conoscenze di tipo semantico, limiti della memoria di lavoro. Perciò, un sistema di tipo 3) sembra avere molto in comune con un “modello mentale”. Ciò che caratterizza la proposta di Johnson-Laird è l’idea che gli esseri umani, oltre a possedere un’incompleta conoscenza del dominio, siano soggetti a restrizioni dovute ad alcune peculiarità dell’architettura cognitiva, che ci mette in grado di reclutare nuovi processori o moduli qualora uno o più di questi siano occupati, possiede un magazzino estendibile di memoria e recupera esaustivamente le computazioni intermedie, mentre dimostra una particolare sensibilità al contenuto delle conoscenze pregresse (che può confliggere con nuove informazioni) e riduce al minimo il carico della memoria di lavoro (elemento, quest’ultimo, responsabile della incompletezza strutturale dei modelli mentali). La “razionalità” che emerge dalla teoria sarà dunque quella propria di un sistema cognitivo che si trova a operare in situazioni di insufficienti conoscenze e limitate risorse computazionali.

Se i critici della teoria dei modelli mentali hanno ripetutamente messo in dubbio che la proposta di Johnson-Laird sia libera da un’impostazione assiomatica, un elemento raramente preso in considerazione (ma confermato da numerosi esperimenti) è che la mente utilizza diverse strategie anche di fronte a problemi di logica e il sistema che costruisce e manipola i modelli mentali incorpora un insieme di strategie già implementabili in diversi programmi per calcolatore. Se consideriamo i programmi che utilizzano i modelli mentali in problemi logici che fanno riferimento all’ordine temporale,

esistono due diverse strategie di risoluzione dei problemi: i) i vari modelli vengono costruiti uno alla volta, indipendentemente l'uno dall'altro, ii) le possibilità alternative sono rappresentate in un singolo modello marcando gli elementi nel modello suscettibili di assumere differenti posizioni (nel tempo) all'interno del modello.

In sistemi come quelli descritti dalla teoria dei modelli mentali il concetto di validità o correttezza di una deduzione acquista un'accezione particolare, perché potremmo sostenere che un tale sistema non sia in grado di fare sempre deduzioni valide così come potremmo pensare ad altre nozioni di validità, che tengano conto del tipo di logica (che potrebbe non essere la logica classica) e dell'architettura utilizzata. Inoltre, è semplicemente un fatto che gli esseri umani non sono sempre capaci di dedurre validamente: non siamo astratti sistemi formali. Infine, se i nostri apparati di alto livello sfruttano la versatilità dei modelli mentali, questi devono essere concepiti come rappresentazioni schematiche che prescindono da almeno qualche aspetto (del contenuto) dell'esperienza percettiva, senso-motoria, spaziale; ovvero, la grana di ciò che è rappresentabile per mezzo di un modello, dipenderà dal tipo di esperienza, ma, a causa di valori di default, un modello non rappresenterà mai tutta l'informazione relativa a un particolare dominio. In vista di queste limitazioni architettoniche, è risultato utile allo scienziato cognitivo considerare i programmi per calcolatore come un'opportuna finestra per la comprensione dei nostri processi mentali:

The program's search for counterexamples works because it has access to the representations of the meanings of the premises. Without these representations, if the program were to change a model, it would have no way to check whether the result was still a model of the premises. Any inferential system that constructs alternative models therefore needs an independent record of the premises. It must either have a memory for their meanings, or be able to return to each premises in order to re-interpret it⁴.

3. *Logica e modelli mentali*

È istruttivo domandarsi in quale modo un sistema artificiale equipaggiato con 'modelli mentali' sia in grado di riconoscere che una formula come $\psi \rightarrow (\varphi \rightarrow \psi)$ – assioma dell'a-fortiori – è una tautologia. Il sistema parte dal condizionale tra parentesi ($\varphi \rightarrow \psi$). Ogni possibile combinazione di φ e $\neg\varphi$ con ψ e $\neg\psi$ corrisponderà a un modello mentale, che il sistema procede poi a integrare con l'antecedente ψ . I 'modelli' contenenti ψ e $\neg\psi$ vengono eliminati, ammettendo invece i modelli che corrispondono al caso in cui l'antecedente è falso e il conseguente è vero e al caso in cui sono entrambi falsi. A questo punto, il sistema elimina le doppie occorrenze (di φ , $\neg\varphi$, ψ , $\neg\psi$) ove presenti, e verifica che i 'modelli' restanti soddisfano la formula.

⁴ Johnson-Laird (2006), p.14.

Secondo Johnson-Laird, la stessa procedura sarebbe eseguita dal dispositivo computazionale umano. Se l'importanza di rappresentare la procedura di ragionamento a partire dai dati linguistici può apparire scontata, non scontata è l'idea di sviluppare modelli al calcolatore che si servano di opportune rappresentazioni di tipo analogico per trarre inferenze. In questa stessa linea si potrebbe ricordare, accanto al programma sviluppato da Johnson-Laird per la logica proposizionale (classica), Tarski's World, un ambiente sviluppato dai matematici John Barwise e John Etchemendy per la logica dei predicati, che riproduce un minimondo composto da oggetti tridimensionali come cubi, sfere, parallelepipedi disposti e mossi su una scacchiera. Tarski's World mostra come la manipolazione, visivamente accessibile, degli stati del dominio risulti essere un utile ausilio per il controllo delle inferenze e il ragionamento deduttivo possa trarre beneficio dalla costruzione di determinati minimondi. Ma un programma di calcolatore che incorpora un isomorfismo tra il piano sintattico e quello semantico difficilmente può corrispondere ai dispositivi rappresentazionali umani.

4. *Menti cartesiane allo specchio*

Johnson-Laird recupera l'idea di modello come struttura matematica in grado di interpretare e rendere veri gli assiomi di una teoria formale: i modelli mentali non rappresentano, in prima battuta, gli stati di cose falsi ("principle of truth"). Ora, se nella comprensione di un testo, possiamo accettare l'idea che la costruzione di un modello mentale o di un insieme di modelli mentali equivalga a cogliere il senso del testo, senso e verità sono connessi. Ma che cosa ci mette in grado di attribuire un valore di verità a una data proposizione? Il teorico dei modelli mentali risponde: i sistemi percettivi. Dunque, incapsuliamo spontaneamente un modello mentale in un modello percettivo di uno stato di cose, e non rappresentiamo situazioni o eventi in grado di falsificare le nostre conclusioni provvisorie.

Ci sono due elementi da combinare insieme. A) È arduo capire come si possa attribuire una reale competenza deduttiva alla mente umana se questa non accede a funzioni di verità e non combina le sue rappresentazioni in accordo con principi logici, mentre il teorico dei modelli mentali non considera la logica parte della nostra architettura: si limita a postulare un meccanismo che interpreta le rappresentazioni mentali basate su enunciati (o meglio, su ciò che resta di un enunciato una volta 'mentalizzato')⁵. B) Se quello che caratterizza la mente umana, e i sistemi cognitivi artificiali embodied, è una semantica

⁵ I dispositivi cognitivi non assiomatici si caratterizzano per una proprietà estranea ai sistemi formali classici: i passi inferenziali possono cambiare nel tempo e sono influenzati da contenuto concettuale, credenze, memoria, variabilità dell'estensione della memoria di lavoro – fattori difficilmente prevedibili. Quando un'inferenza umana termina, non è perché il dispositivo ha dedotto una data proposizione, ma perché non ha altre risorse computazionali da utilizzare.

fondata sull'esperienza (experience-grounded semantics), è un dato di fatto che non facciamo esperienza di funzioni e tavole di verità come vediamo le mele sul tavolo o i quadri appesi al muro, né afferriamo una funzione di verità come afferriamo una palla. Se inoltre i modelli mentali sono ciò che adoperiamo per ragionare, è arduo capire come avremmo potuto sviluppare le varie logiche che oggi conosciamo senza avere esperito un tipo corrispondente di realtà. Ma ammesso che tra mente, logica e realtà ci sia un qualche rapporto di ortogonalità, in qualunque modo modifichiamo il rapporto dobbiamo farlo per mezzo di un punto che appartiene a tutte e tre.

I risultati nel Wason's selection task possono allora ricevere un'interpretazione differente. Gli sperimentatori desiderano spiegare perché, di fronte a situazioni sperimentali progettate per testare le competenze deduttive umane, i soggetti (i loro processi mentali) non seguano regole formali di inferenza. Infatti, di fronte alle carte E, K, 4, 7, e alla regola "se una carta ha una vocale su un lato, allora ha un numero pari sull'altro lato", la maggior parte dei soggetti sperimentali cui sia richiesto di falsificare la regola girando al massimo due carte tra quelle disponibili sceglie o solo la carta E o le carte E e 4, mentre la risposta corretta consiste nella scelta delle carte E e 7. Infatti, se formalizziamo la regola per mezzo della logica predicativa del prim'ordine, ottenendo (per ogni x) (vocale (x) \rightarrow pari (x)), tale proposizione è falsa se esiste una carta che soddisfi (vocale (x)) e \neg pari (x), altrimenti risulta vera; dunque, occorre girare la carta E (vocale) e la carta 7 (\neg pari).

Un'interpretazione per mezzo dei modelli mentali ci dà (come nel caso dell'assioma dell'a fortiori)

Vocale	Pari
\neg Vocale	Pari
\neg Vocale	\neg Pari

e Wason e Johnson-Laird ne deducono che i soggetti non ragionano in accordo a regole formali di inferenza: perché mai dovrebbero?

In realtà, dietro all'analisi dei dati sperimentali vi è la tacita convinzione che esista una sola teoria normativa della razionalità, rispecchiabile nella logica classica: ma questo è esattamente un cognitive bias analogo a quello osservato nel Wason's selection task. Esistono altri sistemi di ragionamento, come NARS, per esempio, che offrono nuove cornici interpretative ai risultati sperimentali sulla competenza deduttiva umana⁶, ma se di fronte a questi risultati possiamo supporre che le architetture computazionali dei soggetti sperimentali siano non-assiomatiche, possiamo anche supporre che seguano una logica non classica che tiene conto di limitate di risorse logico-computazionali e di incompletezza delle informazioni. Si pensi, a questo proposito, alla logica lineare, che nell'interpretazione di connettivi come l'implicazione e la congiunzione presenta analogie con il metodo dei modelli mentali.

⁶ Cfr. Wang (2001) e www.cis.temple.edu/~pwang/

Gli sforzi dei sostenitori della teoria dei modelli mentali sono stati indirizzati a confutare sperimentalmente chi nega l'esistenza della logica mentale. Ma, appunto, quale logica, e per di più, mentale? Lo sviluppo di numerose logiche non classiche ha arricchito la nostra idea della logica, tanto che è molto difficile accordare un ruolo privilegiato a una logica, fino a definirla come la logica interna della mente. Scoprendo che gli esseri umani non organizzano i loro pensieri per mezzo di una qualche logica particolare e che manifestano comportamenti poco 'razionali' (se non si tiene conto delle limitate risorse computazionali della mente umana), gli scienziati cognitivi hanno dato ragione a Frege o hanno confutato il suo argomento contro il psicologismo?

I risultati sperimentali hanno dato impulso alla ricerca di sistemi logici più fedeli alle reali risorse computazionali umane. Sono state studiate logiche compatibili con il ragionamento comune (common sense reasoning): si pensi, per esempio, alle applicazioni in IA della logica lineare o delle logiche non monotoniche. Chi si pone l'obiettivo di studiare la mente umana, non può ignorare la logica, ma come la 'topologia intuitiva' non si lascia esprimere nella teoria degli insiemi, così non possiamo pensare di descrivere esaustivamente per mezzo della logica classica la varietà dei processi mentali coinvolti nel ragionamento umano.

5. *Esseri booleani e automi cartesiani*

Dopo aver mosso queste critiche, torniamo ora al metodo algoritmico per la logica proposizionale sviluppato da Johnson-Laird per evidenziarne i vantaggi. Per estrarre i modelli mentali di una formula, il nostro automa logico va alla ricerca delle istruzioni necessarie per costruire e combinare i modelli. Dopo averle trovate, costruisce i modelli secondo il Principle of truth. Il programma combina modelli mentali in accordo alle regole codificate dalla semantica dei connettivi, cioè, le funzioni selezionano gli oggetti da combinare in base ai loro valori di verità. Nell'uomo questo non sempre accade. Perché non accade? Forse dipende dal tipo di trasformazioni effettuabili sui modelli? Ci sono vincoli topologici su queste trasformazioni e, se sì, quali sono?

Da un punto di vista topologico, una ciambella e una tazza da caffè sono equivalenti (omeomorfe): ciascuna può essere deformata nell'altra grazie a una trasformazione continua, senza tagli o buchi. Lo sono anche i modelli mentali di tali oggetti? Non sembra facile modificare modelli mentali come fossero oggetti di materiale elastico. Inoltre, non è possibile tenere a mente tutte le fasi delle trasformazioni di un modello in un altro e viceversa. Nel mondo reale, gli oggetti deformabili con continuità non sono molti e le operazioni cognitive sulle rappresentazioni mentali degli oggetti hanno a che vedere con informazioni sul materiale che compone gli oggetti, con informazioni sullo spazio che 'contiene' gli oggetti, con la geometria dei nostri movimenti... Di tutto questo ci serviamo per ragionare sugli oggetti di cui parliamo.

Supponiamo ora di presentare a un campione di soggetti sperimentali il seguente argomento: Se inserisco la chiave nella porta e ruoto la chiave in un

certo modo, la porta si apre. La porta non si apre. Cosa dedurre? Molti inferiranno che la serratura è bloccata, o che la chiave non è quella giusta, o che la porta è sbarrata dall'interno, ma praticamente nessuno risponde: Non ho inserito la chiave nella porta oppure Non ho ruotato la chiave in un certo modo – congiuntamente, Ho inserito la chiave e non l'ho ruotata in un certo modo oppure Non ho inserito la chiave e ho ruotato la chiave in un certo modo.

Se si tratta di aprire una porta, quale senso ha ruotare una chiave senza averla inserita nella serratura? I nostri modelli mentali devono confrontarsi con precisi vincoli topologici (forma degli oggetti e forma dei nostri corpi, superfici tangenti, buchi, ecc.) che organizzano il nostro stesso 'mentalizzare' (e 'verbalizzare'). Ruotare una chiave senza inserirla non ha senso per quella che è la funzione delle chiavi, per come sono fatte le porte e per come siamo fatti noi. Né possiamo invertire le opportune sequenze cinematiche nel modello. Solo se facciamo astrazione da tutto ciò, applichiamo il *modus tollens*.

Possiamo effettuare tutte queste operazioni anche su input di tipo linguistico? La struttura manifesta del linguaggio è di tipo sequenziale e le frasi sono prodotte una parola alla volta nel tempo. Se ci affidassimo solo alla *sin-tassi*, non potremmo ragionare come ragioniamo. Ma perché dovrebbe esistere qualcosa nella mente che si uniformi alla *sintassi*? I modelli mentali hanno il vantaggio di implementare determinate operazioni logiche senza incorporare nessun principio logico di natura puramente sintattica, mentre i modelli logici devono essere tutti quanti costruiti per intero. Dagli esperimenti, invece, è risultato che i soggetti costruiscono (di solito) un solo modello e a partire da esso traggono le loro conclusioni. I sostenitori dei modelli mentali spiegano questo fatto dicendo che non adoperiamo la logica per ragionare, così come non adoperiamo il calcolo delle probabilità per prendere le nostre decisioni. Se di fronte a problemi di logica i nostri soggetti non applicano regole logiche, ne dobbiamo concludere che la logica matematica non corrisponde ai processi deduttivi.

Anche la logica matematica sfrutta determinate risorse cognitive, ma non rappresenta il codice genetico del pensiero e non si comprende perché dovrebbe rappresentarlo. La teoria dei modelli mentali ha il merito di aver cercato di descrivere il tipo di processi (mentali) coinvolti nel ragionamento deduttivo attraverso programmi per calcolatore. La teoria è stata estesa anche al ragionamento probabilistico, causale, controfattuale, induttivo: il ragionamento deduttivo non copre, infatti, tutte le nostre capacità inferenziali. In aggiunta, i nostri processi deduttivi seguono meccanismi automatici a volte imprevedibili persino per lo sperimentatore, per effetto di interferenze molteplici.

Il fatto stesso che tali processi siano influenzabili da elementi percettivi, o immaginativi, la dice lunga sulla ricchezza, ancora non adeguatamente esplorata, delle operazioni mentali implicate nel ragionamento. Il contenuto delle nostre percezioni può avere una 'grana' più fine dei nostri formati proposizionali, certe proprietà di tipo spaziale possono non essere del tutto verbalizzabili (Pylyshyn, 2003) e questo potrebbe valere anche per il ragionamento logico-matematico. Eppure, quando ragioniamo, non abbiamo la sensazione che parte del nostro contenuto mentale resti al di fuori della coscienza, e nem-

meno siamo in grado di distinguere quali parti delle nostre rappresentazioni si accordino e quali non si accordino con i principi logici.

Andando oltre il livello proposizionale, la struttura dei modelli mentali è stata arricchita per rappresentare formule predicative. Secondo Yang, Bringsjord (2001), è possibile rappresentare adeguatamente una formula predicativa come Per tutti gli x (IF Px THEN (Fx OR Gx)) all'interno di un modello mentale del tipo [P]F/G.

6. *Forme logiche, strutture, rappresentazioni mentali*

Come è possibile trasferire una particolare struttura logica all'interno di un modello mentale, inteso come una struttura 3D originata da input di tipo percettivo, spaziale o motorio? Questa domanda è la traduzione cognitivista dell'antico problema della natura della forma logica.

Sembra alquanto bizzarra l'idea che, se siamo in grado di introdurre simboli logici in un modello mentale, questo, automaticamente, acquisti una particolare struttura logica che prima era del tutto assente. In matematica, infatti, si conferisce struttura logica a un insieme (non vuoto) quando i suoi elementi si comportano conformandosi a determinate proprietà. Per esempio, un insieme non vuoto su cui è definita un'operazione binaria qualsiasi costituisce un gruppo (acquista la struttura logica particolare di gruppo rispetto a questa operazione) quando i suoi elementi generici si uniformano alle proprietà associative, dell'esistenza dell'elemento neutro e dell'elemento inverso.

Vi sono numerosissimi esempi di gruppi in matematica: l'insieme degli interi rispetto alla somma, il nostro sistema temperato concepito come gruppo con addizione modulo 12, o si pensi al gruppo di simmetrie del corpo umano; esso contiene due elementi, l'identità e la riflessione rispetto a un piano verticale. È possibile costruire una tavola di moltiplicazione per questo gruppo, servendosi dell'operazione $*$ (che significa 'seguito da') dove I denota l'identità e r la riflessione.

*	I	r
I	I	r
r	r	I

È interessante osservare come il gruppo composto da i due numeri interi (1, -1) insieme alla moltiplicazione ordinaria, sia isomorfo al primo:

x	1	-1
1	1	-1
-1	-1	1

Il comportamento del bicondizionale (\leftrightarrow) mostra la medesima struttura matematica:

\leftrightarrow	V	F
V	V	F
F	F	V

Per Tarski, ogni teoria è formalmente corretta e materialmente adeguata se è possibile derivare tutti i bicondizionali della forma: $\alpha \leftrightarrow 'a'$ dove α sta per

il nome di un enunciato del linguaggio che vogliamo analizzare (linguaggio oggetto) e 'α' per la sua traduzione nel metalinguaggio. La precisa simmetria a questo livello è traducibile, sintatticamente, in equivalenza logica, e la stessa relazione geometrica è valida anche per ciò che osserviamo con questi due (e non uno o tre) occhi. Il mondo (inteso come la collezione di oggetti visibili, manipolabili, nella nostra esperienza sensibile) non cambia se lo 'guardiamo' da destra verso sinistra o da sinistra verso destra. La tavola precedente è diversa da quella dell'implicazione materiale, che, storicamente, è il connettivo che ha creato maggiori problemi di adeguatezza. A causa della nostra 'attitudine' alla simmetria, è possibile comprendere perché nel ragionamento quotidiano i comportamenti dei due connettivi (\rightarrow , \leftrightarrow) siano confusi. Secondo la teoria dei modelli mentali, un ragionatore privo di conoscenze logiche tratterà l'implicazione come riferentesi al caso in cui premessa e conclusione siano entrambe vere, e, non andando alla ricerca di altre interpretazioni, finirà per confondere i due connettivi & e \rightarrow .

Esiste inoltre una relazione tra la nozione di oggetto (anche mentale) e la nozione di invarianza rispetto a qualche categoria di azioni, ben descrivibile all'interno della teoria dei gruppi. Tale relazione è stata messa in luce già da Poincaré e da Piaget. In anni più recenti ci si è chiesti se questa relazione può influenzare i processi cognitivi implicati nel ragionamento deduttivo. Che i nostri sistemi cognitivi di base siano dotati di certi schemi e non altri, ha un valore fondamentale per l'analisi "genetica" (in senso piagetiano) della cognizione, alla luce dei rapporti tra patterns di risposta motoria, schemi topologici e ragionamento 'astratto', cfr. Peruzzi (2006). In questa luce, non ha molto senso domandarsi in che modo introdurre una qualche struttura logica in una determinata classe di rappresentazioni mentali, quanto chiedersi come sia possibile che certe operazioni logiche sfruttino precise risorse cognitive e solo alcune di esse. È utile, a questo proposito, segnalare che il linguista George Lakoff e lo psicologo Rafael Núñez hanno proposto di considerare alcuni schemi della logica (terzo escluso, modus ponens, sillogismo disgiuntivo, modus tollens) come risultato della 'proiezione metaforica' di patterns topologici fondati sulle nostre abilità spaziali e visivo-motorie, cfr. Lakoff, Núñez (2005). Si profila così una spiegazione del comportamento di alcuni schemi logici basata sulla possibilità di simulare internamente un'esperienza cognitiva a noi familiare come quella di 'essere dentro' o 'essere fuori' una determinata regione spaziale. Sì, consideri, per esempio, il modus tollens: se immaginiamo due contenitori X e Y come regioni di spazio finite e se X è contenuto in Y, allora, se un elemento α è fuori di Y, α è necessariamente fuori di X. Tale "necessità" viene dunque individuata nella nostra esperienza spaziale e applicata metaforicamente a un dominio astratto come quello logico-matematico.

Oltre a individuare punti di connessione tra la proposta di Johnson-Laird e quella di Lakoff, può essere istruttivo considerare i tipi di operazioni effettuabili su determinate classi di strutture mentali, tra cui i modelli mentali, intesi dinamicamente, e confrontarne il risultato con diversi tipi di inferenze (induttive, deduttive, abduttive). Sarebbe interessante indagare il rapporto

tra tali operazioni e gli schemi topologici sottostanti, in modo da evidenziare come questi siano capaci di condizionare la selezione di certe operazioni logiche e la loro relativa implementazione cognitiva (umana e artificiale), cominciando da alcune di queste operazioni (traslazione, rotazione, rototraslazione, incassamento, scansione, integrazione, inserimento, inversione, cancellazione, sovrapposizione, *fleshing out*).

Un problema emerge quando si desidera conferire una particolare struttura logica alle rappresentazioni mentali semplicemente includendo di simboli logici nelle rappresentazioni, senza dimostrare la liceità dell'operazione. Tale inclusione non spiega affatto l'applicabilità della logica agli oggetti stessi della nostra percezione, né spiega la possibilità della logica. Ci si limita ad assumere che, se è possibile ragionare deduttivamente servendosi di modelli mentali intesi come strutture semi-simboliche, allora, per funzionare, questi devono essere combinati in accordo a principi logici. Tale assunzione può essere legittimamente accettata soltanto per i modelli mentali in dotazione a un computer digitale o a un robot (ove i modelli mentali sono rappresentazioni di alto livello non diverse da altre strutture formali), aggiungendo poi un'interfaccia, con le sue componenti hardware (il 'corpo' del robot) in modo da avere un'architettura cognitiva artificiale in grado di ragionare deduttivamente per mezzo di opportuni modelli mentali (strutture di dati in LISP)⁷.

Per quanto riguarda gli esseri umani, le cose sono molto più complicate. Com'è possibile che delle strutture 3D siano suscettibili di essere combinate in forma di enunciati? Il linguaggio è solo secondario nel ragionamento, limitandosi a esprimere processi mentali che sono di un tipo qualitativamente diverso da quello proposizionale (perché usano schemi di tipo percettivo, motorio e spaziale)? Oppure la struttura del linguaggio incide sull'impossibilità di immaginare determinati stati di cose? Non dimentichiamo che i modelli mentali si originano anche da input di tipo linguistico e il sistema traduce anche l'informazione linguistica in formato spaziale/percettivo.

Recentemente, Johnson-Laird (2008) ha sostenuto che i modelli mentali sono strutture disponibili alla coscienza ma non del tutto ispezionabili come le immagini mentali. Si presentano immediatamente due problemi: se i modelli non sono del tutto ispezionabili da parte della coscienza, come facciamo a intervenire sui processi mentali in corso modificando un modello provvisorio, includendovi altre entità? E quali prove abbiamo, anche di natura neurologica, che modelli 3D incorporino simboli "che permettono di rappresentare concetti astratti"?

I modelli mentali costituiscono le strutture dalle quali vengono proiettate le corrispondenti immagini mentali concepite come rappresentazioni di una scena da un particolare punto di vista. Come un programma usa un vettore tridimensionale per proiettare un'immagine bidimensionale da un particolare punto di vista, così il nostro cervello estrae un'immagine bidimensionale

⁷ <http://webscript.princeton.edu/~mentmod/models.php/>.

da un modello tridimensionale. Quando ragioniamo, le immagini devono poggiare su una rappresentazione del problema e Johnson-Laird ipotizza che tale rappresentazione sia un modello tridimensionale e non un'immagine visiva (bidimensionale). I modelli mentali sono dunque concepiti come "strutture astratte simili ai vettori spaziali tridimensionali" utilizzati nei ben noti programmi di grafica tridimensionale. Se assumiamo che i nostri dispositivi computazionali costruiscano un array tridimensionale suscettibile di combinazione logica, anche includendo in tali strutture "simboli astratti per la negazione, la possibilità e la proprietà", dobbiamo allora chiederci quale, tra i nostri moduli, assumendo la natura modulare del ragionamento, si prende l'incarico di combinare tali strutture in accordo ai connettivi logici. Se un connettivo non può essere rappresentato internamente a un modello, in che modo esso ne vincola la costruibilità? Come fa qualcosa che non può trovare una concreta rappresentazione in un modello a guidare l'esatta combinazione dei modelli? Se la struttura dei modelli è tridimensionale, la geometria dei modelli è anch'essa (almeno) tridimensionale e, se la logica sfrutta la versatilità dei modelli, essa ne sfrutta, kantianamente, anche la geometria, che è poi la stessa del mondo reale a noi accessibile. Di conseguenza, anche i nostri pensieri sarebbero soggetti a trasformazioni geometriche.

7. Pensabilità, alberi di refutazione, geometria

Secondo il teorico dei modelli mentali, le immagini mentali non possono rappresentare concetti astratti come negazione, disgiunzione, congiunzione. La mera compresenza di due rappresentazioni mentali, infatti, non ci dice se siamo di fronte a una congiunzione, a una disgiunzione inclusiva o a una implicazione. Se pensiamo che negare una proposizione equivalga a ruotare, poniamo, di 180 gradi un vettore (modello) in uno spazio tridimensionale, occorre possedere strutture tridimensionali per ragionare. Ma non è così che i modelli mentali sono intesi da Johnson-Laird.

Torniamo all'esempio della chiave e della serratura. Cosa significa che un simbolo possa trovare posto in un modello tridimensionale cinematico di uno o più enunciati? Ciò che percepiamo sono sequenze di scene e sembra che esistano limiti architettonici alla stessa immaginabilità di determinate situazioni. Questo perché il nostro unico modello è ciò che resta del mondo reale dopo il filtraggio altamente strutturato operato dai nostri dispositivi cognitivi che si conserva nella struttura dei modelli mentali. I nostri dispositivi non dispongono di una combinatoria esaustiva di modelli.

Se per ragionare sfruttiamo il mondo reale o suoi opportuni modelli mentali come background e se questo è recuperabile nelle forme logiche, allora la nostra competenza logica dipende intrinsecamente dalla struttura 'logica' della realtà. Inoltre, se il ragionamento logico sfrutta la versatilità dei modelli mentali concepiti come elementi di uno spazio, allora potremmo utilizzare la geometria per descrivere sia la natura geometrica della logica (classica) sia la struttura geometrica delle nostre rappresentazioni, mettendo in luce

le radici geometriche della logica. A questo riguardo è da segnalare che la logica (classica) può essere intesa, cfr. Mizraji (2008), come logica vettoriale (vector logic) sfruttando gli strumenti dell'algebra lineare: i valori di verità diventano vettori, gli operatori logici diventano operazioni algebriche nello spazio vettoriale e i teoremi (tautologie) della logica classica sono dimostrabili utilizzando l'algebra delle matrici. Allora i modelli mentali costituirebbero parti della struttura di un opportuno iperspazio, ove situare un modello di razionalità minimale così com'è descritto dalla teoria dei modelli mentali per la logica proposizionale.

Grazie all'algebrizzazione della logica, è emersa la sostanziale intertraducibilità tra classi di reticoli e logiche proposizionali. Per mezzo della teoria della categorie, è stato poi possibile inquadrare la corrispondenza biunivoca tra algebre di Boole e spazi di Stone (spazi topologici compatti di Hausdorff totalmente sconnessi) in una cornice più generale in cui la dualità acquista significato logico. La teoria dei modelli mentali propone una combinatoria di strutture in cui s'incontrano logica, matematica e cognizione, ma non tiene sufficiente conto degli esiti dell'algebrizzazione della logica.

8. Programmazione e conseguenze logiche

L'impostazione comune alle architetture congeturate dai cognitivisti è riconoscibile negli attuali programmi di grafica vettoriale. Tramite tali programmi un'immagine è descritta mediante un insieme di primitive geometriche (punti, linee, poligoni, curve, sfere, cilindri, toroidi, ellissoidi, ecc). Ogni poligono può essere suddiviso in triangoli (primitive): l'immagine di un poligono qualsiasi viene descritta da un sistema di equazioni, consentendo così una codifica dell'immagine indipendente dalla risoluzione. L'approccio geometrico permette la memorizzazione 'astratta' dell'immagine, guadagnando in risoluzione. Ed è a questa proprietà astratta dei modelli che Johnson-Laird si riferisce quando afferma che i modelli sono strutture indipendenti dal 'punto di vista'.

Se ammettiamo che i modelli diventino 'mentali' in seguito alla memorizzazione di componenti percettive, motorie, spaziali e che mantengano con la realtà un rapporto di 'somiglianza strutturale' dovuto alla conformazione dei nostri corpi e delle nostre menti, com'è possibile recuperare la fonte originaria dei modelli data l'unidirezionalità di questo divenire? È semplicemente una impossibilità biologica o è di tipo logico? E quali sono esattamente i limiti architettonici coinvolti nell'impossibilità di esplicitare l'implicito? Il teorico dei modelli mentali non invoca proprietà di tipo architettonico, quanto, piuttosto, di tipo logico: il metodo proposto da Johnson-Laird per la logica (classica) non è altro che la versione speculare, per non dire simmetrica, del metodo degli alberi di refutazione di Beth che rappresenta un esempio (un algoritmo) della convergenza di metodi sintattici e semantici per la logica classica. L'idea di Beth era quella di fornire un metodo analitico che ricorre alle sole sottoformule di una formula per costruire modelli, esaminando un albero di contro-

esempi possibili e ragionando per assurdo: se ciascun “ramo” dell’albero si rivela contraddittorio, la formula è provata. Quello messo a punto è un metodo esaustivo per dimostrare la validità di una formula mostrando l’impossibilità di portare a termine la costruzione di un modello che la falsifichi.

Il metodo offerto da Johnson-Laird prevede la costruzione dei modelli in accordo al “principio di verità” e sembra semplicemente una versione rovesciata, ‘diretta’, del metodo di Beth. Alla luce di ciò, risulta arduo comprendere cosa intenda Johnson-Laird quando afferma che la mente umana possa ragionare validamente senza ricorrere alla logica, cfr. Johnson-Laird (1983). Da una parte, infatti, si sostiene che sia possibile ragionare su enunciati a condizione di ragionare su modelli mentali di tali enunciati, dall’altra si sostiene che un’inferenza sia possibile da un insieme di enunciati solo se prima si sono costruiti opportuni modelli mentali di tale insieme di enunciati: la costruibilità dei modelli diventa condizione necessaria e sufficiente per la manifestazione della competenza deduttiva. Ma quello dei modelli mentali non è in realtà un metodo esaustivo, perché i modelli nascono già particolarizzati a causa di assunzioni di default (non rappresentano i casi negativi e la loro struttura insieme alle regole di combinazione, non possono essere isomorfe a un calcolo logico).

I soggetti sperimentali non hanno consapevolezza di dimostrare o derivare o dedurre rappresentazioni mentali da altre rappresentazioni. Se ci serviamo di modelli mentali per dedurre modelli da modelli, è lecito domandarsi cosa significhi ‘dedurre’ in questa accezione. Per il teorico dei modelli mentali, significa costruire e manipolare. Se il sistema è, poniamo, il gioco degli scacchi, la dama, o il Sudoku, possiamo parlare di ‘deduzioni’ in base alle regole dei giochi in questione, ma i nostri processi cognitivi dedurrebbero esattamente cosa da cosa? Possiamo anche definire i processi mentali di costruzione e manipolazione coinvolti nella risoluzione dei problemi di logica come ‘deduttivi’, ma cosa ci autorizza a definirli così? Se guardassimo i modelli, vedremmo soltanto scene centrate sull’ego e sentiremmo l’eco del nostro dialogo interno: la difficoltà del problema risiederebbe nel costruire i modelli di situazioni compatibili con le premesse e nel tenere a mente l’insieme delle ‘computazioni intermedie’. La correttezza delle nostre deduzioni dipenderebbe dalla potenza della nostra memoria di lavoro.

Se disponiamo di una combinatoria di modelli mentali non espliciti, non possiamo pretendere di effettuare sempre deduzioni valide. Torniamo, così, all’attualità della critica antipsicologista di Frege, che non era orientata sintatticamente ma semanticamente. Gli esponenti della scuola logica polacca, tra cui Tarski, hanno dimostrato che è possibile studiare le proprietà dei sistemi formali come la coerenza e la completezza, indipendentemente dalla scelta di assiomi e regole d’inferenza, servendosi della teoria degli insiemi, dell’algebra e della topologia. Alla luce di tali risultati, risulta arduo ipotizzare sistemi deduttivi mentali come fanno i teorici della “logica mentale”.

I modelli mentali sono un costrutto teorico candidato a rimpiazzare schemi formali o altre regole deduttive. Per dedurre correttamente, però, i mo-

delli non bastano: occorre dire al sistema cosa è deducibile da cosa e come. Se davvero disponiamo di modelli mentali per catturare la transitività della conseguenza logica, come catturare tale proprietà per mezzo di modelli tridimensionali di stati di cose? Che cosa significa ‘dedurre’ in questa accezione? Se i modelli mentali originano da input di tipo percettivo, essi sfruttano necessariamente i meccanismi neuropsicologici della visione ordinaria: la scena è costruita da dispositivi dotati di scansione ottica per mezzo delle combinazioni di modelli mentali. In che modo le combinazioni vincolano il significato di proprietà come quella di ‘conseguenza logica’? Che senso ha affermare che un modello è conseguenza logica di un altro modello? Sarebbe come dire che un’immagine retinica è conseguenza logica di un’altra.

Un robot booleano che utilizzi i modelli mentali deve prima disporre delle istruzioni (algoritmi) necessarie per costruire i modelli e per combinarli correttamente, ma non sembra che l’essere umano venga al mondo con un tale sistema di istruzioni, eppure riesce molto presto a fare, e con grande facilità, ragionamenti.

Bibliografia

- M.A. Arbib, *La mente, le macchine e la matematica*, Torino, Boringhieri, 1968.
- S. Bozzi, C. Mangione, *Storia della logica da Boole ai giorni nostri*, Milano, Garzanti, 2001.
- C. Cellucci, *Filosofia e matematica*, Roma, Carocci, 2002.
- G. Frege (1923), *Ricerche Logiche*, a cura di M. Di Francesco, Milano, Guerini e Associati, 1988.
- W. Hodges (2005), *How reasoning depends on representations*, www.maths.qmul.ac.uk/~wilfrid/jadavpur2.pdf.
- D. G. Hofstadter, *Concetti fluidi e analogie creative*, Milano, Adelphi, 1996.
- P. N. Johnson-Laird, *Modelli mentali*, Bologna, il Mulino, 1988.
- P. N. Johnson-Laird, *Mental Models, deductive reasoning and the brain*, in S. Gazzaniga (ed.), *The Cognitive Neuroscience*, Cambridge (Massachusetts), Bradford Book, MIT Press, 1996, pp. 999-1008.
- P. N. Johnson-Laird, Y. Yang, *Mental logic, mental models and computer simulations of human reasoning*, in R. Sun (ed.), *Cambridge Handbook of Computational Cognitive Modeling*, Cambridge, Cambridge University Press, 2006, pp. 339-358.
- P. N. Johnson-Laird, *Pensiero e ragionamento*, Bologna, Il Mulino, 2008.
- G. Lakoff, R. E. Núñez, *Da dove viene la matematica. Come la mente embodied dà origine alla matematica*, Torino, Bollati Boringhieri, 2005.
- M. Livio, *The Equation That Couldn't Be Solved*, New York, Simon and Schuster, 2006.
- E. Mizraji, *Vector Logic*, “Journal of Logic and Computation”, Volume 16, pp. 97-121.
- D. Palladino, C. Palladino, *Logiche non classiche: un'introduzione*, Roma, Carocci, 2007.

- A. Peruzzi, *Il significato inesistente*, Firenze, Firenze University Press, 2004.
- A. Peruzzi, *Il lifting categoriale dalla topologia alla logica*, Annali del Dipartimento di Filosofia 2006, Firenze, Firenze University Press, 2005, pp. 51-78.
- Z.W. Pylyshyn, *Seeing and Visualizing. It's Not What You Think*, Cambridge, Massachussets, 2003.
- P. Wang, *Cognitive Logic versus Mathematical Logic*, Third International Seminar on Logic and Cognition, Guangzhon, May, 2004, pp. 1-5.
- P. Wang, *Wason's Cards: What is wrong?*, Third International Conference on Cognitive Science, Beijing, 2001, pp. 1-10.
- Y. Yang, S. Bringsjord, *Mental MetaLogic: a new paradigm in psychology of reasoning*, in L. Chen, Y. Zhou (eds.), *Proceedings of the Third International Conference of Cognitive Science (ICCS 2001)*, Press of the University of Science and Technology of China, Hefei, 2001, pp. 199-204.