

Pre-print submitted to "Sistemi Intelligenti" (06/2013)

Copyright © 2014 by Società editrice il Mulino

The definitive version is available at publisher's home:

<http://www.rivisteweb.it/doi/10.1422/76304>

GIUSEPPE BOCCIGNONE

STRUTTURA, INFORMAZIONE, MODELLI: IL CONTRIBUTO DI VALENTINO  
BRAITENBERG

ABSTRACT

Braitenberg's neuroanatomical research was carried out in the endeavour to identify the network structures specific to a given part of the brain. But, more generally, his work concerns the functional interpretation of brain structures. Yet, since computers came into play in the 1950s, Braitenberg was well aware of their potential for bridging the gap between neuronal structures and mechanisms, namely, for building up models of brain functions. In this paper we shall argue that the relationships between structure, information and models constitute the spine of Braitenberg's systemic-cybernetic thinking. In a nutshell, the quest for a model calls for the interplay between structure and information. Meanwhile, defining information in terms of structure leads to reconsider the meaning of a model at a certain level of explanation. In this perspective, we begin by discussing his noteworthy contributions to functional neuroanatomy, in particular those related to charting the visual cortex, where Braitenberg explicitly addresses the issue of how to

set up a model so to in efficiently account for the underlying neural mechanisms. Then, we consider his unconventional approach to the concept of information, which is shaped in terms of structural information, rendering it a powerful tool for understanding and thus modelling the neural mechanisms of cognitive functions. The paper concludes by suggesting and discussing some implications of Braitenberg's view for most recent and lively debate on models and levels of explanation in the behavioural sciences.

## 1. INTRODUZIONE

Rileggere Braitenberg è porsi il problema del rapporto fra informazione e struttura. E di come questo rapporto condizioni la costruzione di modelli nell'ambito delle scienze cognitive. Questo è il tema di cui qui ci vogliamo occupare.

Ciò detto, è banale constatare che l'opera di Valentino Braitenberg è un menu ricchissimo di fronte al quale esiste solo l'incertezza dove affondare il coltello.

Troviamo lavori comportamentali, piccoli classici, come quello sulla *musca domestica* (Braitenberg e Taddei Ferretti 1966). La biologia teorica (Braitenberg et al 1959; Braitenberg e Lauria 1960; Braitenberg 1974). La passione artigianale per il microscopio e l'analisi dettagliata della corteccia (Braitenberg 1985). La modellistica sintetica e le incursioni nella robotica con i *vehicles* (Braitenberg 1984). La sintesi epistemologica (Braitenberg 1967 bis; Braitenberg 2008).

Diciamo pure che Braitenberg ha avuto la capacità, e la possibilità storica, di divertirsi, riprendendo il titolo di un suo articolo, con teorie, modelli sperimentali e giocattoli cibernetici (Braitenberg 1974).

Ad esempio, nella ricerca che Braitenberg e Taddei Ferretti

(1967) conducono sulla reazione di atterraggio della mosca, vi è proprio il tratto della migliore cibernetica: l'idea che percezione e azione non siano processi separati ma possano essere concepiti nella loro compiutezza solo se considerati come parti integrate dell'anello di percezione/azione in cui si costituisce la dinamica di un agente (biologico o artificiale) situato nel mondo reale (per una disamina storica di queste tematiche rinviamo a Cordeschi 2002).

E' questo un tema ricorrente in Braitenberg: anche nella costruzione dei suoi veicoli (Braitenberg 1984), lo sforzo è di mostrare come comportamenti psicologici complessi possano già scaturire da strutture cerebrali semplici, purchè poste in interazione costante con il mondo. Per inciso, proprio nel congegnare i veicoli, soprattutto i più semplici, fatti di componenti come motorini ruote e sensori, l'elaborazione dell'informazione (almeno nella sua accezione più tradizionale e scontata) è accuratamente evitata a favore di collegamenti diretti tramite fili tra sensori ed attuatori. Una omissione, o un atto di trasgressione concettuale che lascia odor di zolfo: in altre e numerose occasioni il tema dell'informazione è centrale, per esempio per interpretare e mettere ordine nella costellazioni di neuroni che Braitenberg scruta nei suoi vetrini.

Così, riandando ai primissimi lavori pubblicati, frutto della frequentazione con la scuola del fisico Caianiello (si veda per esempio: Braitenberg *et al* 1959; Braitenberg e Lauria 1960) registriamo un'attenzione particolare rivolta alla costruzione di modelli a partire dai dati sperimentali. Attenzione dovuta non solo alla pratica scientifica ma anche, e probabilmente soprattutto, epistemologica. E per certo quelle prime prove non

sono *phrénesies de jeunesse*, giacchè il tema sollevato è motivo costantemente ripreso negli anni. Allora, come riordinare le fila di queste molteplici ispirazioni?

Ci è sembrato che il tema della struttura e del suo non banale rapporto con l'informazione potesse essere il filo conduttore. Un tema che Braitenberg svolge in maniera originale e che lo colloca prossimo a Donald MacKay, una delle menti più brillanti e originali, e probabilmente controverse della cibernetica (Cordeschi 1994). Ed è proprio dipanando questo filo che teorie, modelli sperimentali e artefatti «sintetici» si ricompongono in un quadro coerente.

Non solo: sviluppando il tema e approfondendolo alla luce del dibattito scientifico e metodologico più recente, ci pare di poter affermare che Braitenberg abbia colto e per certi versi anticipato alcune delle criticità su cui a tutt'oggi si discute nell'ambito delle scienze comportamentali e delle neuroscienze cognitive.

## 2. MODELLI E LIVELLI DI SPIEGAZIONE

Dicevamo che il tema della costruzione di un modello spicca fin dall'inizio come centrale nel lavoro di Braitenberg (Braitenberg *et al* 1959; Braitenberg e Lauria 1960). Le motivazioni che conducono alla formulazione di un modello sono quelle classiche (Braitenberg 1985): render conto della miriade di evidenze sperimentali accumulate sulla corteccia cerebrale e suggerire nuove ipotesi e possibili risultati: spiegazione e predizione.

Ma in *Charting the Visual Cortex*, Braitenberg aggiunge un tassello non irrilevante: la nozione di modello efficiente:

«Per modello efficiente intendo un modello che possa essere immediatamente tradotto in programma tale da consentire a un calcolatore che lo esegua di effettuare le stesse operazioni ascrivibili alla porzione di corteccia presa in considerazione» (Braitenberg 1985, 379).

E' evidente in questo passaggio la relazione del concetto di modello in Braitenberg, con il livello di spiegazione della teoria computazionale elaborata nello stesso periodo da David Marr (1982). Marr illustrava la questione con un esempio tratto dalla fisica statistica: così come non è praticabile la formulazione di processi termodinamici a partire da un sistema di equazioni di moto - un'equazione per ciascuna particella, nello stesso modo non è pensabile di ricondurre lo studio del sistema nervoso allo studio dei singoli dei singoli neuroni o delle singole sinapsi. Dal canto suo, Braitenberg sottolinea come sia impossibile pervenire ad una descrizione o modello matematico anche di un singolo neurone corticale, qualora la descrizione sia concepita come un «catalogo» completo delle sue migliaia di sinapsi.

Entrambi (ma forzatamente Braitenberg nelle sue escursioni istologiche) fronteggiano l'arcigno problema di cui narra Borges (1997) in un suo breve racconto. I cartografi dell'impero avevano raggiunto una tale accuratezza che la mappa di una provincia occupava un'intera città e quella dell'impero un'intera provincia: ma quando furono in grado di produrre una mappa così dettagliata da coincidere punto per punto con tutto l'impero, nessuno seppe più che farsene. Una «mappa perfetta» è inservibile, giacché la sua utilità risiede nella capacità di selezionare certe caratteristiche del territorio piuttosto che

altre e di rappresentarle in modo perspicuo. Dove la perspicuità è funzione dell'uso che si intende farne.

La cogenza del problema è tale che, secondo Braitenberg, la strada dell'astrazione per livelli deve essere perseguita ad ogni costo («at all costs»), non escludendo la possibilità che «[...tali costi] possano implicare l'assunzione di ipotesi non perfettamente garantite, ovvero l'esclusione di fatti considerati come generalmente assodati. Una simile deformazione della versione ufficiale è giustificata, qualora costituisca l'unica via percorribile per arrivare ad un modello efficiente, essendo l'efficienza, cioè la fondamentale concordanza con il meccanismo (che è di per sé, ovviamente, efficiente), il criterio supremo per la validità di ogni modello» (Braitenberg 1985, 379).

Su tale base propone una metodologia in cui, fissato il livello di descrizione, i dettagli siano modellabili più propriamente in termini statistici - nella fattispecie, mediante una «nuvola» probabilistica di sinapsi - abbandonando l'analisi a circuiti elettronici equivalenti: intuizione che anticipa con ampio margine una tendenza esibita da alcuni dei lavori più recenti e autorevoli di neurobiologia teorica (Mortimer et al. 2009; Pfister et al. 2010). È un approccio che definisce, con pregevole termine di proprio conio, «neuroanatomia statistica» (Braitenberg e Schuz 1991).

### 3. STRUTTURA E INFORMAZIONE

L'intuizione che prende forma da tali premesse è che non vi sia soluzione di continuità fra l'analisi di un comportamento, lo studio delle strutture nervose, e l'elaborazione di un modello. Se nel lavoro sulla *musca domestica* (Braitenberg e Taddei

Ferretti 1967) viene esperita un'indagine sistemica, la prosecuzione naturale per supportare l'analisi comportamentale è l'anatomia. Tra il 1967 e il 1970 Braitenberg pubblica due lavori, ampiamente citati in letteratura, sulla struttura del sistema visivo della mosca (Braitenberg 1967; Braitenberg 1970). Torna alla sua specialità: quello per cui «viene pagato»: l'osservazione al microscopio di strutture.

Lo studio della struttura è sicuramente costruzione di rigorose architetture ma anche il bisogno di un contenuto vivo, di una lingua concreta e aderente alla vita. La vita stessa è per Braitenberg struttura. Ma, come dicevamo, (ri)comporre la struttura è anche il primo passo nella costruzione di un modello.

Un esempio paradigmatico del suo modo di procedere e della sua efficacia è sintetizzato in *Charting the Visual Cortex* (Braitenberg 1985). Dove, cominciando con un'analisi di struttura e geometria spaziale, propone quale modello della corteccia visiva una macchina computazionale che definisce «di compromesso», perché in grado di analizzare congiuntamente lo stimolo sia nel dominio delle frequenze sia nel dominio dello spazio. Anticipa di poco, o comunque vi arriva per un ragionamento diverso, il lavoro fondamentale di Daugman (1985), il quale, ispirandosi al lavoro pionieristico di Dennis Gabor (1945) ed estendendolo a stimoli visivi (bidimensionali), propone i filtri corticali che diverranno noti come filtri di Gabor, primi esempi della moderna teoria dei filtri «steerable» e delle «wavelet» (Simoncelli et al. 1992), strumenti fondamentali per la visione computazionale.

Ora, ritornando al tema qui affrontato, ci sembra che si possano individuare due vie che dall'analisi strutturale

conducono all'informazione.

La prima è tracciabile bibliograficamente. Il lavoro di Braitenberg è tra i pochissimi citati (insieme con i classici di Hubel e Wiesel) nella celebre trilogia di Linsker (Linsker 1986a; Linsker 1986b; Linsker 1986c), dove si articola il fondamento neurofisiologico del principio di massima conservazione dell'informazione noto come Infomax (Linsker 1988).

Qui è importante notare che anche Linsker, come già Marr e Braitenberg, riconosca nel problema dei livelli di spiegazione (dal sub-cellulare al sistemico, come li definisce) e nei legami tra un livello e l'altro, la questione dirimente per comprendere struttura e funzione del cervello. Proprio prendendo spunto dall'analisi di Marr sulla riduzione della ridondanza come principio unificante dei sistemi percettivi, propone, quale criterio generale di ottimizzazione, la massimizzazione della media temporale della mutua informazione (misurabile sulle attività di due strati neuronali: la sorgente/input e il bersaglio/output).

Tuttavia, pur individuando nel principio di Infomax (e dunque, almeno apparentemente, nella teoria dell'informazione classica à la Shannon) lo sfondo su cui delineare il livello della teoria computazionale, ritroviamo nell'analisi di Linsker un'impostazione che lo accosta più a Braitenberg che a Marr. E' Linsker stesso a sottolinearlo, quando afferma che il proprio lavoro ha messo in evidenza «un insieme di possibilità per combinare indagine sperimentale, modellazione e teoria e farle interagire al fine di progredire nella nostra comprensione dei sistemi percettivi. Queste interazioni non si configurano rigidamente come *bottom-up*, né come *top-down*. Richiedono



un'analisi congiunta piuttosto che lo studio isolato dei livelli di descrizione pertinenti a teoria computazionale, algoritmi e hardware» (Linsker 1990, 276).

Ma il legame profondo tra struttura e informazione, ancora più esplicito di quanto evidenziato da Linsker, nonché da studi successivi che germogliano sul terreno fertile di questo programma di ricerca (e.g., Field 1994; Li e Atick 1994), è segnatamente quello indicato da Braitenberg stesso in maniera lapidaria: «Si può parlare di informazione contenuta in una struttura quando l'azione di questa su altre strutture è determinata in maniera essenziale non dalla mera quantità dei suoi elementi, ma dalla loro disposizione» (Braitenberg 2008).

Non vi è citazione esplicita, ma è palese il legame ideale e profondo con i contenuti della critica mossa da MacKay, durante la conferenza di Macy del 1950, nei confronti del concetto di informazione formalizzato da Shannon pochi anni prima (Shannon 1948). Vale la pena richiamarne i punti essenziali (MacKay 1969).

Muovendo dall'accezione propria del termine informazione, nel senso di ciò che determina la forma, MacKay distingue due modi di tale determinazione: per costruzione e per selezione. Quando si opera per selezione, si rimane nel quadro classico elaborato da Shannon: tuttavia, questa è solo una parte della storia.

L'informazione per MacKay è infatti operazionalmente definibile come ciò che causa un incremento (guadagno) di conoscenza nell'osservatore. A tale scopo, è necessario lavorare per costruzione su due piani: i) fornire una struttura, cioè una disposizione di elementi strutturali che funge da rappresentazione di un dato oggetto o evento (informazione strutturale); ii) stabilire il peso evidenziale (e.g., la

precisione o il grado di risoluzione) degli elementi della struttura (informazione metrica). Il contenuto dell'informazione strutturale e metrica è dunque, a posteriori, un indice del processo di costruzione di una rappresentazione. Per farla breve, una teoria completa dell'informazione deve, secondo MacKay, considerare tre processi: la produzione, la replicazione e l'interpretazione di rappresentazioni.

Il primo processo è quello misurabile in unità di informazione strutturale e metrica: esso determina la dimensione dell'*ensemble* di possibili rappresentazioni che possono essere trasmesse/replicate.

Il secondo processo, la replicazione /trasmissione di informazione, è misurabile mediante la selezione di una rappresentazione nell'*ensemble*, e dunque utilizzando gli strumenti formali predisposti da Shannon (entropia).

Per inciso, va detto che qui Braitenberg si spinge probabilmente oltre il MacKay, quando afferma che «per poter parlare di informazione è decisiva la stabilità del supporto materiale in cui l'informazione è contenuta»; e che «si parla di trasmissione di informazione quando la riproduzione di una struttura dà luogo a copie che contengono la stessa informazione» (Braitenberg 2008). Intendendo che entrambi i fenomeni, informazione strutturale e replicazione di strutture, risultano essenziali per la vita (Braitenberg 2008, 76).

L'analisi di questa riflessione, lungamente meditata e rielaborata da Braitenberg, ci condurrebbe lontano, ma è doveroso ricordare che il contenuto è attualissimo: si veda per esempio, la discussione innescata da Kauffman *et al.* (2007) sulla non applicabilità dell'informazione di Shannon a sistemi biotici. Nel caso di tali sistemi, infatti, l'informazione

strutturale rappresenterebbe più propriamente l'insieme organizzato di vincoli che permette di convertire energia libera in lavoro: condizione essenziale per la crescita e la riproduzione di un organismo biologico vivente. La vita è propagazione di informazione (Kauffman et al. 2007): del che si era accorto benissimo Kant: «Un ente organizzato non è dunque una mera macchina: questa, infatti, non ha che forza motrice e quello, invece, possiede in sé forza formante e tale, invero, da comunicarla alle materie che non l'hanno (da organizzarle): dunque, una forza formante propagatesi, che non può essere spiegata mediante la facoltà del movimento (il meccanismo) soltanto» (Kant 1995, p.597).

Ma torniamo al punto, il terzo tassello per completare il quadro: l'interpretazione della rappresentazione trasmessa/replicata.

#### 4. LO SGUARDO NEL CERVELLO CHE GUARDA

Il legame che l'informazione ha con la struttura nel MacKay è, come ricordavamo, la premessa per una sua definizione operativa: informazione è ciò che causa un guadagno di conoscenza nell'osservatore. Senza osservatore, sia chiaro, non vi è informazione. Su questa posizione si schiera anche Braitenberg che sentenzia: «in una trasmissione di informazione, quindi di significato, l'interpretazione soggettiva da parte del ricevitore è, in linea di massima, sempre decisiva» (Braitenberg 2008, 92).

Sicché, come per MacKay, il concetto di significato («meaning») va precipuamente trovato nell'effetto che una rappresentazione produce sulla dinamica degli stati interni dell'osservatore: l'atto del conoscere si costituisce come attività interna di

risposta adattiva; in particolare, l'atto del percepire si incarna nell'attività interna da cui scaturisce una risposta adattativa corretta.

In tale prospettiva, l'osservatore è caratterizzabile mediante un *ensemble* di stati di «conditional readiness» (MacKay 1969), ovvero mediante l'insieme statistico delle potenziali risposte comportamentali condizionate dall'osservazione: nel corso dell'osservazione alcuni comportamenti diventano più probabili: altri, meno probabili.

Si pone così in essere un meccanismo di selezione reciproca e circolare: l'osservatore, sulla base del proprio stato interno, seleziona il contenuto informativo strutturale e metrico nell'*ensemble* delle possibili rappresentazioni; l'informazione strutturale seleziona un potenziale (nuovo) stato di risposta nell'*ensemble* di stati (lo spazio di risposta adattivo) che l'osservatore può condizionalmente assumere. A sua volta il nuovo stato interno determina la selezione di contenuto informativo, e così via.

La rappresentazione interna «del mondo nella testa» è dunque un processo, o meglio un insieme di processi, e coincide esattamente con il repertorio delle risposte interne evocate. Per dirla con Braitenberg, «La dinamica dello spettacolo che ha luogo nel teatro interiore del mio cervello è una rappresentazione dei nessi causali tra gli eventi nel mondo esterno» (Braitenberg 2008). E' bene sottolineare, a scanso di equivoci, un'avvertenza preliminare di Braitenberg: «quando io qui parlo di cervello, intendo per semplicità l'intero tessuto nervoso; nei vertebrati, quindi, cervello, cervelletto, tronco dell'encefalo, midollo spinale e sistema nervoso periferico[...]» (Braitenberg 2008, 96).

Dunque, per Braitenberg, nel mondo esterno gli eventi si susseguono per causalità, e questa dinamica causale ha una corrispondenza nella dinamica causale del cervello: sulla base di esperienze passate il cervello opera continuamente una previsione sul più probabile stato futuro del mondo, anticipandolo al suo interno: «Quello che in questo momento accade nel mio cervello è una rappresentazione del mondo come esso è ora, con l'aggiunta di un frammento di passato, e anche di un frammento di futuro previsto» (Braitenberg 2008).

Per chi ha dimestichezza con le più recenti tendenze della visione computazionale (cfr. Prince 2012, per una rassegna esaustiva), ma anche della robotica situata (Thrun et al. 2005) e delle neuroscienze teoriche (Penny 2012), è accattivante rileggere questa ipotesi secondo uno schema interpretativo bayesiano: l'osservazione di un evento, in questo preciso istante, opportunamente pesata con la conoscenza a priori - intesa come conoscenza della dinamica dell'evento congiunta alla conoscenza dello stato del medesimo all'istante precedente - consente di effettuare un'inferenza dello stato attuale nonché una predizione dello stato futuro.

E' evidente che questo è uno schema predittivo/ inferenziale del tutto generale. Poi si tratta di vedere come legarlo con l'idea che un agente possa agire nel e sul mondo, e che possa apprendere nel corso della sua interazione con il mondo.

Braitenberg ha le idee chiare e individua un meccanismo che consente di porre in relazione inferenze/predizioni con apprendimento e azioni: il confronto tra ingressi sensoriali e rappresentazione interna. Qui Braitenberg non sembra discostarsi da un tema caro alla cibernetica classica (per esempio, lo stesso MacKay, 1969; ma per una discussione

generale, si rimanda a Cordeschi 2002, e per una disamina specifica a Cordeschi 1994).

Tuttavia, è lecito affermare che il merito di Braitenberg è di puntualizzare come la valutazione dell'incongruenza, della discrepanza tra quello che ci si aspetta e quello che accade si possa costituire come strumento unico e unificante per contemperare apprendimento e azione. In sintesi: il tentativo di minimizzare la differenza tra predizione e osservazione, nel caso dell'apprendimento innesca il meccanismo di revisione delle regole con cui il cervello produce le sue previsioni - cioè quello che Braitenberg denomina «la correzione dell'immagine del mondo»; nel caso dell'azione, scatta *tout court* una «correzione del mondo».

La proposta di Braitenberg, elegante nella sua parsimonia esplicativa - un buon esempio di applicazione del rasoio di Ockham - va presa sul serio; soprattutto alla luce di proposte recenti che coniugano inferenza bayesiana e minimizzazione dell'errore di predizione proprio al fine di unificare apprendimento e azione (Friston 2010). Ma torneremo più avanti su questo punto.

Sembrerebbe invece che, congetturando ad un livello così generale, quello dell'interazione fra un agente e il mondo, di aver smarrito il filo conduttore. In realtà non è così, se si tiene presente il metodo con cui Braitenberg arriva a stendere l'affresco.

Il cervello è una mappa del mondo che nella sua struttura codifica la struttura causale del mondo. Ma, per costruire la mappa, è ancora decisiva la disposizione di (gruppi di) neuroni nel tessuto nervoso: la geometria dei collegamenti. E per Braitenberg lo studio delle diverse disposizioni delle

fibre è il *terminus a quo* per inferire quale elaborazione dell'informazione ha luogo. L'informazione è struttura: d'altro canto, che cosa sono i suoi veicoli se non esempi di variazione incrementale della struttura?

Ancora una volta, è bene ribadirlo, predizione e inferenza non sono oggetti meramente speculativi. Nascono dal lavoro quotidiano al microscopio. Un esempio elegantissimo è reso in *Thoughts on the Cerebral Cortex* (Braitenberg 1974). Affrontando un itinerario guidato dalla geometria dei collegamenti dei neuroni piramidali, Braitenberg «deduce», fra i possibili modi di operazione, l'inferenza e la predizione. Via via dipanando il groviglio di assoni e sinapsi, ricostruisce architetture da cui, in ultima istanza, per ingegneria inversa, traccia il modello computazionale: «I neuroni piramidali sono elementi di memoria. Hanno capacità di apprendere le proprietà di uno stimolo in ingresso, perché hanno la tendenza ad apparire in determinate successioni[...] Gli assoni collaterali dei piramidali codificano le probabilità condizionali di come le proprietà rappresentate si succedano, in altri termini, rappresentano la struttura causale del mondo. [...] Il riconoscimento di accadimenti familiari, l'analisi di condizioni inconsuete, e la predizione del futuro sono operazioni che possono essere dedotte naturalmente e facilmente dalle proprietà descritte »

Questo in sintesi il metodo.

Tuttavia, nel punto in cui siamo arrivati, una questione delicata si pone senza remissione. Abbiamo postulato fin dal principio la stretta relazione che, *prima facie*, intercorre tra il concetto di modello in Braitenberg e il livello di spiegazione della teoria computazionale elaborata da Marr

(1982). Ma in questo raffronto non va dimenticato o trascurato uno dei presupposti dell'impostazione marriana: l'indipendenza fra i livelli.

Con Braitenberg invece scopriamo che la relazione fra struttura e modello è vincolante: al livello di teoria computazionale, d'accordo con Marr, ci si occupa di elaborazione dell'informazione, ma a condizione che la struttura causale del mondo si rifletta in ultima analisi nella struttura causale (e statistica) del tessuto nervoso. Con buona pace di Marr.

Anche in quest'occasione, ci pare, Braitenberg precorre una tendenza che ha guadagnato consensi in questi ultimi anni. E che ci costringe nuovamente là dove siamo partiti: sul terreno dei modelli.

##### 5. ANCORA MODELLI, PER CONCLUDERE

Ricapitolando, vi è un mondo che ha una sua struttura causale: il cervello (ricordiamo l'accezione estesa del termine) è parte di questo mondo e contemporaneamente esibisce una sua propria struttura che gli permette di mappare in modo dinamico la struttura causale del mondo. Possiamo studiare il cervello al livello di neuroni e sinapsi, gruppi di neuroni, aree cerebrali e comportamenti (Braitenberg 1967 bis): ai diversi livelli avremo la possibilità di definire processi di inferenza e predizione, di apprendimento e azione modulabili mediante un meccanismo di confronto e riduzione dell'incongruenza tra l'immagine inferita/predetta e quella che ci consegnano i nostri organi di senso.

Per pesare l'importanza di questa posizione, va ricordato che l'idea della minimizzazione della discrepanza fra predizione e osservazione del mondo - la base dei metodi di codifica



predittiva («predictive coding») - è stata tecnicamente generalizzata da Friston (2010) nel concetto di minimizzazione Bayes-variazionale dell'energia libera di Gibbs (un approccio, detto per inciso, in grado di sussumere l'Infomax di Linsker). Con una certa licenza, si potrebbe affermare che Friston ha articolato in maniera formale il ragionamento svolto da Braitenberg: utilizzare la discrepanza come strumento di controllo sia dell'apprendimento sia dell'azione sul mondo. Da un punto di vista filosofico, la rilevanza del risultato è stata amplificata da Andy Clark (in corso di stampa), arrivando ad individuare nel medesimo meccanismo il viatico per unteoria unificata della mente.

Ora, al di là degli entusiasmi che la (ri)scoperta di meccanismi generali suscita, è doverosa una certa prudenza nel trarre conseguenze epistemologiche. E' stato infatti osservato (Boccignone e Cordeschi 2012) come su un piano squisitamente tecnico un approccio Bayes-variazionale fondato su una procedura di minimizzazione dell'energia libera, non è propriamente un modello (se vogliamo ragionare restando nello schema di Marr). E' piuttosto un algoritmo efficiente per simulare in maniera approssimata e deterministica inferenza e/o apprendimento in un modello probabilistico (in alternativa ad algoritmi di approssimazione stocastica, quali le simulazioni Monte Carlo). In breve: a livello computazionale si identifica il modello bayesiano (e rinviamo a Penny 2013, per una rassegna esaustiva di esempi); successivamente si utilizza una tecnica di ottimizzazione variazionale per effettuare inferenze approssimate. Esempi ben noti in letteratura (Lee e Mumford 2003; Rao 2005; Litvak e Ullman 2009; Penny 2012, per una rassegna) ripercorrono puntualmente tale schema.

Detto questo, è evidente che la vera questione resta la costruzione di un modello computazionale vincolato dalla struttura.

Ora, è curioso notare che se in compagnia di Braitenberg prendiamo per buona questa strada, siamo costretti ad un salto epistemologico rispetto alla proposta di Marr, imputabile proprio alla nozione di struttura.

Infatti, se si esamina la consueta modalità di progettazione di un modello computazionale probabilistico, questa si può grosso modo articolare in due passi: 1) si identificano le variabili aleatorie fondamentali che descrivono il problema; 2) si impongono le dipendenze condizionali tra le variabili, dipendenze atte a catturare i vincoli del problema. Un esempio semplice: se  $S$  è la variabile aleatoria che denota lo stato interno di un soggetto (per esempio, allegro/triste) ed  $E$  è la variabile aleatoria che descrive un'espressione facciale (l'insieme delle microespressioni indotte da  $S$  - sopracciglia abbassate, naso arricciato, ecc. - à la Ekman, per intenderci), il vincolo che uno stato emotivo  $S$  generi un'espressione facciale  $E$  (e non il contrario), è modellabile con la probabilità condizionata di  $E$  dato  $S$ , cioè  $P(E | S)$ . Questo vincolo consente di fattorizzare correttamente la probabilità congiunta  $P(S,E)$  come  $P(S,E) = P(E | S) P(S)$ , dove  $P(S)$  è la probabilità a priori di uno stato emotivo  $S$ . Ora, l'insieme delle dipendenze statistiche (e in particolari contesti, anche causali), nella moderna teoria dei modelli bayesiani è isomorfo alla struttura istanziata nel modello grafico probabilistico, nella fattispecie un semplice grafo diretto a due nodi:  $S \rightarrow E$  (per una dimostrazione formale e una discussione approfondita, si veda Koller e Friedman, 2009).

In questa prospettiva, dunque, esattamente come delineato da Braitenberg, la struttura recita un ruolo fondamentale nella definizione del modello. Tuttavia, identificando la struttura come matrice di costruzione del modello, si introducono alcune criticità rispetto all'impostazione originale di Marr (Boccignone e Cordeschi 2007).

Primo, il livello del modello computazionale perde la sua indipendenza rispetto al livello implementativo.

Si consideri lo schema illustrato in Figura 1. Immaginiamo il mondo esterno costituito, per semplicità, da gruppi o *cluster* di oggetti puntiformi che si muovono su un piano (con riferimento alla figura, in alto a destra). I *cluster* obbediscono ad una semplice dinamica che li vincola a muoversi orizzontalmente, verso destra o verso sinistra. Il problema computazionale per l'osservatore è di inferire la direzione di moto dei *cluster*: a partire dalla sequenza temporale di immagini retiniche, è necessario computare il flusso ottico e, nel contempo, individuare un meccanismo per risolvere il cosiddetto problema dell'apertura ottica che il calcolo locale dei vettori di flusso ottico non è in grado di dirimere. Il modello computazionale, mostrato in alto a sinistra nella figura, è formalizzato in un modello grafico probabilistico che evolve nel tempo (precisamente, una rete Bayesiana dinamica) e cattura la dinamica causale del mondo. Tuttavia, si noti come la sua struttura gerarchica a tre livelli sia ricavata come rappresentazione iconica della gerarchia anatomico-funzionale Retina / Area corticale V1 / Area corticale Medio-Temporale (MT), mostrata in basso a destra.

In breve: per stabilire i vincoli computazionali che consentono di formulare le dipendenze condizionali/causali del modello probabilistico ci si appoggia sull'architettura, sulla anatomia del tessuto nervoso, dunque su quello che Marr definiva il livello dell'implementazione. E' un approccio che rinuncia in

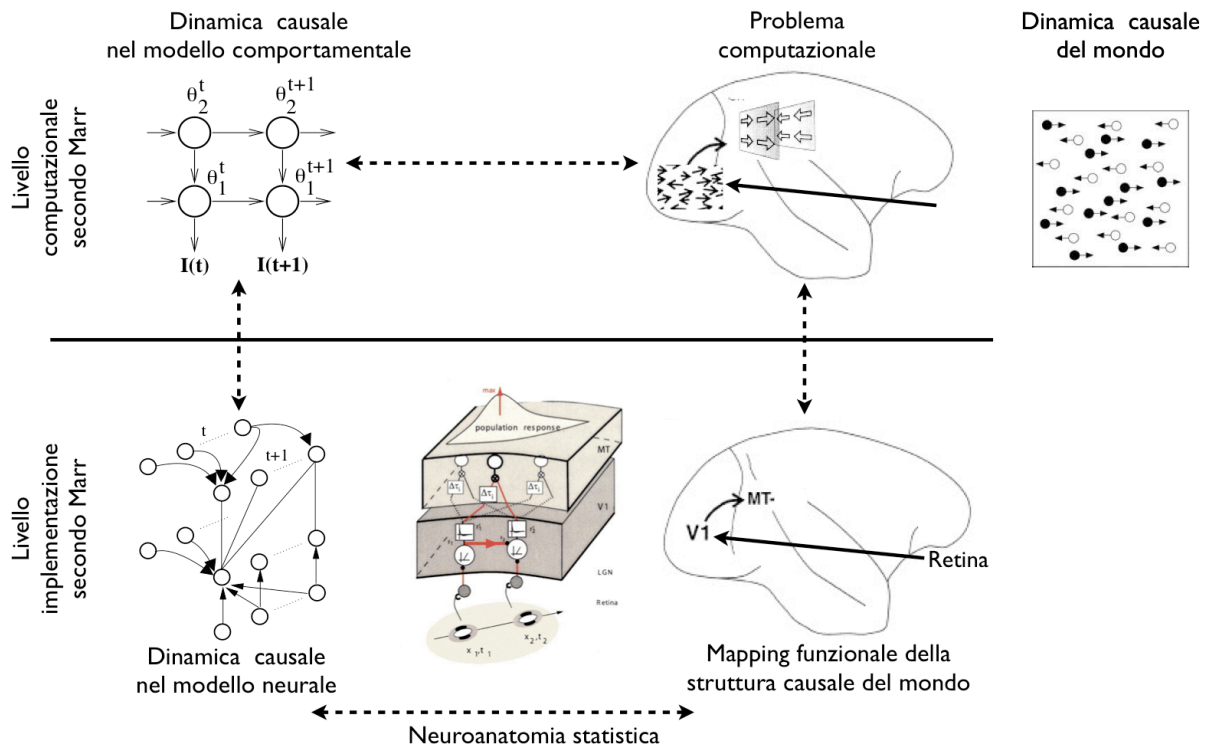


Figura 1. Uno sguardo nel cervello che guarda: modelli e strutture

parte, o del tutto, alla piena indipendenza tra i livelli; per contro consente di imporre, seguendo Braitenberg, una struttura causale nel modello che riflette la struttura causale del mondo: l'immagine del mondo nella testa.

Questo modo di procedere è comune a molti lavori classificabili come approcci bayesiani alle scienze comportamentali e alle neuroscienze cognitive (Penny 2012). Lo stesso esempio delineato nella figura è in parte mutuato da un

lavoro di Rao (2005) (ma si vedano anche Lee e Mumford 2003; Litvak e Ullman 2009).

Un ulteriore elemento, evidenziato da chi sostiene quella che potremmo definire la tesi forte, ontologica, dell'approccio bayesiano, cioè che le computazioni probabilistiche siano realmente implementate a livello neurale, è la possibilità di riduzione dal livello del modello computazionale al livello di architettura neurale (Rao 2005). Questa possibilità è contemplata nello schema presentato: ciascuna variabile aleatoria del modello teorico (in alto a sinistra) viene mappata *top-down* in una rete di neuroni (in basso a sinistra), preservando la medesima struttura causale. Su tale base, per esempio, Rao ha mostrato come le equazioni dell'algoritmo di propagazione di messaggi fra nodi del grafo probabilistico (*Belief Propagation*) possano essere direttamente ridotte alle equazioni che governano la trasmissione di segnali fra neuroni. Peraltro, è opportuno osservare come, nel quadro che stiamo delineando, l'algoritmo di inferenza sia agnostico rispetto allo scopo precipuo della computazione. E' una procedura di scambio di messaggi fra i nodi del grafo del tutto generale: utilizzabile per calcolare, come in questo caso, la direzione del movimento; ma, all'occorrenza, anche la disparità di due immagini stereo. E' la struttura a determinare l'esito della computazione. Tutto sommato un ruolo ben diverso da quello prospettato in origine da Marr (Boccignone e Cordeschi, 2007): e, di fatto, in Figura 1 il livello algoritmico non è neppure rappresentato.

In ultimo, si consideri nell'esempio il segmento che attiene al livello implementativo di Marr. Qui, seguendo Braitenberg, è possibile - sotto la guida di ipotesi generali sulla natura

delle operazioni svolte dalle aree cerebrali prese in esame - determinare un modello teorico della computazione sfruttando le tecniche della neuroanatomia statistica (da destra a sinistra, in basso). Questa possibilità pone con tutta evidenza un ulteriore problema allo schema di Marr: una realizzazione al livello dell'implementazione (disegnata nella figura in basso a sinistra come struttura neurale) può costituire a sua volta un modello computazionale per un livello di spiegazione a granularità più fine.

Ci ritroviamo, dunque, con una pluralità di livelli esplicativi: possibilità che lo stesso Braitenberg aveva già messo in luce molti anni fa (anche se poi in generale, va ricordato, esibiva una preferenza personale per i *cell assemblies*, piuttosto che il livello del singolo neurone, ritenendo i gruppi neurali più utili per fondare modelli computazionali generali, Braitenberg 1959).

Detto per inciso, vi sarebbero ulteriori elementi di discussione da lui individuati con largo anticipo sui tempi: fra tanti, la critica all'utilizzo esclusivo di strutture gerarchiche per comporre l'architettura del sistema nervoso centrale (di certo fondamentali per le aree visive), e l'invito a considerare la possibilità di introdurre congiuntamente strutture eterarchiche (un tema che sta diventando oggi come oggi centrale nella letteratura delle reti su larga scala e che pone seri problemi teorici ai modelli grafici probabilistici, Boccignone e Cordeschi 2012).

Come giudizio conclusivo, ci sentiamo di affermare che Braitenberg ha individuato e apparecchiato sul tavolo diverse questioni cruciali: questioni che a tutt'oggi animano sia il dibattito metodologico ed epistemologico sia la pratica

quotidiana della ricerca. Fra le tante, abbiamo scelto qui di soffermarci sulla centralità della nozione di struttura per definire il problema di elaborazione dell'informazione nei sistemi biologici e artificiali.

Braitenberg ci ha mostrato un percorso: l'esigenza di costruire il modello, porta a ripensare il rapporto fra struttura e informazione. Nel contempo, definire l'informazione in termini di struttura costringe a riconsiderare cosa debba essere inteso per modello.

Un percorso non scontato, in cui la fine incontra il proprio inizio. E sulla cui plausibilità, ci confortano le sue parole: «La circolarità di una storia mi convince di più della sua fondazione in qualche assioma postulato in partenza».

#### RIFERIMENTI BIBLIOGRAFICI

Boccignone G. e Cordeschi R. (2007), *Bayesian models and simulations in cognitive science*, in *Models and Simulations 2*, Tilburg, The Netherlands

Boccignone G. e Cordeschi R. (2012), *Predictive brains: forethought and the levels of explanation*, in «*Frontiers in Psychology*», 3, pp. 511

Borges J.L. (1997), *Sull' esattezza della scienza*, in *Storia universale dell'infamia*, Adelphi, Milano

Braitenberg V., Caianiello E.R., Lauria F., Onesto N. (1959), *A System of Coupled Oscillators as a Functional Model of Neuronal Assemblies*, «*Il Nuovo Cimento*», 11, pp. 278-282

Braitenberg V. e Lauria F. (1960), *Toward a Mathematical Description of the Grey Substance of Nervous Systems*, «Suppl. Nuovo Cimento», 18, pp. 149-165

Braitenberg V. e Taddei Ferretti C. (1966), *Landing reaction of musca domestica induced by visual stimuli*, in «Naturwissenschaften», 53, pp. 155.

Braitenberg V. (1967), *Patterns of projection in visual system of fly. 1. Retina-Lamina projections*, in «Exp. Brain. Res.», 3, pp. 271-298

Braitenberg V. (1967 bis), *On the use of theories, models and cybernetical toys in brain research*, in «Brain. Res.», 6, pp. 201-216

Braitenberg V. (1970), *Order and orientation of elements in the visual system of the fly*, «Kybernetik», 7, pp. 235-242

Braitenberg V. (1974), *Thoughts on the Cerebral Cortex*, «J. theor. Biol. », 46, pp. 421-447

Braitenberg V. (1985) *Charting the visual cortex*, in A. Peters e E. G. Jones (a cura di), *Cerebral cortex*, 3, New York, Plenum Press, pp. 379-414.

Braitenberg V. e Schuz A. (1991). *Anatomy of the cortex: Statistics and geometry*, Berlin, Springer-Verlag.

Braitenberg V. (1984), *Vehicles: Experiments in synthetic*



*psychology*, Cambridge, Bradford Books/MIT Press.

Braitenberg V. (2008), *L'immagine del mondo nella testa*, Milano, Adelphi.

Clark A. (in stampa), *Whatever next? Predictive brains, situated agents, and the future of cognitive science*, in «Behav. Brain Sci.»

Cordeschi R. (1994), *Indagini meccanicistiche sulla mente: la cibernetica e l'intelligenza artificiale*, in V. Somenzi e R. Cordeschi (a cura di), *La filosofia degli automi*, Torino, Bollati Boringhieri, pp. 19–61.

Cordeschi R. (2002), *The Discovery of the Artificial: Behavior, Mind and Machines Before and Beyond Cybernetics*, Dordrecht, Kluwer

Daugman J. G. (1985), *Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters*, in «J. Opt. Soc. Am. A», 2, pp. 1160–1169.

Field D. J. (1994), *What is the goal of sensory coding?*, in «Neural computation», 6, pp. 559–601.

Friston K. (2010), *The free-energy principle: a unified brain theory?*, in «Nature Reviews Neuroscience», 11, pp 127–138.

Gabor D. (1946), *Theory of communication*, in «J. Inst. Electr.

Eng.» 93, 429-457 (1946).

George D. e Hawkins J. (2009), *Towards a Mathematical Theory of Cortical Micro-circuits*, in « PLoS Comput. Biol.», 5.

Kant I. (1995), *Critica del giudizio*, Rizzoli, Milano.

Kauffman S., Logan R.K., Este R., Goebel R., Hobill D. e Shmulevich I., (2007) *Propagating organization: An enquiry*, in «Biol. Philos.», 23, pp. 27-45.

Koller D. e Friedman N. (2009), *Probabilistic Graphical Models: Principles and Techniques*, Cambridge, MIT Press.

Lee T. S. e Mumford D. (2003), *Hierarchical Bayesian inference in the visual cortex*, in «J. Opt. Soc. Am. A», 20, pp.1434-1448

Li Z. e Atick J. J. (1994), *Toward a theory of the striate cortex*, in «Neural computation», 6, pp. 127-146.

Linsker R. (1986a), *From basic network principles to neural architecture: Emergence of spatial-opponent cells*, in « Proc. Natl. Acad. Sci. USA», 83, pp. 7508-7512

Linsker R. (1986b), *From basic network principles to neural architecture: Emergence of orientation-selective cells*, in «Proc. Natl. Acad. Sci. USA», 83, pp. 8390-94

Linsker R. (1986c), *From basic network principles to neural*

*architecture: Emergence of orientation columns*, in « Proc. Natl. Acad. Sci. USA », 83, pp. 8779-83

Linsker R. (1988), *Self-organization in a perceptual network*, in «Computer», 21, pp. 105-117

Linsker R. (1990), *Perceptual neural organization – some approaches based on network models and information-theory*, in «Annu. Rev. Neuro.», 13, pp. 257-281

Litvak S. e Ullman S. (2009), *Cortical Circuitry Implementing Graphical Models*, in «Neural Computation», 21, pp. 3010-3056

MacKay D. (1969), *Information, Mechanism and Meaning*, Cambridge, MIT Press.

Marr D. (1982), *Vision – A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco, W. H. Freeman.

Mortimer D., Feldner J., Vaughan T., Vetter I., Pujic Z., Rosoff W.J., Burrage K., Dayan P., Richards L.J. e Goodhill G.J. (2009), *A Bayesian model predicts the response of axons to molecular gradients*, in « Proc. Natl. Acad. Sci. USA », 25, pp. 10296-10301.

Penny W. (2012), *Bayesian Models of Brain and Behaviour*, in «ISRN Biomathematics», 2012, pp.1-19

Pfister J-P., Dayan P. e Lengyel M. (2010), *Synapses with*

*short-term plasticity are optimal estimators of presynaptic membrane potentials*, in «Nature Neuroscience», 13, pp. 1271-1275.

Prince S. J. D. (2012), *Computer vision: models, learning and inference*, Cambridge University Press

Rao R.P.N. (2005), *Hierarchical Bayesian inference in networks of spiking neurons*, in L. K. Saul, Y. Weiss e L. Bottou (a cura), *Advances in neural information processing systems*, 17 (pp. 1113-1120), Cambridge, MIT Press.

Shannon C.E. (1948), *A mathematical theory of communication*, in «Bell Syst. Techn. J.», 27, 379-423, 623-656

Simoncelli E.P., Freeman W.T., Adelson E.H. e Heeger D.J. (1992), *Shiftable multiscale transforms*, in «IEEE Trans. Inf. Theory», 38, pp. 587-607

Thrun S., Burgard W. e Fox D. (2005) *Probabilistic Robotics*, Cambridge, MIT Press.