

POLITECNICO DI MILANO
Facoltà di Ingegneria dell'Informazione
Corso di studi in Ingegneria Informatica



COMPUTER E CERVELLO A CONFRONTO:
UN'ANALISI CONCETTUALE

Elaborato di laurea di Zennaro Fabio Massimo
Matricola 668997

Relatore: prof. Marco Colombetti
Correlatore: dott. Viola Schiaffonati

Anno Accademico 2006-07

Indice

| | |
|--|----|
| Capitolo 1 – Introduzione | 4 |
| Capitolo 2 – Concetti critici | 14 |
| 2.1 Condizioni di sensatezza e verità..... | 14 |
| 2.2 Definizione di macchina | 17 |
| 2.2.1 Macchina artefatto..... | 21 |
| 2.2.2 Macchina naturale..... | 22 |
| 2.2.3 Macchina come sistema dinamico | 25 |
| 2.3 Attribuzione di funzione | 26 |
| 2.4 Rappresentazione | 28 |
| Capitolo 3 – Il computer | 31 |
| 3.1 Limiti di un computer | 31 |
| 3.2 Categorizzazione delle macchine per input ed output..... | 32 |
| 3.3 Definizione di computer | 36 |
| Capitolo 4 – Cervello e computer | 40 |
| 4.1 Definizione di cervello | 40 |
| 4.3 Rappresentazioni di computer e cervello..... | 42 |
| Capitolo 5 – Il cervello è un computer? | 47 |
| 5.1 Il cervello è un computer..... | 47 |
| 5.2 Il cervello non è un computer | 50 |
| 5.3 Conclusioni | 54 |
| Bibliografia | 56 |

Capitolo 1 – Introduzione

Motivazioni Affrontare un tema come quello dell'identità tra computer e cervello è una sfida affascinante e stimolante.

Affascinante, in quanto è un problema che si rivolge all'analisi di quegli stessi mezzi che permettono all'uomo di conoscere ed apprendere. Si studiano gli strumenti che sono il fondamento di ogni sapere: il cervello, su cui si basa l'intera comprensione della realtà, e il computer, oggi in grado solo di coadiuvare l'operato del cervello in molti campi ma, nelle previsioni più ardite e visionarie, capace di replicare ogni aspetto del cervello umano. Ardite e visionarie in quanto, seppur capaci di suscitare interesse, queste previsioni mancano il più delle volte di solide giustificazioni se non un'ottimistica speranza nell'inarrestabile progresso della scienza e della tecnologia. La risoluzione stessa del quesito che si pone sull'identità tra cervello e computer dovrebbe essere irrinunciabile per sostenere queste predizioni, ma raramente viene fornita una risposta esplicita e coerente.

Stimolante, poiché è una domanda che si svincola dalle strutture e dalle applicazioni specifiche di una particolare disciplina, per permetterci invece di riflettere sulla nostra stessa conoscenza e su ciò che scienza ed ingegneria insieme potrebbero o non potrebbero realizzare. Ci si pone di fronte a quella che potrebbe definirsi una *framework question* (Searle, 1998), ovvero una domanda riguardante i quadri intellettuali all'interno dei quali l'uomo agisce ed opera, una domanda che precede la pratica e l'azione. Infatti, quando si affronta un problema in un qualsiasi campo del sapere, ad esempio l'intelligenza artificiale, si opera sempre in un contesto preciso in cui le risposte alle domande di *framework*, come quella del problema del rapporto computer-cervello, sono date per scontate. In altre parole, la ricerca scientifica è sempre preceduta dalla riflessione e dalla risoluzione di quesiti di natura qualitativa (Newell e Simon, 1976); le risposte a tali domande costituiscono teorie di carattere generale alla base stessa della ricerca. È, ad esempio, il caso della teoria cellulare in biologia o della teoria atomica in chimica; entrambe sono leggi di natura qualitativa, il cui contributo però è fondamentale nell'analisi e nell'interpretazione dei risultati ottenuti all'interno dei rispettivi campi di competenza. Anche la domanda che si affronta vuole dunque essere una domanda di carattere fondante, che si pone prima e al di fuori di un contesto prestabilito.

***Ragioni dello
studio teorico***

Si potrebbe tuttavia essere indotti a pensare che lo studio di una questione affrontata in maniera astratta e su cui si dibatte ormai da anni possa essere superflua o addirittura inutile; potrebbe essere ragionevole ritenere che solo uno studio pratico o applicativo possa fornire delle risposte concrete al problema della possibile identificazione tra cervello e computer. Sebbene il contributo della pratica sia irrinunciabile, si crede che studiare domande ad un tale livello di astrazione non sia affatto privo di utilità, ma anzi presenti degli indubbi lati positivi. Ben lungi dal voler addentrarsi nel problema del rapporto tra pratica e teoria, della precedenza dell'una o dell'altra o della loro complementarità, si ritiene tuttavia opportuno sottolineare sinteticamente alcuni vantaggi di un simile approccio.

Un'impostazione di studio come quella qui proposta è in grado di offrire una visione d'insieme su un argomento e di consentire al lettore di avere una chiara comprensione di contenuti, obiettivi e limiti del problema.

Definire coerentemente ed univocamente i contenuti che saranno parte del dominio del problema è il presupposto di ogni studio rigoroso in quanto consente di stabilire e specificare quali concetti saranno utilizzati nell'indagine.

Definire precisamente gli obiettivi che saranno perseguiti permette di concentrare l'attenzione sul fine prestabilito, evitando di perdersi in questioni irrilevanti, di percorrere strade che non porteranno ai risultati attesi o, addirittura, di non riconoscere il momento in cui si raggiungerà il proprio scopo.

Definire correttamente i limiti consente di evidenziare le possibilità della ricerca, descrivendo fin dove lo studio può spingersi, quali sono, se esistono, quegli ostacoli che con i mezzi teorici e pratici ora disponibili risultano strutturalmente insormontabili.

In definitiva, avere una visione d'insieme del problema significa essere in grado di muoversi con maggiore consapevolezza nel campo stesso del problema, essere focalizzati sugli obiettivi da raggiungere e saper valutare con coerenza difficoltà e risultati.

Questi contributi sono più che mai preziosi nel caso in esame. Infatti il problema del rapporto computer-cervello abbraccia molti ambiti del sapere, riceve contributi da numerosi studiosi e ricorre spesso a concetti a prima vista intuitivi, ma il cui significato preciso è ben lungi dall'essere chiaro e condiviso, ed anzi varia sensibilmente da un'area di studio all'altra. La collaborazione di molti studiosi fa sì che, da una parte, i

confini della materia di studio continuo ad estendersi e, dall'altra, che aumenti il rischio di incomprensioni, rendendo particolarmente difficile l'orientamento nel campo d'indagine. Avere le idee chiare nello studio di un problema come il rapporto cervello-computer è dunque un prerequisito irrinunciabile per poter affrontare concretamente le successive sfide poste dalla pratica.

Definizione dell'obiettivo Ora, l'obiettivo primo di questa tesi è lo studio e l'analisi della domanda apparentemente semplice: “il cervello è un computer?” e della sua possibile risposta. Si vorrebbe cioè valutare se sia possibile stabilire un rapporto di identità tra i calcolatori elettronici e il cervello umano.

Difficoltà dell'obiettivo: Studiare il significato di questa domanda può apparire a prima vista elementare: sembrerebbe trattarsi infatti del semplice confronto tra due oggetti fisici, facilmente distinguibili e identificabili, senza necessità di ricorrere a concetti astratti. In verità queste assunzioni sono errate, e per più ragioni.

1. Ampio orizzonte di conoscenze Primo, uno studio approfondito dell'argomento richiede un vasto sostrato di conoscenze. Non è possibile limitarsi a quello che potrebbe sembrare essere il primo ed unico campo di indagine, quello dell'intelligenza artificiale, ma è inevitabile spaziare abbondantemente anche in altri settori del sapere (Russel e Norvig, 2003). È infatti irrinunciabile il contributo di diverse discipline come biologia e neuroscienze, dedicate allo studio del sistema nervoso, del cervello e del modo in cui viene elaborata l'informazione; psicologia, volta a comprendere come gli uomini pensino ed agiscano; linguistica, interessata a spiegare la relazione tra linguaggio e pensiero ed il modo in cui la conoscenza viene rappresentata; matematica, tesa allo studio dei metodi di ragionamento e ad una loro possibile formalizzazione; filosofia, impegnata nella discussione sulla natura della mente e sulla conoscenza.

2. Conoscenze limitate Secondo, la domanda prevede il confronto fra due oggetti estremamente complessi, cervello e computer, il cui funzionamento risulta enormemente articolato; non solo il comportamento del cervello non è ancora pienamente spiegato e compreso, ma, ironia della sorte, non è possibile trattare del cervello e del computer se non per mezzo dello stesso cervello.

3. Impossibilità di sperimentazione Terzo, per poter trarre conclusioni inconfutabili circa l'identità tra computer e cervello sarebbe necessario poter ottenere anche evidenze pratiche e sperimentali; tuttavia, mentre in alcuni casi è possibile ricorrere

diretta a semplici esperimenti mentali, molto spesso sarebbe necessario il ricorso ad esperimenti concreti che però possono risultare difficilmente realizzabili, o addirittura impossibili da eseguire al giorno d'oggi.

Ridefinizione dell'obiettivo Ne consegue che, data la profondità e la complessità dell'argomento, non si pretenderà di fornire al lettore conclusioni definitive riguardo questo problema. Piuttosto, dal momento che per fornire una risposta affermativa o negativa a questo quesito non risultano sufficienti le sole conoscenze in ambito ingegneristico-informatico, ma servirebbero conoscenze ben al di là del presente ambito di studi, si ridefinirà con maggiore precisione l'obiettivo da perseguire. Si pone quindi, più umilmente, come fine quello di chiarificare il senso della domanda, esaminando ciascuno dei termini coinvolti e spiegandone il significato; definiti i concetti presenti nella domanda, si propongono al lettore una serie di riflessioni attraverso le quali evidenziare e sottolineare similitudini, analogie e differenze tra il cervello umano e i calcolatori elettronici; mostrare, cioè, in quale senso sia corretto parlare di un'identificazione tra cervello e computer e per quali ragioni, invece, questo paragone sembri inadeguato.

Le riflessioni proposte intendono collocarsi all'interno della discussione sull'intelligenza umana e artificiale, e chiaramente all'interno del dibattito sul problema dell'identità tra cervello e computer. Il problema del raffronto tra computer e cervello può essere analizzato da diverse prospettive.

Confronto strutturale È possibile, ad esempio, confrontare computer e cervello da un punto di vista strutturale. Ciò significa valutare l'architettura di un computer e metterla in relazione con un cervello, analizzando le differenze tra i due a livello quantitativo o qualitativo. Un simile paragone è senza alcun dubbio molto utile, specialmente in ambito applicativo, in quanto consente una continua revisione e un progressivo miglioramento nell'implementazione dei computer. Tuttavia esso con difficoltà può tentare di dare una risposta al problema dell'identificazione tra cervello e computer; si tratta infatti di un semplice paragone, incapace di spiegare come l'intelligenza possa realizzarsi nel computer e nel cervello.

Confronto sintattico È possibile, in alternativa, confrontare computer e cervello analizzando il modo in cui essi sono in grado di elaborare simboli. Un computer è infatti in grado di manipolare meccanicamente simboli, senza necessità di attribuire loro un significato. È dunque ragionevole stabilire un paragone tra cervello e computer considerando il loro funzionamento a *livello sintattico*.

Confronto semantico Ma è anche possibile istituire un confronto ad un ulteriore livello che si potrebbe chiamare *livello semantico*. Dal momento che la capacità di attribuire un significato a dei segni è una componente fondamentale dell'intelligenza e del cervello, è possibile indagare in quale modo segni e simboli elaborati dal cervello e dal computer assumano, ammesso che lo facciano, un significato.

Precedenti storici Storicamente, riguardo l'approccio sintattico allo studio dell'intelligenza, il contributo più celebre e rilevante fu senza dubbio quello offerto da Newell e Simon (1976) con l'*ipotesi del sistema fisico di simboli*, presentata durante l'ACM Turing Award. Essi avanzarono questa ipotesi come una possibile legge di tipo qualitativo su cui fondare la definizione stessa di intelligenza e su cui basare le ricerche dell'intelligenza artificiale.

Sistema fisico di simboli Secondo i due ricercatori, gli elementi fondamentali alla base di un *sistema fisico di simboli* sono *simboli e processi*. Si definiscono *simboli* delle entità la cui aggregazione genera *espressioni* o *strutture simboliche*. Tali espressioni sono strutture *fisiche*, nel senso che sottostanno e si relazionano tra loro secondo le leggi della fisica. Si definiscono *processi* una serie di operazioni che operano su una struttura simbolica e permettono, attraverso azioni di creazione, modifica, riproduzione e distruzione, di produrre espressioni da altre espressioni.

Nella teoria del *sistema fisico di simboli*, due concetti sono necessari per evidenziare il rapporto tra simboli, espressioni e oggetti della realtà. Il primo è il concetto di *designazione*: si dice che un'espressione *designa* un oggetto se, data l'espressione, il sistema può o influenzare l'oggetto o agire in maniera dipendente dall'oggetto. Il secondo è il concetto di *interpretazione*: un sistema può *interpretare* un'espressione se l'espressione designa un processo e se, data l'espressione, il sistema può eseguire tale processo; in altri termini, un sistema può *interpretare* se data un'espressione può eseguire il processo che tale espressione designa.

Un sistema costituito da simboli e processi, in cui sia possibile la designazione e l'interpretazione e che goda di proprietà di completezza e chiusura, si definisce un *sistema fisico di simboli*. Ora, secondo Newell e Simon (1976), un tale sistema fisico di simboli ha i mezzi necessari e sufficienti per un'azione generale intelligente. Affermare che i mezzi sono *necessari* significa che ogni sistema che esibisce un comportamento intelligente è, in ultima analisi, un sistema fisico di simboli. Affermare

che i mezzi sono *sufficienti* significa che ogni sistema fisico di simboli può essere organizzato in modo da mostrare un comportamento intelligente. Infine con *azione generale intelligente* si intende un'azione che manifesta lo stesso livello di intelligenza di un'azione umana.

Secondo questa ipotesi, dunque, il requisito fondamentale dell'intelligenza è quello di essere in grado di elaborare e manipolare simboli. Questa idea è confortata da una serie di intuizioni e applicazioni che avevano segnato il cammino dell'informatica e dell'intelligenza artificiale fino agli anni Settanta. Newell e Simon (1976) sostengono infatti che l'evoluzione dell'informatica, segnata dallo sviluppo della logica formale, delle macchine di Turing, del concetto di programmi risidenti in memoria (*stored program*) e dell'elaborazione di liste (*list processing*), abbia condotto inequivocabilmente all'idea che l'intelligenza possa considerarsi il prodotto di un sistema fisico di simboli.

Nel tempo l'ipotesi del sistema fisico di simboli è stata oggetto di diverse critiche. Molti studiosi hanno contestato la teoria cercando di analizzarne i limiti. Alcuni degli argomenti mossi contro la proposta di Newell e Simon sono sinteticamente i seguenti:

**Critiche al
sistema fisico di
simboli**

1. Argomento del *symbol grounding*. L'intelligenza è più della semplice manipolazione di simboli; un computer è inevitabilmente privo di comprensione dal momento che può solo elaborare semplici segni privi di significato (Searle, 1980; Harnad, 1990);
2. Argomento dei *processi non-simbolici*. I processi su cui si fonda l'azione intelligente, e in particolare la percezione, sono non-simbolici e non possono essere riprodotti su un computer digitale (Brooks, 1991);
3. Argomento dell'*inadeguatezza della computazione*. La computazione non è un modello appropriato per l'intelligenza; cervello e computer sono strutturalmente troppo differenti perché un computer possa riuscire ad essere intelligente (Dreyfus);
4. Argomento dell'*intelligenza senza mente*. Gran parte dei cosiddetti comportamenti intelligenti sono in realtà comportamenti *senza mente*, prodotto di una reazione fisica volta a fornire le risposte più adatte ed efficaci alle sfide poste dall'ambiente (Pollack, 2006).

**Attualità del
sistema fisico di
simboli**

Nonostante queste difficoltà, l'ipotesi del sistema fisico di simboli resta ancora una tesi difesa da diversi studiosi (Nilsson, 2006). Alcuni sono propensi a riconoscere l'importanza di processi non-simbolici per modellare l'intelligenza; sono cioè disposti a sostenere che un sistema fisico di simboli sia *non sufficiente* per originare un comportamento

intelligente, ma sono convinti che esso sia comunque una parte *necessaria* dell'intelligenza.

**La critica del
symbol
grounding**

Delle critiche sopra elencate, la prima ha particolare rilievo in questa discussione in quanto introduce a quello che è stato definito come *livello semantico* di confronto tra cervello e computer; essa sottolinea che un sistema fisico di simboli non è in grado di riprodurre l'intelligenza umana in quanto privo di *intenzionalità*, ovvero della capacità di avere una rappresentazione dell'informazione in relazione con il mondo esterno (Harnad, 1990). Indipendentemente dal comportamento, il sistema risulta privo di comprensione. Questa argomentazione è stata avanzata da J. R.

**L'esperimento
della stanza
cinese**

Searle (1980) e sostenuta attraverso l'esperimento mentale della cosiddetta *Stanza Cinese (Chinese Room)*; l'esperimento di Searle suggerisce di immaginare una persona rinchiusa all'interno di una stanza; quest'uomo, in grado di parlare correttamente l'inglese, ma assolutamente privo di qualsiasi conoscenza della lingua cinese, ha il compito di ricevere, attraverso una fessura, un foglio con dei simboli e restituire, attraverso un'identica apertura, una pagina riportante una risposta. Per poter scrivere il foglio di risposta l'uomo può avvalersi di una serie di tabelle riportanti delle regole scritte in inglese con le quali è in grado di far corrispondere ad ogni oscuro simbolo ricevuto uno o più misteriosi segni. Ora, l'operatore che interagisce con la stanza, avendo inserito una domanda in cinese e avendo ricevuto una risposta sensata sempre in cinese, sarà portato a pensare che ciò che vi è all'interno della stanza, metafora di un computer, abbia compreso il quesito e abbia quindi formulato una risposta coerente. Invece, come l'esempio mostra, l'uomo all'interno della stanza ha semplicemente elaborato dei simboli e non vi è stato alcun istante in cui abbia mostrato di comprenderne il significato. Contrariamente a quanto affermerebbe il test di Turing (Turing, 1950; Sterrett, 2000), secondo cui dall'analisi del comportamento sarebbe possibile definire l'intelligenza di una macchina, l'agire di un computer non è invece sufficiente a provarne l'intelligenza: sono infatti richieste un'intenzionalità e un'intelligenza che non possono che risiedere nella mente del programmatore.

Alcune critiche sono state rivolte anche all'argomento della stanza cinese, in particolare dagli studiosi di intelligenza artificiale, nel tentativo di difendere la visione di Newell e Simon. Tra le repliche più interessanti si ricordano in particolare:

**Critiche
all'esperimento
della stanza
cinese**

1. *La replica del sistema.* La comprensione del cinese non può e non deve essere attribuita al traduttore umano, bensì a tutto il sistema nel suo insieme;
2. *La replica del robot.* Si collochi il programma all'interno di un

robot. Quest'ultimo sarebbe allora in grado di interagire con l'ambiente mostrando di avere una comprensione della realtà.

3. La *replica del simulatore del cervello*. Si implementi il programma su una rete in grado di simulare le attivazioni dei neuroni a livello sinaptico che avvengono nel cervello di un uomo che parla cinese. In questo modo la macchina sarebbe dotata di comprensione esattamente come un essere umano.

**Controcritiche a
sostegno
dell'esperimento
della stanza
cinese**

Queste repliche, tuttavia, si mostrano di modesto rilievo, in quanto non riescono a dimostrare realmente in che modo un computer possa sviluppare una comprensione propria (Searle, 1980).

Nel caso della *replica del sistema*, J. R. Searle suggerisce che la stanza in cui avviene l'esperimento sia eliminata e che tutti gli strumenti al suo interno usati dall'uomo siano studiati e interiorizzati dallo stesso. Ora, rimossi tutti gli elementi del sistema e lasciato solo l'uomo, quest'ultimo si comporterà come nel caso precedente, si limiterà, cioè, a compiere un lavoro di manipolazione di simboli, privo di una reale comprensione.

La *replica del robot* si fonda sull'idea che l'intelligenza non sia una semplice manipolazione di simboli, ma comprenda anche l'interazione con l'ambiente. In realtà, ciò che il robot fa è sempre e semplicemente manipolare simboli, eventualmente ricavati da sensori, e in base ai risultati formulati attivare degli attuatori; di fatto, tuttavia, non si spiega come il robot potrebbe avere una comprensione della realtà.

La *replica del simulatore del cervello* presenta il problema, secondo Searle, di concentrarsi su aspetti non rilevanti del funzionamento del cervello. Finché si simulerà la struttura formale del cervello attraverso le attivazioni dei neuroni e non si studieranno invece le sue proprietà causali e l'abilità di produrre stati intenzionali non si potrà avere una macchina intelligente.

**Collocazione
della tesi**

Ora, le riflessioni su computer e cervello che saranno presentate in questo testo intendono collocarsi all'interno del dibattito sull'intelligenza umana e artificiale. Si assumerà come fondamento l'idea che non sia possibile descrivere l'intelligenza solo a livello sintattico. Fatto dunque tesoro dei risultati storici, si propone nelle prossime pagine un'analisi del rapporto tra cervello e computer focalizzato sugli aspetti semantici, con l'obiettivo di analizzare e spiegare in quale modo l'uomo e la macchina possano dare un significato ai simboli che elaborano.

**Struttura della
tesi**

Prima di addentrarsi nell'argomento, si espone sommariamente l'impostazione della tesi nel tentativo di offrire al lettore una visione d'insieme della discussione.

Nel secondo capitolo vengono presentate le *condizioni di sensatezza e verità* della domanda "Il cervello è un computer?"; vengono cioè indicati quei vincoli che la domanda deve possedere per poter ammettere una risposta. Posti questi requisiti, si passa all'analisi dei concetti fondamentali per lo studio del problema affrontato. Si esaminano in primo luogo quei concetti che sono ritenuti indispensabili per poter dare una definizione di *computer* e *cervello*. Per un'analisi a livello fisico si definisce innanzitutto il concetto di *macchina*. Si procede quindi specificando sia alcune sottoclassi del concetto di *macchina*, come *macchina artefatto* e *macchina naturale*, sia il concetto più generale di *sistema dinamico*. Passando poi a livello istituzionale, ovvero al livello di quei concetti necessari per la descrizione di una realtà dipendente da atti istituzionali o normativi, si presentano i concetti di *attribuzione di funzione* e *rappresentazione*.

Con queste conoscenze, nel terzo capitolo, si affronta il problema della definizione di un *computer*. Presi in considerazione i limiti di una macchina ed identificati input e output, viene proposta una categorizzazione delle macchine in base ai loro ingressi ed uscite. Secondo questa distinzione viene quindi proposta una definizione di computer come (*macchina artefatto*) *bitrasduttore, in cui l'output simbolico è una funzione dell'input simbolico*.

Nel quarto capitolo, nel tentativo di stabilire una relazione tra cervello e computer, vengono riesaminati alcuni concetti necessari alla definizione di cervello e viene ulteriormente approfondita la definizione di computer. Si propone una definizione di cervello in analogia con la precedente definizione di computer. Si definisce, cioè, il cervello come (*macchina naturale biologica*) *bitrasduttore, in cui l'output simbolico è una funzione dell'input simbolico*. Infine si mette in evidenza la relazione tra le rappresentazioni di un computer (*rappresentazioni derivate*) e quelle del cervello (*rappresentazioni implicite*).

Infine nel quinto capitolo, senza pretese di dare una risposta definitiva alla domanda oggetto di studio, vengono proposte delle riflessioni sulle analogie e sulle differenze tra computer e cervello. In primo luogo si analizzano gli aspetti per i quali sembra corretto definire il cervello un

computer. Si riportano alcune osservazioni introdotte nei capitoli precedenti e si discute dell'apparente conformità tra il ragionare umano e il calcolare artificiale. In seguito si sottolineano anche le ragioni per cui non sembra opportuno definire il cervello un computer. Si accenna ad alcuni problemi estremamente complessi che sembrano costituire un arduo ostacolo alla speranza di dare una risposta al problema in esame: il problema della coscienza, la teoria dell'acrasia, il problema del libero arbitrio e il problema della totalità della razionalità umana. La conclusione cui si arriverà sarà infine che, riconosciute le differenze tra le rappresentazioni di un computer e quelle del cervello, al momento attuale, *il computer non è un cervello*.

Capitolo 2 – Concetti critici

In questo capitolo vengono espone le condizioni di sensatezza e verità da rispettare per poter rispondere alla domanda “il cervello è un computer?”; esse costituiranno un punto di riferimento per le riflessioni dei capitoli successivi e saranno una garanzia della coerenza delle affermazioni proposte. Seguono poi le definizioni di alcuni concetti fondamentali per lo studio di computer e cervello: se ne presentano alcune necessarie per una definizione dei concetti a livello fisico, ovvero a livello di oggetto puramente materiale (*macchina, macchina artefatto, macchina naturale e sistema dinamico*), ed altre fondamentali invece per un’analisi a livello normativo, ovvero a livello di quelle realtà la cui esistenza è dovuta ad un accordo in un gruppo di “osservatori” (*attribuzione di funzione, rappresentazione*).

2.1 Condizioni di sensatezza e verità

Necessità e ruolo delle condizioni

Per poter valutare il rapporto tra computer e cervello e fornire delle indicazioni atte a rispondere al quesito “Il cervello è un computer?” è necessario che vengano poste alcune condizioni sulla domanda. Si vogliono cioè stabilire e precisare elementari *condizioni di sensatezza e condizioni di verità* che ci consentano di affrontare la domanda adeguatamente.

Dal momento che il problema affrontato implica l’analisi e il raffronto tra due enti distinti, la dichiarazione di queste condizioni si concretizza in vincoli sulla definizione dei concetti di computer e cervello. Assunta una definizione condivisa e intuitiva di cervello come organo, l’attenzione è posta soprattutto sulla definizione di computer. In verità, si tratterà a lungo della definizione di computer, non solo perché essendo un prodotto umano complesso risulta difficile da delimitare con precisione, ma anche perché, contrariamente al cervello, è un concetto proprio dell’ambito di studi di colui che scrive. Si vuole, cioè, fornirne una descrizione accurata e precisa, che, indipendentemente dal contesto di questa tesi, possa essere una valida definizione di *computer*.

Le condizioni di seguito espone servono dunque per stabilire alcuni

parametri che la successiva definizione di computer deve rispettare.

Condizioni di sensatezza Con *condizioni di sensatezza* si intendono una serie di vincoli senza i quali la domanda risulterebbe semplicemente priva di senso. Qualora si scoprisse che le condizioni di sensatezza non possano essere rispettate, sarebbe inutile e superfluo proseguire nello studio. Una domanda insensata non ha infatti una soluzione positiva o negativa, semplicemente non ammette alcuna risposta.

Sembra, tuttavia, che scartare la domanda “Il cervello è un computer?” come insensata sia una conclusione indebita e controproducente. Infatti, la ricorrenza di tale domanda e la frequenza con cui il cervello viene paragonato ad un computer piuttosto che ad una qualsiasi altra macchina o oggetto, lasciano supporre che questa domanda abbia, in fondo, un qualche senso. Sembra, cioè, che esistano delle ragioni condivise per sostenere un paragone tra computer e cervello, ragioni che non possono essere ignorate, ma anzi costituiscono un interessante oggetto di studio. D'altra parte l'accostamento tra computer e cervello e, più in generale, tra macchina e uomo, è stato presente, in maniera più o meno imprecisata, nella mente degli studiosi fin dall'alba dell'informatica e anche prima. Basti ricordare che originariamente la cosiddetta “macchina di Turing” era stata concepita dal suo ideatore come una modellizzazione non di una macchina, ma di un essere umano che per risolvere un problema applica, senza esercitare ingegno o creatività, un metodo rigoroso composto da più istruzioni (Turing, 1936); un altro fatto degno di nota e alla base ancora oggi di incomprensioni da parte di alcuni studiosi (Copeland, 2002), è il fatto che quando Turing, all'interno di quei testi e discussioni su cui si fonderanno i primi teoremi dell'informatica teorica, usa i termini *computer*, *computabile* o *computazione* lo fa sempre in riferimento ad un calcolatore umano e non, in senso moderno, ad una macchina.

Condizioni di verità Con *condizioni di verità* si intendono invece un insieme di condizioni per cui la domanda ammette una risposta positiva. Poste, dunque, tali condizioni, se tutte fossero rispettate si potrebbe concludere che la domanda ha una risposta affermativa; se anche una sola fosse invece violata, la risposta non potrebbe che essere negativa. È necessario sottolineare che stabilire delle condizioni di verità opportune è un'operazione molto complessa. A seconda delle scelte è infatti possibile giungere ad una risposta positiva piuttosto che ad una negativa. È dunque inevitabile che la determinazione di tali condizioni di verità possa essere essa stessa oggetto di dibattito.

Nel caso in questione si ricercano le condizioni di verità per cui sia possibile affermare che un cervello è un computer. Come detto, la risoluzione di questo quesito si fonda su un'analisi accurata e precisa dei termini. Le condizioni sono poste dunque principalmente sulla definizione del concetto di computer. È necessario che la definizione sia quanto più puntuale possibile. Si evitano perciò da un lato definizioni troppo generali, dall'altro definizioni troppo stringenti.

**Rischi
nell'utilizzo di
definizioni
troppo generali**

Nel primo caso, infatti, se si usasse una definizione troppo generica, ne potrebbe conseguire che ogni oggetto risulti essere un computer. In questo modo la definizione non sarebbe utile in quanto fallirebbe nell'indicare quelli che sono i tratti distintivi di un computer e nell'isolare questa categoria dal resto del mondo. Non solo: se la definizione fosse talmente generale da includere tra i computer anche il cervello, la domanda "il cervello è un computer?" risulterebbe banalmente vera; ma in questo caso rientrerebbero nella definizione anche altri oggetti che generalmente non vengono identificati come computer: per assurdo potrebbe accadere che macchine, come un televisore o un'automobile, o addirittura altri oggetti non considerati macchine, come una sedia o come le pagine su cui è stampata questa tesi, si debbano considerare dei computer. È evidente che in questo caso la definizione di computer sarebbe inutile in quanto si riferirebbe ad una classe troppo ampia di oggetti eterogenei senza riuscire a separare i computer dal resto della realtà.

**Rischi
nell'utilizzo di
definizioni
troppo
particolari**

Nel secondo caso, se si fornisse una definizione troppo limitata del concetto di computer, si potrebbe correre il rischio che le macchine comunemente indicate come computer non rientrino in questa definizione. Si pone come requisito minimo che alla classe dei computer debbano appartenere tutti i computer di cui si ha esperienza quotidianamente, come desktop e laptop. Più precisamente, rifacendosi ad una descrizione di computer intuitiva e forse semplicistica, richiedere che tutti gli oggetti generalmente chiamati computer facciano parte della classe dei computer significa che la definizione dovrebbe abbracciare ogni sistema in grado di eseguire computazioni. In questo gruppo rientrerebbero dunque non solo i calcolatori elettronici sopra citati, ma tutti quei dispositivi che indipendentemente dalla complessità e dalle tecnologie adottate sono creati per eseguire computazioni.

Considerare la complessità irrilevante significa ammettere che nella classe dei computer possano rientrare anche dispositivi in grado di eseguire computazioni ma il cui potere di calcolo sia più limitato rispetto ai moderni computer; è il caso di macchine calcolatrici o, addirittura, di

strumenti come l'abaco.

Considerare le tecnologie irrilevanti significa che la definizione di computer dovrebbe essere sufficientemente estensiva da comprendere non solo i comuni calcolatori elettronici ma anche altre implementazioni di computer come i calcolatori meccanici, quali l'abaco o il *Difference Engine* e l'*Analytical Engine* di Babbage, o i calcolatori idraulici, come il computer idraulico sviluppato dal MIT (Blikstein, 2007). La lista potrebbe poi essere ulteriormente allungata considerando anche implementazioni d'avanguardia ancora oggetto di studio, come calcolatori ottici, chimici, quantistici o basati sul DNA (Frixione e Palladino, 2004).

In definitiva l'obiettivo sarà quello di evitare una definizione inadeguatamente astratta, tale per cui le implementazioni reali sopra citate non risultino parte della classe dei computer.

2.2 Definizione di macchina

Problema della definizione Quando si tenta di definire un ente una delle prime operazioni è quella di *predicare il genere*; si tratta cioè di identificare una classe di enti più ampia all'interno della quale rientri l'oggetto da definire, così da assegnargli una collocazione ben precisa e delineare i primi caratteri comuni e differenti rispetto ad enti di altri generi.

Ora, quando si fa riferimento ad un computer, i primi termini usati nel tentativo di darne una definizione sono generalmente *macchina* (*machine*) o *dispositivo* (*device*). Ad esempio si considerino alcune definizioni di computer:

*Computer. A device that consists of one or more associated processing units and peripheral units [...] that can perform substantial computations [...]*¹

(Geraci e Wilson, 1995 : 18)

¹ *Computer. Un dispositivo che consiste di una o più unità di elaborazione e di unità periferiche [...] in grado di realizzare computazioni considerevoli [...]* (trad. propria)

*Computer. Any **device** capable of solving problem by accepting information, applying prescribed operations to the information, and supplying the results of these processes [...]*²

(Fritz, 1963 : 154)

*Un calcolatore moderno è una **macchina** elettronica [...] in grado di eseguire calcoli in modo automatico, [...] ovvero di prendere in ingresso informazioni in formato digitale (o numerico), di elaborarle (o processarle) secondo le regole stabilite da un elenco predefinito di istruzioni macchina, e infine di mandare in uscita i risultati dell'elaborazione [...]*

(Hamacher et al., 2002 : 2)

Macchina e dispositivo I termini *dispositivo* e *macchina*, usati per descrivere la classe di oggetti nella quale rientrano i computer, vengono considerati in questa tesi come sinonimi; da questo momento ci si riferisce ad un computer sempre con il vocabolo *macchina*.

Definizione di macchina Con il termine *macchina* si definisce un *sistema fisico che realizza un processo causale*. Il termine *sistema* designa un *insieme di entità che costituiscono un'unità, in cui ciascun componente interagisce con altri componenti e concorre a realizzare un processo causale*. Il riconoscimento dell'esistenza di un insieme di componenti distinti e separati dall'ambiente è il presupposto per la definizione di un sistema. L'atto di individuazione di un sistema è un'azione arbitraria, dipendente dal giudizio di un osservatore o dall'accordo di più soggetti. Generalmente, un insieme di enti o aspetti della realtà sono raggruppati in un sistema da parte di un agente in base ad un criterio relazionale, ovvero in seguito al riconoscimento dell'esistenza di legami di qualche natura tra gli oggetti studiati. Se rappresentassimo la realtà oggetto di studio come un grafo, il procedimento di identificazione di un sistema potrebbe essere paragonato ad un'operazione di partizionamento: considerando solo certi tipi di relazione (generalmente di tipo causale) è possibile schematizzare la realtà analizzata appunto come un grafo, in cui

² *Computer. Qualsiasi **dispositivo** capace di risolvere un problema accettando informazioni, applicando operazioni preimpostate alle informazioni e restituendo il risultato di questi processi [...]* (trad. propria)

i vertici costituiscono gli elementi della realtà e gli archi costituiscono le relazioni oggetto di studio; quando un sottoinsieme di vertici possiede molte relazioni interne al sottoinsieme e poche relazioni esterne è possibile isolarlo e considerarlo un sistema; si noti che questa è solo una possibilità: le valutazioni sul numero di archi posseduti da un dato sottoinsieme e la decisione di considerare tale sottoinsieme di vertici un sistema sono tutte operazioni arbitrarie. Non vi è nulla nella realtà, per come è stata presentata l'operazione di definizione di un sistema, che imponga che un certo sottoinsieme di enti costituisca necessariamente un sistema.

**Vincoli sulla
definizione di
macchine**

Nel definire un *sistema* come *un'insieme fisico di parti in grado di realizzare un processo causale*, sono esplicitamente posti al concetto di sistema alcuni vincoli.

Si è imposto che il sistema di cui si tratta sia un sistema fisico: dopotutto, trattando di oggetti concreti e materiali come il computer e il cervello sarebbe stato superfluo considerare anche un'accezione astratta del concetto di *sistema*; si stabilisce dunque che il sistema sia un'entità la cui identità e i cui confini siano fisicamente distinguibili.

Si è imposto, poi, che il *sistema realizzi un processo causale*; si definisce la *causalità*, in maniera elementare ed intuitiva, senza l'ardire di addentrarsi in tale problema, come *la connessione tra due eventi, in virtù della quale il secondo consegue dal primo*. Dunque due enti sono tra loro in relazione causale se esiste una connessione di tipo razionale-deduttiva (per cui la causa è la ragione dell'effetto) o temporale-empirica (per cui esiste costanza e uniformità nel rapporto di successione tra i due enti). Affermare dunque che il *sistema realizzi un processo causale* significa asserire che tutti i prodotti o i risultati del sistema sono determinati da dinamiche causali.

**Limiti della
definizione di
macchine**

È evidente che la definizione di macchina appena data è estremamente generale ed insufficiente per definire con precisione un computer. Anche se non è necessario aggiungere ulteriori dettagli sulla fisicità di un computer dopo averlo definito una macchina, risulta tuttavia indispensabile fare delle precisazioni sulla sua architettura logica o sulla sua funzione. Questo è inevitabile dal momento che esistono numerosi tipi di macchine, come ad esempio gli elettrodomestici, tutti in grado di realizzare processi causali, ma pochi dei quali possono considerarsi dei computer.

Poiché si intende fornire una definizione di computer che sia indipendente dalla sua implementazione, non sembra opportuno porre dei

Definizione di macchina in base all'architettura limiti attraverso la descrizione dell'architettura di un calcolatore. Ad esempio, affermare che un computer sia la realizzazione fisica di una macchina di von Neumann, ovvero una macchina seriale costituita a livello elementare da una CPU, una memoria, un bus e da periferiche ausiliarie di input ed output, vincola eccessivamente la definizione di computer; infatti, in primo luogo, computer strutturalmente differenti, come potrebbero essere reti neurali o calcolatori basati sul DNA (Frixione e Palladino, 2004), potrebbero non rientrare nella definizione di computer o richiederebbero l'enunciazione e la dimostrazione di un principio di equivalenza con una macchina di von Neumann. In secondo luogo, ma non meno importante, sembra difficile ridurre il cervello all'architettura di una macchina di von Neumann; se questo non fosse possibile risulterebbe estremamente arduo, almeno a livello strutturale, istituire un raffronto tra cervello e computer.

Definizione di macchina in base alla funzione Sembra invece più accettabile l'idea di definire un computer delineandone la funzione. Questo permetterebbe infatti di fornire una descrizione di computer che sia indipendente da qualsiasi struttura fisica o da qualsiasi architettura. Tuttavia, la grande libertà implementativa unita alla difficoltà di identificare con precisione la funzione di un computer rischiano di condurre ad una definizione eccessivamente vaga, nella quale possano rientrare moltissimi oggetti, se non addirittura tutti gli oggetti esistenti. Se infatti ci si limita a considerare i computer come *macchine in grado di eseguire computazioni*, qualsiasi oggetto, come ad esempio il celebre muro di Searle o il secchio di Hinckfuss (Copeland, 1996), potrebbero considerarsi dei calcolatori; questa è infatti una conseguenza delle seguenti osservazioni proposte da J. R. Searle (1990, 1992):

1. Per qualsiasi oggetto esiste una qualche descrizione dell'oggetto per cui l'oggetto risulta essere un computer digitale.
2. Per qualsiasi programma e qualsiasi oggetto sufficientemente complesso, esiste una descrizione dell'oggetto per cui l'oggetto risulta implementare il programma.

È possibile allora trovare all'interno di un muro configurazioni di molecole isomorfe alla struttura formale di Wordstar ed affermare che tale muro implementi Wordstar (muro di Searle). Parimenti, è possibile trovare tra le molecole in un secchio d'acqua esposto al sole delle configurazioni per cui sia corretto concludere che il secchio stia eseguendo una computazione (secchio di Hinckfuss). Ora, sebbene entrambi questi argomenti si prestino a critiche, essi tuttavia esortano a considerare con attenzione la definizione della funzionalità di un computer, evitando di adottare descrizioni eccessivamente semplicistiche

o addirittura inappropriate.

Sottoinsiemi del concetto di macchina Date le difficoltà appena enunciate, prima di passare all'analisi della funzionalità di un computer, sembra più opportuno concentrarsi ancora per qualche tempo sulla classe delle macchine cercando di affinarne la definizione. Avendo scartato l'analisi architetture e avendo momentaneamente sospeso lo studio della funzione di un computer, si è deciso di analizzare con maggior attenzione il concetto di macchina, sia esaminando una possibile partizione delle macchine tra *macchine artefatto* e *macchine naturali*, sia considerando le macchine come *sistemi dinamici*.

Nel considerare le macchine artefatto e le macchine naturali si valuta sia l'adeguatezza di queste classi nel definire computer e cervello sia i confini delle stesse. Si cerca cioè di identificarne i limiti fisici per poter poi distinguere con facilità in che modo la macchina interagisca con il mondo. Definire il contorno è un risultato di importanza capitale: solo dopo questo atto è possibile distinguere l'oggetto di studio dal resto della realtà e riconoscerne l'identità, ovvero capire cosa è parte dell'oggetto di studio e cosa non lo è. Inoltre, dopo aver separato la macchina dall'ambiente, è possibile distinguere efficacemente e fondatamente quali sono i suoi input e i suoi output e studiare così in che modo la macchina si relazioni con il mondo esterno.

2.2.1 Macchina artefatto

Definizione di artefatto Il primo sottoinsieme di macchine che si considera è quello delle *macchine artefatto*. Rientrano in questa categoria tutte le macchine create dall'uomo, o più precisamente, secondo la definizione di *artefatto*, tutte le *macchine fatte ad arte*, ovvero plasmate, interamente o parzialmente, dall'uomo secondo un preciso disegno; è quindi indispensabile che antecedentemente alla realizzazione di un artefatto esista un progetto. Dal momento che raramente l'uomo agisce privo di un disegno o di una finalità, sembrerebbe che il concetto di macchina artefatto non aggiunga molto alla definizione di computer. In verità l'introduzione del concetto di progetto apporta un contributo fondamentale in quanto permette di definire un oggetto in base all'idea secondo cui era stato concepito. Se un artefatto è un'implementazione valida di un'idea, si possono identificare limiti e funzioni dell'artefatto nel progetto stesso. Infatti, quando un uomo concepisce un artefatto,

definisce già implicitamente quali sono i confini di tale oggetto. Si assumerà dunque che i limiti di un artefatto siano quelli definiti dall'artefice; ciò risulta particolarmente conveniente, in quanto facilita enormemente il compito di identificazione dei confini e, conseguentemente, di distinzione degli input e degli output.

Limiti della definizione di artefatto È evidente che tale definizione si applica perfettamente al concetto di computer, ma risulta del tutto inadeguata nel caso del cervello: non solo perché è inappropriato considerare il cervello un prodotto artificiale, ma anche perché non sembra sensato (se non all'interno di qualche sistema fondato sulla fede) affermare che il cervello sia l'implementazione di un progetto. La classe di macchina artefatto è dunque esclusiva per il concetto di computer; per la definizione di cervello sarà invece necessario ricorrere ai concetti seguenti.

2.2.2 Macchina naturale

Definizione di macchina naturale Complementare alla classe delle macchine artificiali è ovviamente la classe delle macchine naturali. Intuitivamente, in tale classe è possibile far rientrare tutti quei *sistemi fisici che realizzano un processo causale che non siano stati creati o modificati dall'uomo*. È evidente che, data la generalità della definizione di macchina adottata, la classe delle macchine naturali contenga al proprio interno numerosi oggetti che possono considerarsi macchine.

Confini di una macchina naturale Carattere comune di tutte le macchine naturali è quello di non avere confini ben precisi e inequivocabili; differentemente dagli artefatti, non esistendo (o comunque non essendo in alcun modo noto) un progetto a cui tali macchine si conformano, è necessario definirne i confini in modo arbitrario; è cioè rimessa all'uomo la decisione di stabilire quali siano i limiti di una macchina naturale. Si noti che affermare che la definizione dei confini è un'operazione *arbitraria* non significa che qualsiasi scelta di confini ha il medesimo valore; *arbitraria* significa soltanto che tale delimitazione non esiste implicitamente in natura ma è imposta dall'uomo; è poi compito di quest'ultimo effettuare una scelta arbitraria e *razionale* per il riconoscimento di tali confini, così da ottenere una delimitazione coerente di una data macchina naturale; una definizione *arbitraria e irrazionale*, oltre a non essere condivisa, rischia infatti di condurre a delle conclusioni inconsistenti (ad esempio, se si considerasse come macchina naturale un bacino idrico, sarebbe una scelta *arbitraria e*

razionale quella di considerare come confini della macchina i confini dettati dal bacino e dai suoi affluenti; sarebbe, invece, una scelta *arbitraria e irrazionale* quella di considerare, senza giustificazione alcuna, il bacino e solo una parte dei suoi affluenti).

Un'ulteriore classificazione delle macchine naturali che si potrebbe adottare per facilitarne lo studio, sarebbe quella che distingue *macchine naturali biologiche* e *macchine naturali non-biologiche*.

Definizione di macchina naturale biologica Tra le *macchine naturali biologiche* rientrano gli organismi e gli esseri viventi considerati come macchine; più precisamente, volendo comprendere in questa sottoinsieme tutte le macchine biologiche, si potrebbe ricorrere ad un concetto, quello di *autopoiesi*, spesso utilizzato per caratterizzare i sistemi viventi in biologia e rinominare quindi il sottoinsieme delle *macchine naturali biologiche* come sottoinsieme delle *macchine naturali autopoietiche*. *Autopoietico* significa letteralmente “*in grado di produrre sé stesso*” e, secondo la definizione data dai biologi Varela e Maturana (1980 : 78-79) *una macchina autopoietica è una macchina (definita come unità) organizzata come una rete di processi di produzione (trasformazione e distruzione) dei componenti tale che:*

Definizione di autopoiesi

1. *attraverso le sue interazioni e trasformazioni si rigeneri continuamente e realizzi la rete di processi (relazioni) che la costituisce;*
2. *costituisca la macchina come un'unità concreta nello spazio in cui i componenti esistono specificando la rete come il dominio topologico della sua realizzazione.*

Un sistema autopoietico è dunque tale se presenta proprietà di chiusura, auto-organizzazione e autonomia. L'esempio più evidente di sistema autopoietico è quello della cellula biologica; le cellule sono infatti organizzate in strutture definite (nucleo e organuli) racchiusi all'interno di un preciso confine (la membrana cellulare); questa struttura, scambiando molecole ed energia con l'ambiente, è in grado di produrre tutti quei componenti che, a loro volta, permettono il sussistere dell'organizzazione interna della cellula.

È evidente che nel concetto di autopoiesi è insito anche il concetto di limite: non può esistere una macchina autopoietica che non abbia un confine ben delimitato, in quanto senza tale confine non potrebbe costituire un'unità. Riuscire ad identificare una macchina autopoietica permette dunque di riconoscere conseguentemente anche i limiti del sistema stesso e distinguerne con facilità gli input e gli output, risolvendo così il problema dell'identificazione arbitraria dei confini.

**Limiti della
definizione di
macchina
naturale**

Il concetto di autopoiesi è dunque un concetto strettamente legato alla biologia. Esso è stato introdotto per tentare di dare una risposta a cosa caratterizzi l'organizzazione e la struttura degli esseri viventi. Nonostante i tentativi di applicare il concetto di *autopoiesi* in altri campi, come in economia o nello studio dei fenomeni sociali, risulta molto arduo usarlo in riferimento ai computer (Boden, 2000). I moderni computer infatti non soddisfano la definizione di autopoiesi: essi potrebbero considerarsi sistemi chiusi, ma mancano delle fondamentali proprietà di auto-organizzazione ed autonomia; organizzazione e autonomia di un computer sono infatti imposte esternamente dall'uomo.

Tuttavia, sebbene non applicabile al concetto di computer, la definizione di macchina naturale autopoietica continua a rivestire un notevole interesse in quanto potrebbe essere applicata al cervello. In verità l'attribuzione di *autopoiesi* non ad un organismo nel suo insieme, ma ad una sua sottoparte, ovvero un organo, potrebbe essere oggetto di dibattito. Intuitivamente il cervello (o, se necessario, per estensione, l'intero sistema nervoso) sembrerebbe dotato di quelle proprietà di chiusura, auto-organizzazione e autonomia richieste ad un sistema autopoietico; in base a questa considerazione e avendo affermato che l'estensione della classe delle macchine naturali biologiche coincide con quella delle macchine naturali autopoietiche, in questa tesi si assume che il cervello sia una macchina autopoietica e sia dunque possibile ricorrere al concetto di autopoiesi per evidenziarne i confini. D'altra parte, qualora fosse smentito, a livello biologico o concettuale, che il cervello sia una macchina autopoietica, sarà possibile considerare il cervello come una macchina naturale biologica non autopoietica ed identificarne arbitrariamente i confini.

**Definizione di
macchina
naturale
non-biologica**

Il secondo sottoinsieme delle macchine naturali è quello costituito dalle *macchine naturali non-biologiche*. Esso è il complemento del sottoinsieme precedente ed è costituito da tutte le macchine non viventi. Per l'identificazione dei confini delle macchine naturali non-biologiche valgono le considerazioni date in generale per le macchine naturali: è necessario cioè un riconoscimento dei limiti *arbitrario e razionale* senza la possibilità di ricorrere a concetti come quello di autopoiesi. Oltre l'esempio elementare riguardo l'identificazione dei confini di un bacino idrico riportato sopra, non si fanno ulteriori considerazioni sui criteri di *arbitrarietà e razionalità* che dovrebbero determinare la delimitazione di una macchina naturale non-biologica; infatti, dal momento che né il computer né il cervello rientrano in questo sottoinsieme, tale problema risulta ininfluenza rispetto allo scopo della tesi.

2.2.3 Macchina come sistema dinamico

Definizione di sistema dinamico

Un approccio differente allo studio e alla definizione di una macchina può essere quello che si basa sul concetto di *sistema dinamico*. Con *sistema dinamico* si intende un modello matematico di un oggetto che interagisce con la realtà esterna per mezzo di insiemi di ingressi ed uscite e che possiede una sua precisa configurazione interna, definita da uno stato; la relazione tra ingressi ed uscite, mediata dal concetto di stato, è descritta per mezzo di una relazione matematica.

Il termine sistema dinamico è per la verità piuttosto ingannevole, in quanto spesso utilizzato sia per riferirsi alla realtà modellata da un dato sistema, sia al modello matematico vero e proprio. In questa tesi, si fa uso del termine di sistema dinamico nella prima accezione: non essendo interessati alla formulazione e allo studio di precisi modelli matematici, si usa il termine sistema dinamico per fare riferimento ad un oggetto della realtà che è descrivibile secondo il formalismo dei *sistemi dinamici*.

Introducendo il concetto di sistema dinamico, si è affermato che esso costituisce un approccio alternativo allo studio delle macchine; mentre nei paragrafi precedenti si sono analizzate le macchine da una prospettiva funzionale, studiare una macchina come un sistema dinamico significa analizzarle invece da un punto di vista matematico e modellistico. Questo metodo di studio può essere applicato indistintamente a tutte le sottoclassi di macchina identificate: è cioè possibile descrivere come sistema dinamico sia una macchina artefatto sia una macchina naturale, biologica e non. Il concetto di sistema dinamico non serve dunque a raffinare la definizione di macchina, quanto piuttosto a proporre un punto di vista alternativo per lo studio e la descrizione delle macchine.

Limiti della definizione di sistema dinamico

La delimitazione di un sistema dinamico, così come la precedente definizione di macchina, comporta implicitamente la definizione dei confini di un oggetto; il riconoscimento di ingressi e di uscite presuppone infatti l'identificazione di limiti che garantiscano l'esistenza stessa degli ingressi e delle uscite. In quanto tale, la definizione di una macchina come sistema dinamico non apporta un contributo significativo per la delimitazione di un sistema giacché assume gli stessi confini già definiti dal concetto di macchina.

2.3 Attribuzione di funzione

Definizione di attribuzione di funzione

Per poter riprendere lo studio della *funzione* di un computer, abbandonato a favore dell'esame delle macchine artefatto e naturali e dei sistemi dinamici, è necessario introdurre almeno sommariamente alcuni concetti propri dell'ambito istituzionale-sociale, come il concetto di *attribuzione di funzione*. Per affrontare consistentemente il problema della funzione di un computer è infatti inevitabile comprendere cosa significhi attribuire una funzione e in quale modo questo procedimento si realizzi.

Rifacendosi ad una certa forma di realismo (Searle, 1998), si definisce *attribuzione di funzione* la capacità dell'uomo (e di alcuni animali di livello superiore) di utilizzare certi oggetti come strumenti. Riconoscere ad un ente la natura di mezzo o utilizzarlo per raggiungere un determinato scopo significa attribuire a tale ente una *funzione*.

Una prima osservazione che segue da questa definizione è che l'attribuzione di funzione è un'operazione dipendente dal soggetto: una funzione esiste solo in quanto riconosciuta da un agente. Questa affermazione si fonda sull'assunzione di una visione non-teleologica della natura, secondo la quale in natura non esistono funzioni o fini; ciò che l'uomo scopre e studia in natura sono legami di causa-effetto, le funzioni o le finalità esistono solo in relazione ad un osservatore. Sinteticamente, le cause sono osservatore-indipendenti, le funzioni sono osservatore-dipendente, nel senso che le prime esistono indipendentemente dall'uomo, mentre le seconde sono determinate da un osservatore.

Categorizzazione delle attribuzioni di funzione

A seconda del loro rapporto con l'oggetto cui si riferiscono è possibile distinguere due differenti forme di attribuzione di funzione:

- *Attribuzione di funzione d'uso*: si definisce *funzione d'uso* una funzione attribuita ad un oggetto, la cui fisica è sufficiente a realizzare la funzione stessa (ad esempio, una penna, la cui funzione di scrittura è garantita dalla natura stessa della penna);
- *Attribuzione di funzione di status*: si definisce *funzione di status* una funzione attribuita ad un oggetto, la cui fisica non è sufficiente a realizzare la funzione stessa (ad esempio, una

banconota, la cui funzione di potere di scambio non è determinata dalla natura del pezzo di carta che costituisce la banconota).

È possibile cioè classificare le attribuzioni di funzioni tra funzioni garantite dalla struttura fisica dell'oggetto cui sono riferite e funzioni che richiedono invece un'accettazione o un riconoscimento da parte di una comunità di riferimento. Per queste ultime, può essere proposta un'ulteriore tassonomia, volta a facilitarne lo studio:

- *Attribuzione di funzione di status semantica*: si definisce *funzione di status semantica* l'attribuzione ad un oggetto del potere di significare qualcosa (ad esempio, una parola, la cui funzione di significare è data dalla semantica della parola);
- *Attribuzione di funzione di status deontica*: si definisce *funzione di status deontica* l'attribuzione ad un oggetto di proprietà deontiche (ovvero poteri, obblighi, diritti e così via).

La presente tesi conduce a considerare con maggiore attenzione *attribuzioni di funzione di status semantiche*, in particolare in relazione agli input e agli output di un computer.

Definito il concetto di attribuzione di funzione, è ora possibile tentare di applicarlo ai termini precedentemente definiti e proporre delle riflessioni nel caso di macchine artefatto e macchine naturali.

**Attribuzione di
funzione negli
artefatti**

Come già detto, una macchina artefatto è un oggetto creato secondo il preciso progetto di un uomo; questo significa che nel momento in cui viene concepito, l'artefatto ha già una propria funzionalità ben delineata. In generale l'attribuzione di funzione data dal creatore di un artefatto è la più rilevante e significativa; essa infatti è definita a priori ed è il fine per cui l'artefatto viene creato. Dunque, sebbene chiunque possa attribuire ad un artefatto una funzione specifica in un qualsiasi momento, l'attribuzione di funzione proposta dall'artefice sarà sempre considerata in questa tesi la funzione di riferimento.

**Attribuzione di
funzione negli
organismi**

Nel caso di una macchina naturale parlare di *funzione* richiede di considerare separatamente macchine naturali biologiche e macchine naturali non-biologiche. Nel caso di queste ultime si assume che una macchina naturale non biologica (ad esempio, un fiume) non abbia una funzione implicita, ma al più una assegnatagli dall'uomo (ad esempio, nel caso del fiume, l'utilizzo dell'acqua per il raffreddamento in uno scambiatore di calore). Il caso delle macchine naturali biologiche risulta

invece molto più problematico; su questo punto infatti il pensiero di biologi e filosofi è spesso discorde. In particolare quando si fa riferimento non tanto ad un organismo nel suo insieme, ma più specificatamente ad un organo come il cervello, diviene difficile accordarsi sul concetto di *funzione*. Per come è stata precedentemente definita l'attribuzione di funzione, essa risulta essere inevitabilmente agente-dipendente; coerentemente con questa affermazione, secondo Searle (1998), anche gli organi umani non hanno, dunque, implicitamente una funzione (ad esempio, è oggettivo che il cuore causa il pompare del sangue e questo a sua volta è all'origine di molte altre relazioni causali, ma affermare che la *funzione* del cuore sia quella di pompare il sangue è possibile solo in una teleologia che dà importanza alla vita e alla sopravvivenza). D'altra parte, i biologi sono invece convinti che la funzione di un organo non sia agente-dipendente, ma sia una proprietà oggettiva di un organo; la loro spiegazione si fonda sui concetti di *evoluzione* e *fitness* (Rosenberg e Bouchard, 2002). Con *fitness* si intende la capacità di un organismo di adempiere il suo obiettivo primario, ovvero quello di propagare il proprio patrimonio genetico; con *evoluzione* si intende invece il manifestarsi di generazione in generazione di variazioni genetiche casuali; quelle mutazioni in grado di conferire vantaggi adattativi agli organismi tendono poi ad avere una maggiore probabilità di propagazione. *Fitness* ed *evoluzione* permettono agli esseri viventi di svilupparsi e di adattarsi alla grande complessità e diversità dell'ambiente naturale. Ora, in quest'ottica, la funzione di qualsiasi organo è subordinata al fine ultimo di favorire la propagazione del proprio patrimonio genetico; anche il cervello avrebbe perciò intrinsecamente questa funzione. Differentemente da certe forme di pensiero realista, i biologi sono dunque propensi a riconoscere l'esistenza di una funzione intrinseca negli organismi, indipendentemente da qualsiasi attribuzione osservatore-dipendente.

2.4 Rappresentazione

Definizione di rappresentazione

Un secondo concetto di importanza fondamentale per poter studiare il rapporto tra cervello e computer è quello di *rappresentazione*; esso permette di analizzare in quale modo possa essere attribuito un significato a dei simboli e in quale modo essi possano riferirsi al mondo reale. Con il termine *rappresentazione* si vuole infatti indicare ciò che in filosofia, e in particolare da alcuni pensatori come Anscombe e Searle, è definita *intenzionalità*; si è tuttavia deciso di utilizzare il termine

rappresentazione in quanto il vocabolo *intenzionalità*, eredità della tradizione e risultato di una traduzione dal tedesco poco felice, risulta in italiano ingannevole; si definisce, infatti, *stato intenzionale*, o *rappresentazione* in seguito, uno stato mentale che sia in relazione con il mondo esterno; si tratta dunque della facoltà attraverso cui la mente è in grado di dirigersi verso gli oggetti e gli elementi della realtà. L'argomento di maggiore interesse riguardo le rappresentazioni è perciò la loro capacità di significare qualcosa; a questo riguardo è possibile distinguere due tipi fondamentali di rappresentazioni:

**Categorizzazione
delle
rappresentazioni**

- *Rappresentazioni intrinseche*: si definiscono *rappresentazioni intrinseche* quelle rappresentazioni che si realizzano e assumono un contenuto all'interno della mente dell'uomo per mezzo della coscienza; in quanto tali, esse sono sempre dipendenti dal soggetto, ma indipendenti dagli osservatori (ad esempio, avere fame, per cui l'appetito dipende esclusivamente dal soggetto, indipendentemente da qualsiasi osservatore)
- *Rappresentazioni derivate*: si definiscono *rappresentazioni derivate* quelle rappresentazioni cui è attribuito un contenuto da parte di un gruppo di osservatori; in quanto tali, esse sono sempre dipendenti dagli osservatori, che assegnano un significato basandosi sulle proprie rappresentazioni intrinseche (ad esempio, una frase, il cui significato esiste solo in relazione agli osservatori)

Un errore comune è quello di attribuire impropriamente la capacità di avere rappresentazioni anche a cose inanimate; tale attribuzione è tuttavia metaforica, modellata in analogia con l'uomo (ad esempio, l'aver fame di sostanze nutritive attribuito ad una pianta); quest'ultimo uso del concetto di rappresentazione non è corretto e non sarà dunque tenuto in considerazione nel seguito.

**Rapporto tra
rappresentazione
e realtà**

Secondo la *teoria corrispondentista della verità*, le rappresentazioni sono legate alla realtà da una relazione semantica (Marian, 2005; Colombetti, 2003); esse hanno cioè un *contenuto*, che si riferisce ad una parte di realtà, ed un *tipo* (o *modo intenzionale*), che descrive in quale modo una rappresentazione si relaziona alla realtà (ad esempio, credenza, speranza, desiderio).

A seconda del *tipo* è possibile identificare due diverse *direzioni di adattamento* tra rappresentazione e realtà:

- *adattamento mente-a-mondo*, secondo cui la rappresentazione di un osservatore si adatta alla realtà (ad esempio, il conoscere una

caffettiera: la rappresentazione della caffettiera nella mente del soggetto è modellata secondo la caffettiera già esistente);

- *adattamento mondo-a-mente*, secondo cui, in maniera opposta, il mondo si adatta alla rappresentazione di un osservatore (ad esempio, il progettare una caffettiera: la rappresentazione della caffettiera esiste nella mente del soggetto e sarà il modello di riferimento per una futura caffettiera reale).

Dal *tipo* e dalla *direzione di adattamento* conseguono infine condizioni di verità, realizzazione o, più in generale, *condizioni di soddisfazione*; ogni rappresentazione possiede infatti delle condizioni in base alle quali è possibile affermare che essa rappresenti o meno una parte di realtà (Searle, 1998).

Ora, basandosi sulle definizioni date, è finalmente possibile sia analizzare la funzione di un calcolatore sia affrontare rigorosamente e precisamente il confronto tra i concetti di computer e cervello.

Capitolo 3 – Il computer

Nel precedente capitolo sono state poste le basi per affrontare la definizione del concetto di computer: si è stabilito il *genere*, ovvero si è dimostrata l'appartenenza del computer alla classe delle *macchine*; sono state considerate alcune sottoclassi in grado di descrivere con maggiore precisione un computer, tra cui quella di *macchina artefatto*; si è cercato di predicare le *differenze specifiche* di un computer, ovvero quei caratteri che lo distinguono da tutti gli altri oggetti appartenenti alla medesima classe; si è scartata un'analisi fondata sulla struttura fisica o sull'architettura logica dal momento che una simile caratterizzazione sarebbe risultata piuttosto sterile quando confrontata con il cervello; si è proposto di fornire una definizione basata sulla funzione realizzata da computer e cervello. In questo capitolo sarà dunque approfondito lo studio di questa funzione per poter infine giungere ad una definizione completa e adeguata di computer.

3.1 Limiti di un computer

Delimitazione di un computer

Sulla base dei risultati precedenti si è definito il computer come *macchina artefatto*. In quanto *oggetto plasmato ad arte e secondo un progetto*, un computer presenta dei limiti ben precisi e netti. Tali confini costituiscono le interfacce per mezzo delle quali è possibile interagire con un computer e attraverso le quali transitano input ed output; gli esempi più manifesti di questi confini sono costituiti da quelle che vengono comunemente definite periferiche di input (ad esempio, una tastiera) e periferiche di output (ad esempio, uno schermo).

Si vuole sottolineare che la delimitazione di un computer risulta tanto lineare non per la semplicità intrinseca nel riconoscimento dei limiti, ma semplicemente per le considerazioni precedenti. In verità riconoscere i limiti di un oggetto significa prima di tutto essere in grado di riconoscerne l'identità, operazione tutt'altro che banale. La facilità con cui è stato possibile identificare i confini deriva esclusivamente dalla definizione che è stata data di *macchina artefatto*, che permette di rifarsi al progetto dell'oggetto per constatarne l'identità e dunque i limiti.

Attraverso le periferiche di input ed output un computer è in grado di

Input e output di un computer ricevere e restituire *simboli*, dove con *simbolo* o *segno* si identifica un qualsiasi artefatto a cui è attribuito un significato. Più precisamente, un computer è in grado di gestire *input ed output simbolici*, ovvero una sequenza di entità semplici, spesso dette *token*, cui è attribuito un significato. Una *sequenza di token* è una sequenza di elementi atomici, entità minime nel momento dell'analisi; la natura di tali token non è argomento di interesse, in quanto strettamente in relazione con la tecnologia con cui è realizzato un computer: potrebbero essere differenze di potenziale in un computer elettronico così come pressioni all'interno di un computer idraulico. Si sottolinea che il termine *simbolico* vuole indicare proprio input o output generici a cui può essere attribuito un significato e non ha alcuna relazione con gli input o gli output del calcolo simbolico. L'attribuzione di significato a tali token avviene secondo i principi di un'*attribuzione di funzione*, o più correttamente di una *attribuzione di funzione di status semantica*: attraverso l'accettazione condivisa è attribuito ad ogni *token* o insieme di *token* un preciso significato comprensibile per un operatore umano; è questo il modo in cui in un comune computer elettronico si assegna un significato ad input ed output: prendendo in considerazione delle differenze di potenziale e utilizzando una codifica condivisa (ovvero una serie di *attribuzioni di funzione di status semantica* concordate da più osservatori) è possibile far corrispondere, ad esempio, un insieme di voltaggi ad un valore numerico o ad una lettera.

3.2 Categorizzazione delle macchine per input ed output

Si potrebbe ora supporre che ciò che distingue un computer da altre macchine artefatto sia proprio la natura degli input che riceve e degli output che produce. Per poter sostenere questa ipotesi è però necessaria una più accurata analisi di input ed output.

Input e output fisici Nel caso delle macchine artefatto e, conseguentemente, dei computer è possibile distinguere due forme principali di input ed output: *fisici* o *simbolici*. Si definiscono *input o output fisici* gli input o output materiali, risultati di forze concretamente applicate da un agente o esercitate da una macchina; non sono oggetto di alcuna *attribuzione di funzione di status semantica* e il loro riconoscimento è indipendente da eventuali osservatori (ad esempio, un indumento pulito, come output di una lavatrice, è privo di alcun significato condiviso). Si definiscono, invece, **Input e output simbolici** *input o output simbolici* gli input o output anch'essi materiali ed esercitati

o applicati fisicamente, ma latori di un significato; sono cioè oggetto di *attribuzioni di funzioni di status semantiche* e possiedono in quanto tali un significato (ad esempio, la temperatura, input di una moderna lavatrice, la cui regolazione avviene generalmente ruotando una manopola fino ad una determinata tacca; l'atto di posizionare la manopola su un preciso numero ha un valore simbolico e non fisico, in quanto, sebbene l'azione nel suo complesso sia materiale, essa non causa direttamente la variazione della temperatura; piuttosto, il riconoscimento della posizione della manopola e l'elaborazione di questa informazione da parte della macchina fa sì che la temperatura della lavatrice sia correttamente impostata).

**Categorizzazione
delle macchine**

Basandosi su questa distinzione tra input ed output fisici e simbolici è possibile suggerire una distinzione e una categorizzazione delle macchine secondo il tipo di ingressi ed uscite che esse gestiscono:

- *Trasduttori-sensori*: si definiscono *trasduttori-sensori* le macchine in grado di ricevere un input di tipo fisico e restituire un output di tipo simbolico (ad esempio, un termometro, il cui input fisico è costituito da un flusso di calore e il cui output simbolico è costituito da un valore numerico realizzato dall'altezza raggiunta da una colonna di mercurio su una scala graduata e leggibile da un osservatore).
- *Trasduttori-effettori*: si definiscono *trasduttori-effettori* le macchine in grado di ricevere un input simbolico e restituire un output di tipo fisico (ad esempio, un termostato, il cui input simbolico è costituito da un valore numerico impostabile da un agente su una scala tramite la rotazione di una manopola e il cui output fisico è un flusso di calore che mantiene costante la temperatura di un ambiente; oppure una televisione, il cui input simbolico è la selezione di un canale realizzato dalla pressione di un pulsante e il cui output fisico è un'immagine, non simbolica, ma fisica che riproduce iconicamente la realtà).
- *Bitrasduttori*: si definiscono *bitrasduttori* le macchine in grado di ricevere input simbolici e restituire altri output simbolici (ad esempio, un computer, il cui input simbolico è generalmente costituito da simboli realizzati dalla pressione di un tasto e il cui output simbolico è costituito a sua volta da simboli comunemente realizzati da un'immagine mostrata su uno schermo).

**Osservazioni
sulla
categorizzazione
della macchine:** La categorizzazione delle macchine proposta è chiaramente essenziale e concisa. Sebbene essa risulti utile e funzionale a chiarire la funzione di un computer e delle macchine in generale, è necessario fare qualche considerazione.

1. Relazione tra input ed output

Prima di tutto è fondamentale sottolineare che la partizione della macchine proposta è fondata solo su input ed output. Una definizione più precisa della classe di macchine richiede non solo l'identificazione di input ed output ma anche il riconoscimento di un legame tra input ed output; è necessario cioè che input ed output non siano tra loro scorrelati, ma vi sia una *funzione* o, più in generale, una *relazione* che li lega. Si richiede cioè che vi sia un nesso, materiale a livello fisico o logico a livello simbolico, tra input ed output; l'importanza di questa connessione sarà ripresa e analizzata nel corso del prossimo capitolo in relazione al funzionamento di computer e cervello.

2. Input ed output energetici

Si osservi che nel catalogare le macchine è stata trascurata in questa trattazione l'esistenza di input ed output di tipo energetico che permettono il funzionamento della macchina stessa (ad esempio, il flusso di corrente elettrica che alimenta un elettrodomestico o la dispersione di calore generata da computer). Questi input ed output sono dati per scontati, in quanto, sebbene fondamentali per il corretto funzionamento di ogni macchina, non concorrono a specificare la funzione di una macchina; essi costituiscono piuttosto un requisito necessario per il funzionamento stesso di ogni macchina.

3. Asimmetria tra sensori e effettori

Si noti anche che nonostante l'apparente complementarietà, esiste un'asimmetria tra *trasduttori-sensori* e *trasduttori-attuatori*.

In un sensore può esistere un unico input fisico in relazione con un unico output simbolico; è, cioè, idealmente possibile che l'input fisico sia direttamente collegato con un output fisico a cui è assegnato un significato. Si consideri il già citato caso di un termometro: la cessione di energia sotto forma di calore da parte di una stanza (input fisico del termometro) causa una variazione di volume del mercurio (output fisico del termometro); l'altezza finale della colonna di mercurio è poi messa in relazione con un valore (output simbolico del termometro); in conclusione, l'output del termometro è unico e ad esso è attribuito un significato da parte di un osservatore.

Un attuatore, invece, pur generando un output fisico, dispone generalmente di più input. Infatti, esso riceve tipicamente un input fisico connesso ad un output fisico analogamente a quanto avviene in un sensore; ma oltre a questo si aggiunge l'input simbolico, utilizzato per controllare la macchina; questo input simbolico non è ottenuto assegnando un significato al pre-esistente input fisico, ma si tratta quasi sempre di un input supplementare. Si consideri il già citato caso del termostato: esso riceve in ingresso la temperatura di una stanza,

esattamente come il termometro (input fisico del termostato), ed imposta la temperatura dell'ambiente secondo i desideri dell'utente (output fisico del termostato); ora l'input simbolico non è ottenuto assegnando un significato alla temperatura ricevuta in ingresso, ma aggiungendo un input completamente nuovo, l'input simbolico della manopola del termostato usato per controllare l'intera macchina.

L'asimmetria appena espressa diventa evidente confrontando la rappresentazione grafica del funzionamento di un sensore, offerta in figura 3.1, con quella del funzionamento di un effettore, esposta in figura 3.2.

Per completezza si evidenzia infine che nella precedente categorizzazione delle macchine sono assenti alcune classi considerate di secondaria importanza.

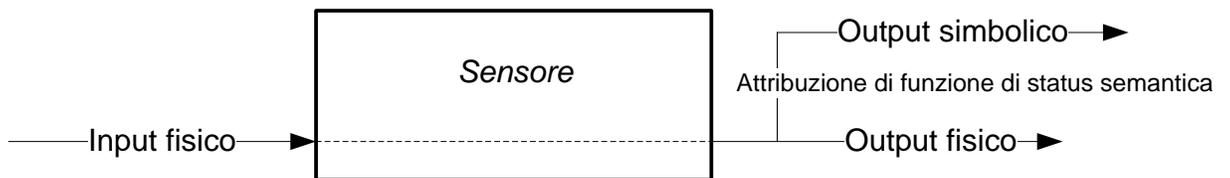


Figura 3.1 – Sensore

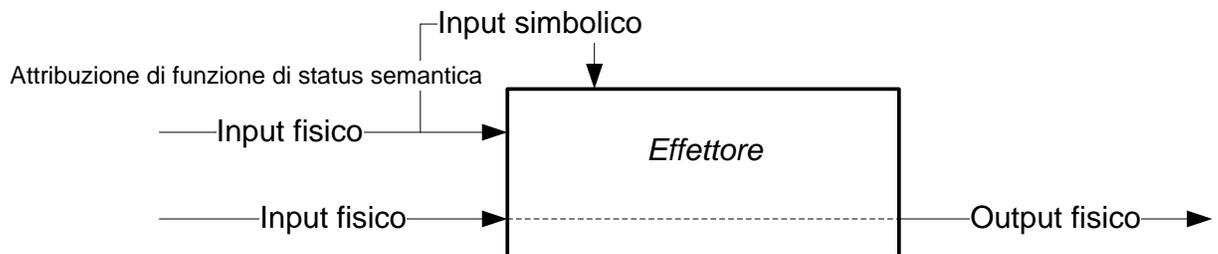


Figura 3.2 – Effettore

4. Classi di macchine di secondaria importanza

Si è, ad esempio, tralasciata la possibile classe di macchine che ricevono solamente input fisici e restituiscono solo output fisici; simili macchine non sono in grado di elaborare input o output dotati di un significato e dunque il loro studio risulta, per il fine proposto, sterile. Sono inoltre state trascurate anche altre possibili classi derivanti dalla composizione delle precedenti, come macchine in grado di ricevere sia input fisici che simbolici così come di restituire output fisici e simbolici; si tratta cioè di macchine dotate di controlli simbolici ausiliari per favorire l'interazione (ad esempio, macchine utensili dotate di controllo numerico). Come già visto raffrontando sensori ed effettori, questi ultimi risultano la maggior parte delle volte macchine dotate di un controllo. Per quanto riguarda i sensori, anch'essi potrebbero essere dotati di controlli in grado di gestire l'output; si ritiene tuttavia che la presenza di questo input ausiliario per controllare l'uscita non sia influente, nel senso che non è sufficiente per trasformare un sensore in un bitrasduttore: infatti, nonostante la presenza di uno o più controlli, finché l'output simbolico è una funzione dell'input fisico la macchina si può considerare un sensore.

Fatte queste precisazioni sulla categorizzazione delle macchine, e in particolare su sensori ed effettori, ora l'attenzione è volta alla classe di maggiore interesse per lo scopo di questa tesi, ovvero a quella dei bitrasduttori.

3.3 Definizione di computer

Nel paragrafo precedente si è esposta la definizione di una macchina bitrasduttrice e si è portato come esempio di tale dispositivo un computer. Si vuole ora mostrare che ogni computer risulta essere un'istanza di un bitrasduttore.

Il computer è un bitrasduttore

Si consideri dunque il computer come un bitrasduttore, ovvero come un generico oggetto in grado di ricevere degli input simbolici e restituire attraverso elaborazioni più o meno complesse altri output simbolici; tale definizione rispetta tutte le condizioni di sensatezza e verità precisate nel capitolo precedente: essa è infatti sufficientemente ristretta da non comprendere né oggetti comunemente non considerati computer (ad esempio, gli elettrodomestici) né il cervello che, come già detto, non si considera una macchina artefatto; inoltre è sufficientemente ampia da comprendere tutte le tipologie esistenti di computer indipendentemente da tecnologia e complessità; rientrano infatti nella classe di bitrasduttori sia

computer realizzati con tecnologie diverse (ad esempio, elettronici e meccanici) sia artefatti di potenza computazionale più limitata (ad esempio, una calcolatrice) fino al caso limite dell'abaco, che può considerarsi una macchina in grado di ricevere input simbolici (realizzati tramite il movimento di sfere che rappresentano delle quantità) e di restituire output simbolici (realizzati dalla posizione delle sfere che indica un risultato) sebbene con potenza computazionale praticamente nulla.

Si noti tuttavia che, anche se è possibile affermare che un computer è necessariamente un bitrasduttore, non è possibile tuttavia affermare il contrario, ovvero che ogni bitrasduttore è un computer. Quest'ultima affermazione è smentita, come si dirà nel prossimo capitolo, dal fatto che anche il cervello, che si considera per il momento differente da un computer, rientra nella classe dei bitrasduttori.

Uno studio più accurato del computer e del suo funzionamento come bitrasduttore permette di evidenziare con maggiore dettaglio il funzionamento dello stesso. È possibile distinguere due livelli ai quali un bitrasduttore opera: un livello *fisico* e un livello *simbolico*. La presenza di questi due differenti livelli di funzionamento è una conseguenza diretta dell'uso e dell'elaborazione di input ed output simbolici, che, come già spiegato, sono simbolici in quanto latori di significato, ma pur sempre realizzati fisicamente.

Livello fisico del computer

Vi è dunque in ogni computer un livello base, il livello fisico, all'interno del quale avviene l'elaborazione e la computazione vera e propria. Ogni input ricevuto, sia esso una tensione o il movimento di un ingranaggio, viene manipolato fisicamente in modo da fornire un output. Il risultato prodotto è la conseguenza dell'elaborazione sintattica e meccanica degli input ricevuti, senza alcun riferimento al loro possibile significato. Così ridotto, il computer risulta ben lontano da un bitrasduttore, più simile ad una semplice macchina in grado di ricevere input fisici e restituire output analogamente fisici; in questa condizione, non avrebbe senso parlare di simboli né fare il computer oggetto di un'*attribuzione di funzione di status semantica*; al contrario, un simile dispositivo è esclusivamente oggetto di un'*attribuzione di funzione d'uso* in quanto è la semplice meccanica della macchina a realizzare la funzione che lega l'input all'output.

Livello simbolico del computer

Il secondo livello è invece il livello simbolico, all'interno del quale input ed output assumono il significato che è loro comunemente assegnato. È a questo punto che attraverso un'opportuna *attribuzione di funzione di status semantica* l'input e l'output fisici sono messi in corrispondenza con un input ed un output simbolici. Tra input simbolico e output simbolico esiste ora una corrispondenza, fondata ovviamente sul legame tra input fisico ed

output fisico, ma a cui è possibile assegnare un significato; o meglio, è possibile interpretare il legame tra input simbolico ed output simbolico come la realizzazione di una funzione simbolica che, dato un valore simbolico, è in grado di restituire un risultato anch'esso simbolico.

Graficamente è possibile rappresentare un bitrasduttore con i suoi due livelli, fisico e simbolico, come schematizzato in figura 3.3. In questa immagine le frecce orizzontali nel livello fisico e nel livello simbolico rappresentano, rispettivamente, la presenza di un processo fisico e di una funzione simbolica che collegano input ed output. Le frecce verticali che mettono in relazione i due livelli identificano invece l'*attribuzione di funzione di status semantica* attraverso la quale è assegnato un significato ad input ed output; il ruolo e l'importanza di queste frecce sarà ripreso e approfondito in seguito, in particolare nel contesto di un confronto con il cervello.

Centralità del livello simbolico Una delle proprietà più importanti di una macchina come quella rappresentata è che, in un'implementazione corretta e in condizioni di funzionamento normali, un utente può interagire con un bitrasduttore, e quindi con un computer, esclusivamente a livello simbolico.

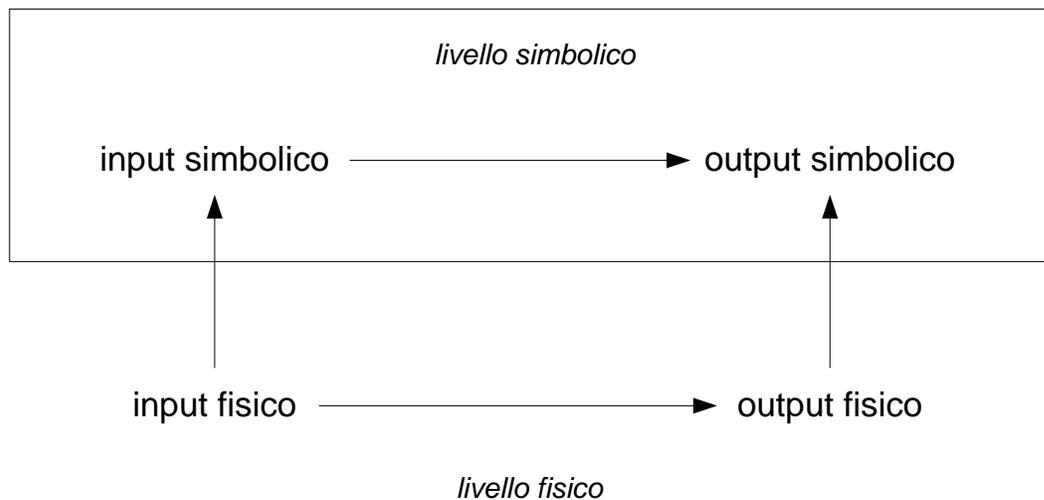


Figura 3.3 – Bitrasduttore

Questo significa che per poter ottenere dei risultati da un bitrasduttore è sufficiente fornire input simbolici e acquisire output simbolici; è possibile, cioè, utilizzare o programmare un computer senza prendere in considerazione il livello fisico, ma concentrandosi semplicemente sul livello simbolico.

Definizione di computer In conclusione, coerentemente con quanto esposto finora, è possibile riassumere tutte le osservazioni fatte nella seguente definizione di computer. Si definisce, dunque, computer un *bitrasduttore (macchina artefatto)* in cui *l'output simbolico è una funzione dell'input simbolico*.

Capitolo 4 – Cervello e computer

Definito il computer, si intende ora eseguire una simile analisi anche per il cervello al fine di fornirne una descrizione valida ed efficace; quest'ultima, insieme con la definizione di computer, sarà il fondamento per istituire un confronto significativo a livello semantico.

4.1 Definizione di cervello

**Genere e
differenze
specifiche del
cervello**

Nei capitoli precedenti sono state introdotte una serie di definizioni molto utili per affrontare il problema dell'identità di cervello e computer. Tutti i concetti enunciati sono tuttavia stati presentati quasi esclusivamente in relazione con il problema della definizione di un computer. Dopotutto uno degli obiettivi centrali era proprio quello di trovare una definizione di computer che fosse valida a prescindere dal contesto di questa tesi. Sono stati fatti perciò pochi riferimenti al possibile impiego di tali concetti per descrivere il cervello; è necessario quindi spiegare con maggiore precisione e dettaglio come questi termini possano essere utilizzati per predicare *genere e differenze specifiche* del cervello, così come fatto nel caso del computer.

Rifacendosi alla terminologia già utilizzata e definita nel secondo capitolo, si ritiene conveniente applicare il concetto di *macchina* anche al cervello. A prima vista l'uso di tale termine, comunemente utilizzato per indicare prodotti umani, potrebbe sembrare inadeguato e fuori luogo per definire il cervello. Tuttavia, considerando la definizione data, sembra che applicare questa categoria al concetto di cervello non sia affatto errato.

**Il cervello come
macchina**

Avendo definito nel secondo capitolo una *macchina* (par. 2.2) come *un insieme di entità che costituiscono un'unità, in cui ciascun componente interagisce con altri componenti e concorre a realizzare un processo causale*, la stessa definizione sembra applicabile anche nel caso del cervello; è possibile infatti riconoscere nel cervello più parti distinte e considerarlo dunque come un insieme di elementi, più o meno complessi, in relazione tra loro e che costituiscono un'unità; inoltre, tali parti non possono che essere causalmente connesse tra loro. La definizione di

cervello come macchina è dunque perfettamente sensata.

Tuttavia, come nel caso del computer, essendo il concetto di macchina troppo generale, si cerca di precisare con maggior dettaglio le specificità del cervello ricorrendo alle sottoclassi di macchina già definite nel corso del secondo capitolo.

***Il cervello come
macchina
naturale
biologica***

Posto che la definizione di cervello come *macchina artefatto* è del tutto inammissibile, dal momento che il cervello non può essere considerato alla stregua di una creazione umana implementata secondo un preciso disegno, è spontaneo fare ricorso alla classe delle *macchine naturali*. Quest'ultima è proprio la sottoclasse delle macchine che meglio si adatta a descrivere il concetto di cervello: il cervello risulta infatti essere una macchina non creata né modellata dall'uomo e dunque non può che essere una macchina naturale e, più precisamente, una *macchina naturale biologica*. In base a quanto espresso nella definizione (par. 2.2.2), il cervello è dunque una *macchina naturale biologica* e una *macchina naturale autopoietica*; sembra infatti coerente attribuire il concetto di autopoesi anche al cervello non in quanto lo si considera una forma di organismo, ma in quanto organo dotato, fino a prova contraria, delle proprietà di chiusura, auto-organizzazione e auto-contenimento generalmente richieste ad un sistema autopoietico. Assumendo questa definizione possiamo dunque riconoscere i limiti del cervello con la stessa linearità con cui sono stati evidenziati i confini di un computer; tali limiti corrispondono ai limiti dell'organo del cervello definiti dalla biologia.

***Il cervello come
macchina
sistema***

Come ogni macchina anche il cervello può, nonostante la sua complessità, essere descritto come un *sistema dinamico*. Si può cioè rappresentare il cervello definendo un modello matematico di sistema dotato di ingressi e uscite, in cui la relazione tra ingressi ed uscite risulta però non nota e, al limite, persino non deterministica. La possibilità di considerare il cervello come un sistema sembra offrire lo spunto per un interessante confronto a livello modellistico tra computer e cervello; tuttavia la conoscenza tuttora parziale del cervello rende questo confronto difficilmente perseguibile.

Determinato *genere e differenze specifiche* anche del cervello, è ora possibile condurre un'analisi sui suoi limiti analogamente a quanto fatto per il computer. Basandosi sugli studi biologici e sul concetto di macchina naturale autopoietica, è possibile identificare il cervello con l'insieme delle cellule nervose che lo costituiscono. Ancora una volta si noti che l'identificazione del cervello e il riconoscimento dei confini è

reso possibile dal concetto stesso adottato per descrivere il cervello: ovvero il concetto di macchina autopoietica che permette di definire unità e limiti di un sistema autopoietico.

**Input ed output
del cervello**

In base a questa delimitazione risultano facilmente identificabili input ed output, in entrambi i casi costituiti da stimoli elettrici. Tali impulsi si possono considerare dei *token*, al pari, ad esempio, delle differenze di tensione in ingresso o in uscita di un computer; essi costituiscono dunque degli input e degli output simbolici, in quanto oltre ad essere fisici e materiali, risultano avere un potere rappresentativo così come input ed output di un computer sono oggetto di un'*attribuzione di funzione di status semantica*.

**Il cervello è un
bitrasduttore**

Ne consegue che, essendo sia gli input che gli output simbolici, è possibile far rientrare anche il cervello all'interno della classe dei *bitrasduttori*. È possibile cioè descrivere il cervello come una macchina dotata due livelli: un livello *fisico*, costituito dall'insieme dei neuroni, oggetto di *attribuzione di funzione d'uso*, che manipolando fisicamente gli impulsi in ingresso generano un output anch'esso fisico; ed un livello *simbolico*, in cui ciascun token possiede all'interno del cervello un potere rappresentativo implicito per mezzo del quale si esplica la funzione simbolica che mette in relazione ingresso ed uscita. Analogamente a quanto detto per il computer, è dunque possibile relazionarsi con il cervello esclusivamente a livello simbolico, limitandosi a fornire e a ricevere input e output simbolici, trascurando così il livello materiale di realizzazione fisica.

**Definizione di
cervello**

È in conclusione possibile condensare queste osservazioni nella definizione di cervello. In maniera assolutamente simmetrica rispetto a quanto fatto con il computer, si definisce allora il cervello un *bitrasduttore (macchina naturale biologica) in cui l'output simbolico è una funzione dell'input simbolico*.

4.3 Rappresentazioni di computer e cervello

Le definizioni date a questo punto di *computer* e *cervello* permettono di istituire una consistente identità tra i due, sia a livello costitutivo che funzionale: entrambi appartengono al medesimo *genere* di oggetti (*macchine*) pur conservando *differenze specifiche* proprie (il primo è una *macchina artefatto*, il secondo una *macchina naturale autopoietica*) ed

inoltre condividono la medesima funzionalità data dall'appartenenza alla stessa classe di macchine (*bitrasduttori*). Si evidenzia che in questo modo sono state soddisfatte le condizioni di sensatezza poste all'inizio della tesi: computer e cervello non sono due oggetti totalmente dissimili tra i quali non ha senso stabilire un paragone, ma, al contrario, hanno molti tratti in comune come le loro definizioni evidenziano.

A questo punto, avendo rispettato sia le condizioni di sensatezza (con la definizione di computer e cervello) sia le condizioni di verità (con l'identificazione di computer e cervello con un bitrasduttore) è possibile procedere all'analisi di quello che è stato definito come *livello semantico* di computer e cervello; è qui, invero, che dopo aver sottolineato diverse similitudini si segnaleranno le prime e più significative differenze tra i due concetti.

**Problema delle
rappresentazioni**

Si vuole dunque analizzare e comprendere in quale modo ai simboli elaborati da cervello e computer sia attribuito un significato. O meglio, assunto che il significato è assegnato per mezzo di un'azione arbitraria di *attribuzione di funzione* da parte di un agente nel caso del computer o dal potere rappresentativo di input ed output nel caso del cervello, si vuole esaminare in quale modo sia possibile che ad un segno o ad un insieme di token sia attribuito un significato in relazione con il mondo reale. Questo conduce inevitabilmente ad un'analisi non solo del computer o del cervello ma anche della loro relazione con l'agente che attribuisce significato e alle *rappresentazioni* che egli possiede.

**Rappresentazioni
del computer**

Nel caso di un computer è stato già accennato come attraverso una *attribuzione di funzione di status semantica* sia possibile far corrispondere input ed output fisici alle relative controparti simboliche. Questa attribuzione è eseguita da un osservatore o da una collettività, che interagendo con il computer, assegna un contenuto ai simboli elaborati. Attraverso questa operazione, i token ricevuti in ingresso e restituiti in uscita dal computer possono fare riferimento ad una realtà esterna sia al computer sia all'agente che interagisce con la macchina. È evidente che i simboli ricevuti e restituiti dal computer non hanno intrinsecamente un significato né lo assumono all'interno del computer stesso; piuttosto il loro unico significato è quello che ottengono in relazione alla collettività o all'osservatore che opera con il computer. Riferendosi al contenuto dei simboli elaborati da un computer, si deve dunque parlare di *rappresentazioni derivate*: input ed output ottengono infatti un contenuto per mezzo di agenti che, attraverso le proprie *rappresentazioni intrinseche*, attribuiscono ai simboli in ingresso ed uscita condizioni di soddisfazione.

Riprendendo e completando l'immagine presentata nei capitoli precedenti per raffigurare un bitrasduttore e il suo funzionamento a livello fisico e simbolico, è possibile aggiungere ulteriori dettagli per evidenziare il contributo degli osservatori nell'attribuire significato agli input e agli output simbolici del computer, come in figura 4.1. In questa immagine, l'icona dell'agente vuole evidenziare il ruolo fondamentale ed imprescindibile di un osservatore che opera con un computer: solo attraverso la sua interazione con il computer (simboleggiata dalla freccia punteggiata tra l'agente e l'input o l'output del computer), input ed output simbolici risultano essere delle rappresentazioni derivate in grado di riferirsi al mondo esterno (simboleggiata dalla freccia punteggiata tra il livello simbolico del computer e la realtà).

Rappresentazioni del cervello

Il funzionamento del cervello risulta profondamente differente. Anche nel caso del cervello sono presenti, come in ogni bitrasduttore, input fisici e output fisici; per mezzo del potere rappresentativo implicito che possiedono, i token ricevuti in ingresso e restituiti in uscita sono messi in corrispondenza con input ed output simbolici; in questo modo ai token è assegnato un significato preciso. Fino a questo punto, le operazioni del computer e del cervello sembrano essere simili; tuttavia è completamente diverso il modo in cui i simboli elaborati dal computer e i simboli elaborati dal cervello si relazionano con il mondo.

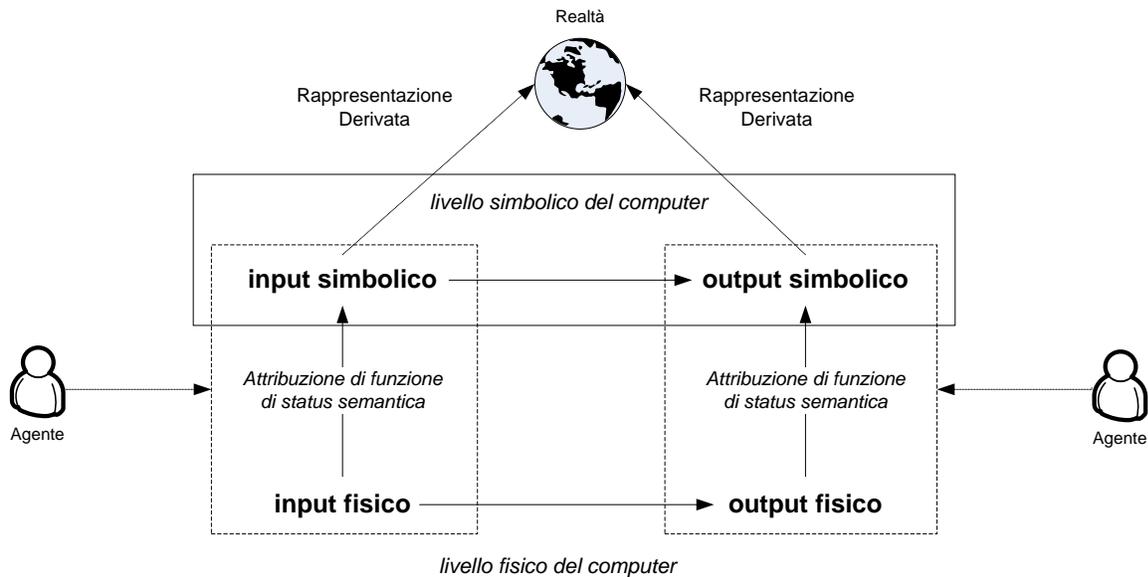


Figura 4.1 – Computer

Infatti, mentre il computer necessita di un agente e si riferisce alla realtà esterna per mezzo di *rappresentazioni derivate*, il cervello funziona senza bisogno dell'intervento di alcun osservatore esterno. I simboli elaborati dal cervello assumono un significato all'interno del cervello stesso senza che uno o più agenti debbano attribuirne o riconoscerne il significato. Si tratta dunque non più di *rappresentazioni derivate*, bensì di *rappresentazioni intrinseche*; esse assumono cioè il loro contenuto nella mente stessa del soggetto.

Ancora una volta, riprendendo l'immagine utilizzata per raffigurare un bitrasduttore, è possibile sintetizzare graficamente il funzionamento del cervello come in figura 4.2. L'accostamento delle due immagini, quella del computer (figura 4.1) e quella del cervello (figura 4.2), consentono di cogliere a colpo d'occhio le differenze che intercorrono tra il funzionamento del computer e del cervello.

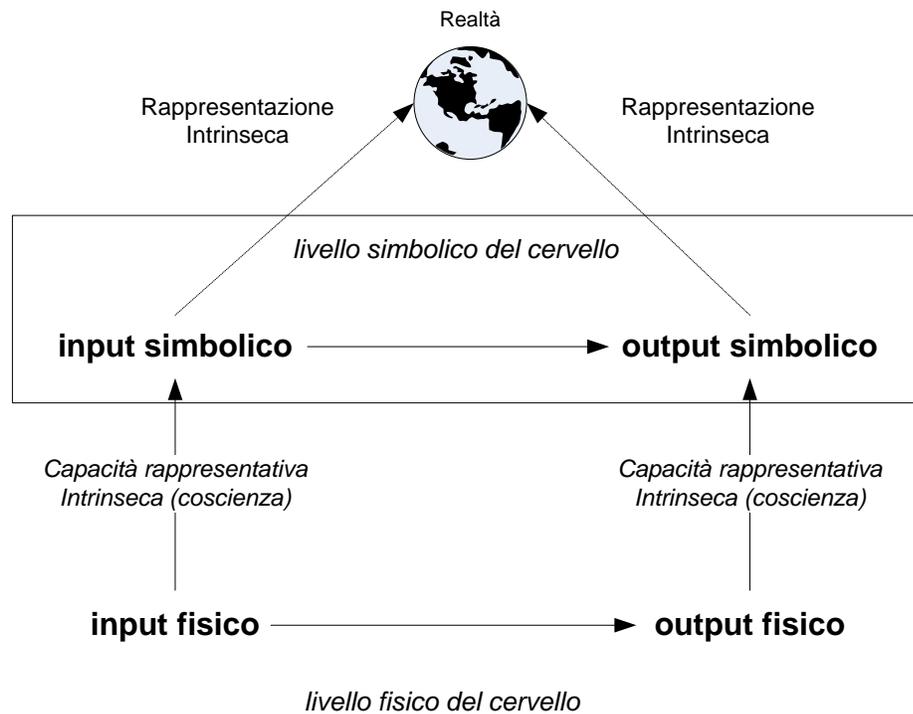


Figura 4.2 – Cervello

Nel caso di quest'ultimo, è evidente l'assenza di un agente che si relaziona a input ed output simbolici e attraverso il quale i simboli si riferiscono alla realtà; piuttosto, all'interno del cervello stesso e per mezzo di rappresentazioni intrinseche, input ed output simbolici sono in grado di fare riferimento ad una realtà esterna al cervello (freccia punteggiata tra il livello simbolico del cervello e la realtà). Riassumendo, sebbene le rappresentazioni di computer e cervello possano avere lo stesso *contenuto* (ad esempio, l'affermazione dell'esistenza di una caffettiera), lo stesso *modo intenzionale* (ad esempio, credere), la stessa *direzione di adattamento* (ad esempio, adattamento mente-a-mondo) e le stesse *condizioni di soddisfazione* (ad esempio, l'esistenza della caffettiera), ciò che le differenzia è che le prime sono sempre *osservatore-dipendenti*, le seconde, invece, pur esistendo solo in relazione ad un soggetto, sono sempre *osservatore-indipendenti* (ad esempio, nel computer, il credere nella presenza di una caffettiera dipende dalla presenza di un osservatore che assegna ai simboli elaborati questo preciso significato; nel cervello umano non è invece necessario alcun agente, poiché per mezzo del potere rappresentativo intrinseco del cervello, questa asserzione assume significato direttamente nella mente del soggetto).

Dopo aver dimostrato le similitudini di fondo tra cervello e computer a livello fisico e sintattico (ovvero come siano elaborati i simboli) e dopo aver evidenziato una profonda differenza concettuale tra cervello e computer a livello semantico (ovvero come i simboli assumano un significato), per completare le osservazioni sull'identità tra cervello e computer verranno presentate nel corso del prossimo capitolo ulteriori differenze e similitudini sia a livello teorico che pratico.

Capitolo 5 – Il cervello è un computer?

In questo capitolo si vogliono offrire al lettore ulteriori riflessioni per poter valutare sensatamente la domanda se il cervello sia o meno un computer. In particolare verranno sia esposti nuovi argomenti riguardo ad analogie e differenze tra computer e cervello sia riprese e approfondite le osservazioni delle pagine precedenti.

5.1 Il cervello è un computer

Tralasciando per il momento le modalità con cui è assegnato significato ai simboli elaborati da computer e cervello e al tipo di *rappresentazione* su cui si basano, le similitudini tra il funzionamento del cervello e del computer sono profonde. In entrambi i casi ci si trova infatti di fronte a delle implementazioni reali di bitrasduttori; per operare con queste macchine è, come già detto, sufficiente relazionarsi ed operare esclusivamente a livello simbolico.

**Funzionamento
del computer in
condizioni
ordinarie**

Nel caso del computer è possibile interagire fornendo esclusivamente input simbolici e ricevendo parimenti output simbolici, ignorando cioè del tutto la presenza dello strato fisico. Questo significa che per un utente l'unico livello significativo è quello simbolico; al livello fisico è richiesto solo di assicurare la validità e la consistenza del livello superiore; il livello fisico deve, cioè, essere una semplice implementazione che assicuri un mapping convenzionale e sensato tra input ed output, in modo che un osservatore possa riconoscere l'esistenza di un funzione simbolica tra input ed output. In generale tale funzionamento è garantito finché il livello fisico è in grado di operare in condizioni regolari e ordinarie; la presenza di malfunzionamenti o guasti a livello fisico può causare comportamenti anomali sia a questo livello, che, conseguentemente a livello simbolico; in tali condizioni i risultati forniti da un computer risulterebbero non corretti o del tutto privi di senso. In sintesi, è possibile precisare la precedente affermazione riguardo l'interazione con un computer come segue: il computer è un bitrasduttore con il quale, *in condizioni di funzionamento ordinario*, è possibile interagire a livello esclusivamente simbolico; è necessario invece prendere in considerazione e studiare il livello fisico solo in presenza di incoerenze; ad esempio, se il risultato di una semplice somma fosse non corretto,

sarebbe necessario ricercarne le cause ad un livello inferiore; nella pratica, la ricerca di errori viene generalmente condotta analizzando anzitutto una serie di livelli inferiori, tutti simbolici (livello del *codice di alto livello*, livello del *codice macchina*, livello del *microcodice*) e solo alla fine, qualora queste verifiche non abbiano avuto successo, si dovrebbe valutare la correttezza del livello fisico vero e proprio.

**Funzionamento
del cervello in
condizioni
ordinarie**

Il cervello presenta un funzionamento molto simile. Anche nel caso del cervello è infatti possibile interagire considerando solo il livello simbolico; questo è, dopotutto, il comportamento di ogni uomo che si relaziona ai suoi simili per mezzo di input e output simbolici, con parole o gesti, quasi sempre non sapendo nulla del reale funzionamento del cervello. Ma l'analogia non si ferma a questo punto. Esattamente come nel caso del computer, il livello fisico è ignorato solo finché gli output simbolici forniti dal cervello risultano coerenti. Nel caso di un computer, tuttavia, è molto semplice giudicare se ad un input simbolico corrisponda il corretto e sensato output simbolico: si tratta cioè semplicemente di valutare se ad un dato ingresso corrisponde l'uscita che era stata progettata. È evidente, invece, che stabilire quando input ed output elaborati dal cervello siano tra loro coerenti è un compito estremamente arduo; è assunzione comune ritenere che input e output siano tra loro coerenti se risultano legati da un *criterio di razionalità*; ora, definire la *razionalità* indipendentemente dal cervello è un'operazione piuttosto complessa: ne è dimostrazione il fatto che spesso si definisce la *razionalità* come il regolare funzionamento del cervello; è chiaro che nel presente caso questa definizione risulterebbe circolare: si affermerebbe infatti che il cervello funziona correttamente quando funziona razionalmente, ovvero secondo il regolare funzionamento del cervello. Volendo evitare di incorrere in questo circolo vizioso, si considera *razionale* un comportamento per il quale è possibile identificare delle motivazioni e degli input che giustifichino un dato output; questo significa che, assegnati degli input, se è possibile identificare delle connessioni o delle relazioni (principalmente di tipo causale) a livello semantico attraverso cui spiegare un comportamento, quest'ultimo può generalmente considerarsi razionale; ad esempio, in una giornata piovosa (input) la decisione di utilizzare un ombrello (output) mentre si cammina in un luogo scoperto si considera un comportamento razionale poiché è possibile identificare un motivo per l'apertura dell'ombrello: la volontà di non bagnarsi che funge da ragione dell'azione intrapresa.

**Problema della
definizione di
razionalità**

Si noti che la ragione addotta per spiegare un comportamento non deve necessariamente essere condivisa: essa deve giustificare un output, ma non per questo deve essere considerata una ragione accettata da ogni persona; in altre parole, di fronte a input identici, persone diverse

possono assumere comportamenti differenti in base a ragioni differenti, ma ugualmente giustificabili e sensate.

Si sottolinea anche che le ragioni presentate per spiegare un comportamento devono essere sempre a livello simbolico; non sarebbe cioè corretto tentare di spiegare un comportamento facendo riferimento al livello fisico; in quest'ultimo caso non solo si perderebbe la simmetria con il funzionamento del computer, ma soprattutto si sposterebbe l'ambito dell'analisi dal livello simbolico al livello fisico, in cui, tuttavia, non risulta corretto parlare di razionalità, giacché la razionalità lega input ed output simbolici.

Si tenga presente, infine, che questa trattazione di razionalità è molto semplificata; l'esempio riportato (apertura dell'ombrello a causa della pioggia) è utile per raffigurare un comportamento manifestamente razionale ma la sua semplicità potrebbe essere ingannevole: in realtà giustificare un comportamento umano, ovvero trovare delle ragioni che lo spieghino, è un'operazione estremamente intricata che richiede di considerare con attenzione e interpretare una molteplicità di input simbolici presenti e passati.

Ora, come detto, finché input ed output risultano legati tra loro coerentemente secondo il criterio di razionalità appena esposto, è possibile relazionarsi con il cervello umano esclusivamente a livello simbolico; nel momento in cui tale criterio di razionalità non sia più soddisfatto è invece necessario, come nel caso del computer, considerare anche il livello fisico; incoerenze a livello simbolico possono infatti essere spesso spiegate da malfunzionamenti o problemi a livello fisico. Comportamenti irrazionali, anche non patologici, potrebbero essere causati da eccessivo stress o stanchezza, oppure da scompensi chimici o ormonali (ad esempio, malumori e reazioni scontrose possono frequentemente essere dovute ad un malessere fisico). Dunque, anche il cervello può essere più precisamente definito come segue: il cervello è un bitrasduttore con il quale, *in condizioni di funzionamento ordinario*, è possibile interagire esclusivamente a livello simbolico; è necessario invece prendere in considerazione e studiare il livello fisico solo in presenza di incoerenze.

Analogie tra le operazioni di computer e cervello

Oltre a questa similitudine tra computer e cervello in merito al loro funzionamento in condizioni ordinarie e non, è inevitabile aggiungere qualche osservazione anche sull'apparente analogia tra le operazioni eseguite dal cervello e quelle effettuate da un computer. Idealmente, infatti, sia i procedimenti del cervello umano che di un computer dovrebbero essere caratterizzati da rigore e consequenzialità; in entrambi

i casi dovrebbe, cioè, esistere un criterio di razionalità o logicità che lega le azioni e per mezzo del quale è possibile spiegare tutte le operazioni intraprese. Come già proposto da Thomas Hobbes (1655) in *Computatio sive logica*, nella prima parte dell'opera *De corpore*, il *ragionare* umano e il *calcolare* artificiale sarebbero due operazioni equivalenti; tale convinzione si rispecchia nell'etimologia stessa dei termini: *ragione* deriva infatti dal latino *ratio*, vocabolo usato per indicare un conto o un calcolo, ed inoltre termini come ragionamento o sillogismo indicavano originariamente proprio delle forme di calcolo. Pur assumendo come valido questo accostamento tra il *calcolare* artificiale e il *ragionare* umano, esso risulta utile per indicare dei punti di contatto tra il cervello e il computer ma è tuttavia parziale e insufficiente per essere il fondamento di un paragone compiuto: il *ragionare* non è infatti l'unica operazione compiuta dal cervello e ridurre tutta l'attività di quest'ultimo a calcolo potrebbe essere riduttivo; in altre parole, sarebbe più corretto, parlare di un'analogia tra le operazioni del computer ed un sottoinsieme delle possibili operazioni del cervello.

5.2 Il cervello non è un computer

Rappresentazioni e problema della coscienza

Una prima significativa differenza tra cervello e computer è quella esposta nel capitolo precedente a proposito della relazione con la realtà. Si è cioè affermato che mentre il computer è in grado di riferire la realtà per mezzo di *rappresentazioni derivate*, il cervello ricorre invece a *rappresentazioni intrinseche*. Tale differenza introduce direttamente al problema della *coscienza*, di cui si fa solo un breve accenno. In base alle definizioni date nel quarto capitolo, una *rappresentazione intrinseca* è una rappresentazione che assume significato nella mente dell'uomo per mezzo della coscienza. Una delle caratteristiche strutturali della coscienza, secondo la forma di pensiero realista che è stata presa come riferimento (Searle, 1998), è proprio quella della *intenzionalità o rappresentatività*, ovvero la possibilità di essere diretta o riferirsi alla realtà esterna. A causa della sua soggettività ontologica, la coscienza esiste solo in relazione ad un soggetto e solo all'interno di un soggetto è in grado di attribuire significato. È la presenza della coscienza, dunque, che permette alle rappresentazioni del cervello di riferirsi direttamente al mondo esterno, così come è l'assenza della coscienza a richiedere la presenza di un agente che relazioni input ed output simbolici prodotti da un computer con la realtà esterna.

**Differenza tra
cause e ragioni**

Si riconsideri poi il paragone tra il funzionamento (e il malfunzionamento) di cervello e computer offerto nel paragrafo precedente. Si è affermato che valutare la coerenza tra input simbolici e output simbolici in un computer risulta molto semplice (si tratta cioè semplicemente di valutare se ad un dato ingresso corrisponda l'uscita che era stata progettata), mentre nel caso del cervello è estremamente complesso (è necessario valutarne la razionalità attraverso la ricerca e l'analisi della fondatezza di un insieme di ragioni). Questa disuguaglianza è dovuta al fatto che l'output di un computer è determinato da *cause*, mentre quello di un cervello è conseguenza di *ragioni*. Ciò significa che mentre nel computer ogni causa è sufficiente per spiegare un output, altrettanto non vale nel cervello: la presenza di una o più *ragioni* in grado di giustificare o di spiegare un'azione è necessaria, almeno in condizioni ordinarie, ma non sufficiente per determinare un output. Nel caso dell'agire umano si parla infatti di *acrasia* o *weakness of will* (*debolezza di volontà*) per indicare quel fenomeno per cui l'uomo è in grado di agire anche contro le sue migliori ragioni e contro il giudizio più fondato (Buss, 2002); nonostante tutte le motivazioni possibili per agire diversamente, è quindi possibile che un uomo decida di agire contro la sua ragione e compire azioni che risultano ingiustificate o addirittura irrazionali (ad esempio, la decisione di fumare di una persona pur consapevole delle conseguenze è un caso di debolezza di volontà).

**Teoria
dell'acrasia**

In definitiva, dunque, mentre le azioni di un computer hanno sempre delle *cause* precise e sufficienti, l'agire umano ha bisogno di motivazioni minime e l'insieme di tutte le ragioni possibili e fondate per giustificare un comportamento è comunque insufficiente per asserire l'inevitabilità del comportamento stesso.

**Problema del
libero arbitrio**

Queste riflessioni sembrano inoltre condurre direttamente al problema del *libero arbitrio* e del *determinismo*, su cui si fondano diverse argomentazioni circa una possibile differenza sostanziale tra cervello e computer. Escludendo da queste osservazioni computer in fase di studio, come i computer quantistici, sul cui funzionamento sarebbe necessario discutere, le operazioni di tutti i restanti moderni computer sono chiaramente deterministiche: ogni risultato, persino qualora si chieda al computer di generare numeri casuali, o più precisamente, appunto, *pseudo-casuali* (Knuth, 1981), è necessariamente e univocamente determinato dagli input ricevuti e dallo stato della macchina. Un'analisi simile difficilmente può essere condotta nei riguardi del cervello: in generale quando si fornisce un input al cervello umano è impossibile stabilire quale sarà il suo output; al più è possibile stimare lo spazio dei possibili output, ovvero di tutte quelle uscite che è razionale attendersi

dato un preciso input. In verità è impossibile affermare se questo raggio di possibili uscite sia dovuto ad una forma di libertà o sia una semplice illusione; è possibile infatti che il comportamento umano appaia non deterministico esclusivamente a causa dell'ignoranza dell'osservatore: è probabile infatti che quest'ultimo non conosca tutti i reali input del cervello o giudichi a livello simbolico casuali alcuni processi che a livello fisico sono invece strettamente causali; è cioè ipotizzabile che con una completa consapevolezza di tutti gli input e una profonda conoscenza del funzionamento del cervello sia possibile relazionare deterministicamente input ed output; si potrebbe addirittura immaginare che un simile sapere ci potrebbe permettere di risolvere quell'apparente incongruenza nel comportamento umano che è l'*acrasia*: si potrebbero cioè spiegare le azioni dettate da una debole volontà a livello fisico o per mezzo di input ora ignoti. In conclusione non è possibile asserire nulla di indiscutibile, ora come ora, riguardo al determinismo del cervello.

Esiste anche un'ulteriore differenza tra il funzionamento del computer e del cervello sempre a livello simbolico.

**Totalità del
legame di
razionalità**

All'interno di un computer la freccia che lega input ed output simbolico è una freccia completamente determinata che esprime una relazione di causalità tra input ed output; la sua coerenza e validità può essere facilmente stimata valutando se l'output simbolico ottenuto è uguale al risultato atteso in corrispondenza dell'input simbolico fornito. Oltre a godere di queste proprietà, il legame tra input ed output nel computer è un legame totale: l'output del computer è determinato solo e soltanto dagli input ricevuti; esclusi malfunzionamenti, non esiste, cioè, la possibilità che l'output simbolico di un computer possa variare dipendentemente da altri fattori che non siano gli input.

All'interno del cervello la freccia che lega input ed output non solo è, come già detto, non completamente determinata, ma gode della proprietà di totalità solo in via ipotetica. Non è infatti certo che il legame tra input simbolico ed output simbolico del cervello possa essere completamente spiegato a livello razionale; o meglio, anche se a posteriori è possibile quasi sempre, in assenza di malfunzionamenti, addurre una spiegazione razionale per un dato comportamento, questo non esclude che, nel momento in cui si è assunto il detto comportamento, la scelta sia stata influenzata o determinata dallo stato del livello fisico del cervello. Potrebbe quindi essere possibile che un output simbolico del cervello umano sia in parte determinato dal corrispettivo input simbolico, in parte causato da processi a livello fisico che condizionerebbero l'output (ad esempio, di fronte a due alternative razionalmente giudicate equivalenti da un soggetto, la decisione di propendere per l'una piuttosto che per

l'altra potrebbe essere dovuta a fenomeni di livello fisico che si realizzano nel cervello; reazioni chimiche, quali la produzione di serotonina, potrebbero spingere il soggetto ad intraprendere una scelta piuttosto che un'altra). In questo modo, tali contributi del livello fisico non potrebbero considerarsi dei malfunzionamenti (in quanto il legame razionale tra input ed output simbolico sarebbe preservato), ma sminuirebbero il valore dell'interpretazione razionale del comportamento umano (che perderebbe il suo valore di interpretazione totale e in grado di spiegare l'agire umano in ogni suo aspetto) e soprattutto costituirebbero un notevole elemento di differenza tra il funzionamento del computer e del cervello. Per evidenziare questo funzionamento profondamente differente, si riporta in figura 5.1 questa visione alternativa del funzionamento del cervello.

Esistenza di un livello intermedio Infine, in riferimento al caso di malfunzionamento esposto in precedenza, diversamente da un computer, potrebbe essere possibile spiegare spesso il comportamento irrazionale del cervello non scendendo direttamente dal livello simbolico al livello fisico, ma considerando prima un ipotetico livello intermedio, identificabile con l'inconscio della psicanalisi e di altre teorie della psicologia clinica.

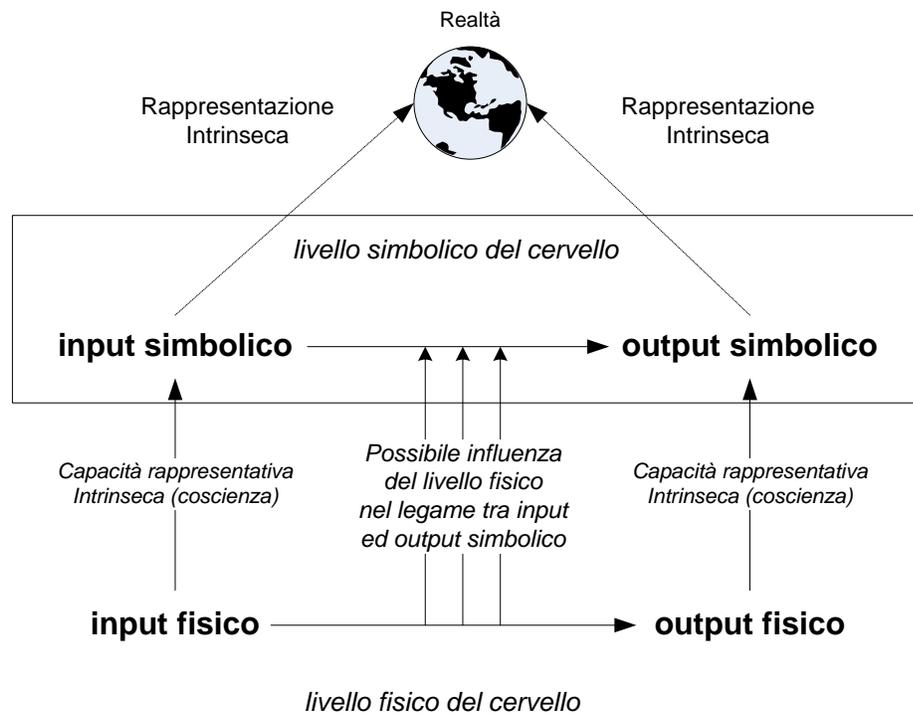


Figura 5.1 – Concezione alternativa del cervello

Sarebbe cioè questo il livello dei contenuti definiti appunto inconsci, ovvero un insieme di conoscenze ed esperienze non consapevoli, simboliche ma non presenti al livello superiore, per mezzo delle quali sarebbe possibile spiegare razionalmente comportamenti apparentemente irrazionali. Questo livello permetterebbe cioè di spiegare quei comportamenti a prima vista non sensati, che neanche il soggetto è in grado di giustificare, ma che non dipendono da malfunzionamenti a livello fisico. È evidente tuttavia che i confini tra questo presunto livello intermedio, il livello fisico e il livello simbolico risultano piuttosto nebulosi: sarebbe infatti necessario precisare sia quando un contenuto del cervello risulta conscio o inconscio, sia descrivere quale rapporto intercorre tra i contenuti inconsci e il livello fisico, ovvero se e quando i primi possano essere ridotti al semplice funzionamento fisico.

5.3 Conclusioni

Si è dunque approfondito il paragone a livello simbolico tra cervello e computer e si sono presentati altri argomenti citati per difendere o negare l'identità di cervello e computer; l'obiettivo dichiarato era quello di fornire una serie di ragioni per cui sostenere in conclusione che “il cervello è un computer” o “il cervello non è un computer”. Si sono forniti argomenti validi e consistenti per sostenere entrambe le posizioni; tuttavia, le condizioni poste all'inizio di questa tesi, quelle di verità in particolare, richiedevano che tutte le condizioni fossero soddisfatte affinché la risposta alla domanda il “cervello è un computer?” potesse essere positiva; anche una sola delle diverse differenze enumerate nel paragrafo precedente sarebbe stata sufficiente per giungere ad una conclusione negativa. Inoltre, pur sospendendo il giudizio su argomentazioni come quella dell'*acrasia* o del *libero arbitrio* e considerandole dunque non sufficientemente provate per sostenere la differenza tra cervello e computer, permangono comunque quelle profonde differenze a livello di rappresentazioni tra cervello e computer che sono state il centro di questa trattazione. L'unico vero argomento approfondito in questa tesi da poter essere addotto come ragione dell'asserzione “il computer non è un cervello” è dunque quello sviluppato riguardo la differenza tra le *rappresentazioni derivate* di un computer e le *rappresentazioni intrinseche* di un cervello.

In conclusione la risposta tratta da queste pagine di riflessioni non vuole essere la lapidaria affermazione che “il cervello è un computer” o “il cervello non è un computer”, affermazioni che di certo richiederebbero maggiore studio e approfondimento, ma la semplice osservazione che *non sarà possibile parlare di un computer che sia un cervello* finché non si sia risolto il problema della differente semantica e delle *rappresentazioni* usate da cervello e computer.

Bibliografia

- Blikstein, P. (2007). Programmable Water, http://www.blikstein.com/paulo/projects/project_water.html (ultimo accesso: 13 maggio 2007).
- Boden, M. A. (2000). Autopoiesis and Life, *Cognitive Science Quarterly*, 1, 117-145.
- Brooks, R. A. (1991). Intelligence without representation, *Artificial Intelligence*, 47, 139-159.
- Buss, S. (2002). Personal Autonomy, *The Stanford Encyclopedia of Philosophy (Winter 2002 Edition)*, <http://plato.stanford.edu/archives/win2002/entries/personal-autonomy/> (ultimo accesso: 13 maggio 2007).
- Colombetti, M. (2003). Le rappresentazioni.
- Copeland, B. J. (1996). What is computation?, *Synthese*, 108, 335-359.
- Copeland, B. J. (2002). The Church-Turing Thesis, *The Stanford Encyclopedia of Philosophy (Fall 2002 Edition)*, <http://plato.stanford.edu/archives/fall2002/entries/church-turing/> (ultimo accesso: 13 maggio 2007).
- Fritz, W. B. (1963). Selected Definitions, *Communications of the ACM*, 4 (6), 152-158.
- Frixione, M., Palladino, D. (2004). *Funzioni, macchine, algoritmi*, Carocci, Roma.
- Geraci, A. K., Wilson, P. C., ed. (1995). *Standard Glossary of Computer Hardware Terminology*, IEEE, New York.
- Hamacher, V. C., Vranesic, Z. G., Zaky, S. G. (2002) *Computer Organization, fifth edition*, McGraw-Hill. Trad. it. 2005, *Introduzione all'architettura dei calcolatori*, McGraw-Hill, Milano.
- Harnad, S. (1990). The Symbol Grounding Problem, *Physica D*, 42, 335-346.
- Hobbes, T. (1655). *De Corpore*, Londra.
- Knuth, D. E. (1981). *The Art of Computer Programming – Vol. 2: Seminumerical Algorithms*, Addison-Wesley.
- Marian, D. (2005). The Correspondence Theory of Truth, *The Stanford Encyclopedia of Philosophy (Fall 2005 Edition)*, <http://plato.stanford.edu/archives/fall2005/entries/truth-correspondence/> (ultimo accesso: 13 maggio 2007) .

- Maturana, H. R., Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*, D. Reidel Publishing Co., Boston.
- Newell, A., Simon, H. A. (1976). Computer Science as Empirical Inquiry: Symbols and Search, *Communications of the ACM*, 19, 113-126.
- Nilsson, N. J. (2006). The Physical Symbol System Hypothesis: Status and Prospects, <http://ai.stanford.edu/~nilsson/OnlinePubs-Nils/PublishedPapers/pssh.pdf> (ultimo accesso: 13 maggio 2007).
- Pollack, J. B. (2006). Mindless Intelligence, *IEEE Intelligent Systems*, 3 (21), 50-56.
- Rosenberg, A., Bouchard, F. (2002). Fitness, *The Stanford Encyclopedia of Philosophy (Winter 2002 Edition)*, <http://plato.stanford.edu/archives/win2002/entries/fitness/> (ultimo accesso: 13 maggio 2007).
- Russel, S, Norvig, P. (2003). *Artificial Intelligence: A Modern Approach, Second Edition*, Prentice Hall. Trad. it. 2005, *Intelligenza Artificiale Vol. 1 Un Approccio Moderno*, Pearson Italia, Roma.
- Searle, J. R. (1990). Is the Brain a Digital Computer, *Proceedings and Addresses of the American Philosophical Association*, 64, 21-37.
- Searle, J. R. (1992). *The Rediscovery of the Mind*, MIT Press, Cambridge, Massachusetts.
- Searle, J. R. (1998). Minds, brains, and programs, *Behavioural and Brain Sciences*, 3, 417-457.
- Searle, J. R. (1998). *Mind, Language and Society*, Basic Books, New York. Trad. it. 2000, *Mente, linguaggio, società*, Raffaello Cortina, Milano.
- Sterrett, S. G. (2000). Turing's Two Tests for Intelligence, *Minds and Machines*, 10, 541-549.
- Turing, A. M. (1936). On computable numbers with an application to Entscheidungsproblem, *Proceedings of the London Mathematical Society, Series 2*, 42, 230-265.
- Turing, A. M. (1950). Computing machinery and intelligence, *Mind*, 59, 433-460.