

Gli zombie non possono esserci

Marco Giunti

University of Cagliari

email: giunti@unica.it

homepage: <http://edu.supereva.it/giuntihome.dadacasa/>

-
- 1 **IL PROBLEMA DELLA COSCIENZA FENOMENICA NELL'ULTIMO QUARTO DEL NOVECENTO**
 - 2 **LA DISTINZIONE FRA COSCIENZA FENOMENICA E COSCIENZA PSICOLOGICA**
 - 3 **I PROBLEMI DELLA COSCIENZA SECONDO CHALMERS**
 - 3.1 *IL PROBLEMA FACILE DELLA COSCIENZA*
 - 3.2 *IL PROBLEMA DIFFICILE DELLA COSCIENZA*
 - 3.3 *IL PROBLEMA SUPERDIFFICILE DELLA COSCIENZA*
 - 4 **L'ARGOMENTO DEGLI ZOMBIE FENOMENICI E LA SFIDA DI CHALMERS**
 - 5 **CHE COS'È UN'ESPERIENZA IN GENERALE?**
 - 6 **DUE DIFFICOLTÀ NELLA DEFINIZIONE DI ESPERIENZA**
 - 6.1 *ESPERIENZE ASSENTI*
 - 6.2 *ESPERIENZE STRANE*
 - 7 **CHE COS'È UN'ESPERIENZA COSCIENTE?**
 - 8 **PROPOSTA DI ANALISI DEL CASO PARADIGMATICO DI ESPERIENZA COSCIENTE: L'ESPERIENZA DEI QUALIA**
 - 9 **TEORIA DELLA COSCIENZA FENOMENICA E FISCALISMO**
 - 9.1 *LA SPIEGABILITÀ RIDUTTIVA FISICA DELLE ESPERIENZE COSCIENTI, L'IMPOSSIBILITÀ LOGICA DEI MONDI ZOMBIE E DI QUELLI ANGELICI*
 - 9.2 *L'IPOTESI FISCALISTA FORTE E IL TEOREMA DI RIDUCIBILITÀ FISICA*
 - 10 **CONCLUSIONE: IL PROBLEMA DIFFICILE DELLA COSCIENZA NON È UN PROBLEMA SCIENTIFICO, MA FILOSOFICO**
- APPENDICE: CONSISTENZA DELLA TEORIA ASSIOMATICA**
- RIFERIMENTI BIBLIOGRAFICI**
- NOTE**
-

Parole chiave

spiegazione riduttiva; sopravvenienza; problema difficile della coscienza; gap esplicativo; zombie; possibilità logica; concepibilità; ontologia; fiscalismo; materialismo; dualismo; qualia; concetto di coscienza

Sommario

Partendo dalla distinzione fra coscienza fenomenica e coscienza psicologica, propongo un quadro formale entro cui si possa analizzare l'argomento degli zombie fenomenici di Chalmers. Utilizzando tale quadro formale, chiarisco il problema che l'argomento di Chalmers pone a chiunque si proponga di (i) costruire una teoria della coscienza fenomenica e (ii) sostenere la spiegabilità riduttiva delle esperienze coscienti nei termini dei fatti fisici. Passo quindi ad ampliare tale quadro formale, fornendo una vera e propria

teoria assiomatica della coscienza. La spiegazione della coscienza fenomenica provvista da tale teoria non è affatto incompatibile con una prospettiva fisicalista ma, al contrario, una volta inserita in un quadro fisicalista minimale, permette di provare la spiegabilità riduttiva fisica delle esperienze coscienti e, conseguentemente, l'impossibilità logica dei mondi zombie. Concludo quindi che l'intuizione di Chalmers secondo cui la coscienza fenomenica eccederebbe la portata di qualunque teoria fisica è priva di fondamento. Infine, metto in luce una possibile origine di tale errata percezione, e cioè il non aver riconosciuto che il problema difficile della coscienza non è un problema di spiegazione scientifica ma, piuttosto, un problema filosofico di esplicazione concettuale.

English title

Zombies cannot be there

Keywords

reductive explanation; supervenience; hard problem; explanatory gap; zombie; logical possibility; conceivability; ontology; physicalism; materialism; dualism; qualia; concept of consciousness

Short abstract

On the basis of the distinction between phenomenal and psychological consciousness, I propose a formal framework where we can express and analyze a strong form of Chalmers' zombie argument. By employing such formal framework, I make clear the kind of problem that this argument poses to anyone who is willing to (i) construct a theory of phenomenal consciousness and (ii) maintain the reductive explainability of phenomenal consciousness by physics. I then extend such formal framework so as to provide a theory of consciousness in axiomatized form. The explanation of phenomenal consciousness provided by this theory is by no means inconsistent with a physicalist perspective. In fact, once the theory is supplemented with a minimal physicalist assumption, we can prove the reductive explainability of phenomenal consciousness by physics and, as a consequence, the logical impossibility of zombie worlds. I then conclude that Chalmers' intuition, according to which phenomenal consciousness is beyond the scope of any physical theory, is not tenable. Finally, I hint at a possible source of such erroneous perception, that is, not realizing that the hard problem of consciousness is not a problem of scientific explanation but, rather, a philosophical problem of conceptual explication.

Long abstract

I propose a simple formal possible world ontology where we can naturally express and analyze a strong form of Chalmers' zombie argument. By employing such formal framework, I make clear the kind of problem that this argument poses to anyone who is willing to (i) construct a theory of phenomenal consciousness and (ii) maintain the reductive explainability of phenomenal consciousness by physics.

I then extend such formal ontology so as to provide a theory of consciousness in axiomatized form. The key idea of my theory is to show how, within an appropriate extension of the proposed possible world ontology, we can make sense of the distinction between first person or internal facts of consciousness (*i.e. phenomenal consciousness*) and third person or external ones (*i.e. psychological consciousness*), in a natural manner. The formal machinery which allows such an extension is a straightforward adaptation of the standard axioms of the lambda calculus. According to my theory, being internal

(phenomenal) or external (psychological or functional) is not an intrinsic feature of a fact of consciousness but, rather, depends on whether or not a special relationship between the fact and a reference system holds. Thus, the very same fact of consciousness turns out to be internal (or phenomenal) with respect to a reference system, and external with respect to other systems. It is a theorem that, for any fact of consciousness, there is exactly one system for which the fact is internal. I call such a privileged system *the subject* of the fact of consciousness; the fact is external for any other system.

My theory of phenomenal consciousness is by no means inconsistent with the claim that phenomenal consciousness is reductively explainable by physics. In fact, once the theory is supplemented with a minimal physicalist assumption, according to which the *external* (or third person) facts of consciousness are logical consequences of the physical facts, we can *prove* that the internal (or first person) facts are logical consequences of the physical facts as well. A corollary of this theorem is the logical impossibility of a zombie world (*i.e.* a world where all the physical facts of our world hold, but no internal fact of consciousness holds). These results pose a dilemma to Chalmers and all zombies' friends: if a zombie world is logically possible, then, contrary to Chalmers' position, the *external* facts of consciousness are not entailed by the physical facts or, equivalently, *psychological* consciousness is not reductively explainable by physics. Conversely, if psychological consciousness can be reductively explained by physics, then zombie worlds are logically impossible.

I conclude that Chalmers' intuition, according to which phenomenal consciousness is beyond the scope of any physical theory, is not tenable. Finally, I hint at a possible source of such erroneous perception, that is, not realizing that the hard problem of consciousness is not a problem of *scientific explanation* but, rather, a *philosophical* problem of conceptual *explication*.

1 IL PROBLEMA DELLA COSCIENZA FENOMENICA NELL'ULTIMO QUARTO DEL NOVECENTO

A partire dalla seconda metà del '900, in un primo momento con la diffusione di posizioni comportamentiste (Ryle 1949), poi con la teoria dell'identità (Place 1956; Smart 1959; Lewis 1966, 1972; Armstrong 1968, 1981b) e infine con l'affermarsi del funzionalismo (Putnam 1960, 1964, 1965, 1967a, 1967b), la filosofia della mente di tradizione analitica aveva posto sempre più in secondo piano la dimensione fenomenica, esperienziale, o soggettiva della coscienza. Infatti, secondo tali posizioni, la mente doveva essere studiata da un punto di vista strettamente oggettivo, esterno, o in terza persona. Per il comportamentismo si trattava di mettere in luce i rapporti nomologici diretti fra comportamenti esternamente osservabili; la teoria dell'identità riduceva gli stati mentali a stati fisici; e infine, per il funzionalismo, occorreva cogliere i rapporti fra input e output, esternamente accertabili, attraverso la mediazione di stati mentali interni, ma altrettanto oggettivi.¹

Tuttavia, verso la metà degli anni '70, il quadro cominciò lentamente a cambiare. Da un lato, furono elaborati una serie di argomenti, ormai divenuti classici,² che si proponevano di mostrare come la dimensione fenomenica, esperienziale, o soggettiva della coscienza sia tendenzialmente irriducibile a quella psicologica, funzionale, o oggettiva. Dall'altro, cominciarono ad interessarsi al problema della coscienza non soltanto psicologi e filosofi, ma anche neuroscienziati, fisici, matematici e, in generale, ricercatori attivi in quasi tutti i campi delle scienze cognitive.

Il crescente interesse per il problema della coscienza sfociò in un primo convegno che si tenne a Tucson nel 1994. I convegni di Tucson sono poi diventati appuntamenti periodici che si tengono con frequenza biennale.³ Retrospectivamente, il primo convegno di Tucson può essere considerato l'atto di nascita di un nuovo campo di ricerca interdisciplinare denominato *Consciousness Studies* (*Studi sulla coscienza*). Uno dei suoi centri propulsori è la *University of Arizona*, a Tucson, dove nel 1998 si è costituito il *Center for Consciousness Studies*, i cui direttori sono David Chalmers, Stuart Hameroff e Alfred Kaszniak. Ned Block è invece il presidente dell'ASSC (*Association for the Scientific Study of Consciousness*), che tiene anch'essa regolari congressi con cadenza annuale⁴. Infine, sono state fondate anche nuove riviste, espressamente dedicate agli studi sulla coscienza. Fra queste, il *Journal of Consciousness Studies*,⁵ *Psyche*,⁶ *Consciousness and Cognition*⁷ e *Consciousness and Emotion*.⁸

2 LA DISTINZIONE FRA COSCIENZA FENOMENICA E COSCIENZA PSICOLOGICA

Il problema della coscienza, nella forma in cui si è imposto nell'ultimo scorcio del '900, consiste prima di tutto nel riconoscere la distinzione, almeno concettuale, fra due tipi di fatti della coscienza: (i) i *fatti interni* (o *in prima persona*) della coscienza, o le *esperienze coscienti*; (ii) i *fatti esterni* (o *in terza persona*) della coscienza, o le *funzioni coscienti*. L'insieme dei fatti interni della coscienza è la *coscienza fenomenica*; l'insieme dei fatti esterni della coscienza è la *coscienza (neuro-)psicologica*.

La distinzione fra coscienza fenomenica e coscienza psicologica è stata messa in luce da una serie di argomenti classici: l'argomento del *What is it like to be?* (Nagel 1974), i *qualia assenti* (Block 1978) e lo *spettro invertito* (Block 1978; già presente in Locke 1690), l'argomento di *Mary* o della *conoscenza* (Jackson 1982), l'argomento degli *zombie fenomenici* (Chalmers 1996, e inoltre si veda la nota 2). Ciascuno di questi argomenti, oltre a chiarire la distinzione *concettuale* fra esperienze coscienti e funzioni coscienti, si propone anche, più o meno esplicitamente, di provare che tale distinzione è una distinzione *reale*.

Per comprendere in che cosa consista tale distinzione è forse opportuno richiamare l'esempio paradigmatico di fatto interno della coscienza, e cioè l'esperienza dei *qualia* o delle *qualità immediatamente vissute*.

L'esperienza dei qualia può essere convenientemente descritta attraverso il seguente schema generale:

(Q) l'esperienza cosciente di **qualcuno** che [vede / sente / gusta / odora / prova] la **qualità** di **qualcosa**

Sostituendo nello schema (Q) un valore determinato per i termini in grassetto e scegliendo il più appropriato fra i cinque verbi di percezione, si ottengono, per esempio, le seguenti descrizioni di specifiche esperienze di qualia:

1. *l'esperienza cosciente di Maria che vede il verde di un prato di montagna;*
2. *l'esperienza cosciente di Maria che gusta il dolce di una zolletta di zucchero;*
3. *l'esperienza cosciente di Giorgio che sente il timbro del suono di un violino;*
4. *l'esperienza cosciente di Giorgio che sente la rugosità di un tronco di un albero;*
5. *l'esperienza cosciente di Maria che odora il profumo di una rosa;*
6. *l'esperienza cosciente di Maria che prova il dolore di una martellata sul pollice;*

7. *l'esperienza cosciente di Giorgio che prova il piacere di fumare un sigaro;*
8. *l'esperienza cosciente di Giorgio che prova il caldo di una giornata di agosto;*
9. *l'esperienza cosciente di Maria che prova il freddo di un pezzo di ghiaccio.*

Un minimo di riflessione su questi esempi basterà a convincerci che, a ciascuna esperienza di qualia da essi descritta, corrisponde⁹ un fatto esterno della coscienza. Considerando l'esempio 6, al dolore coscientemente vissuto dall'interno da Maria, corrisponde uno specifico stato di attivazione di una parte rilevante del suo sistema nervoso periferico e/o centrale, nonché, tipicamente, una serie di comportamenti (verbali, motori, etc.) esternamente accertabili. E' proprio questo complesso (stato neurofisiologico + eventuali comportamenti) che, nella sua organizzazione funzionale, costituisce il fatto esterno, o la funzione cosciente, corrispondente alla coscienza interna del dolore da parte di Maria.

Inoltre, l'osservazione precedente sulla corrispondenza fra esperienze dei qualia e funzioni coscienti può in effetti essere generalizzata a tutti i fatti di coscienza: un qualsiasi fatto di coscienza, se è interno, ha un corrispondente esterno;¹⁰ viceversa, se è esterno, ha un corrispondente interno.¹¹

Chiarito in questo modo che i fatti interni e i fatti esterni della coscienza sono almeno concettualmente distinti, rimane tuttavia aperta la questione se essi lo siano anche realmente. In altre parole: i due concetti di fatto interno ed esterno della coscienza individuano davvero due insiemi diversi di fatti, o invece questi concetti sono soltanto due modi diversi di descrivere un identico insieme? Riprenderò successivamente questa questione (vedi sez. 9.1).

3 I PROBLEMI DELLA COSCIENZA SECONDO CHALMERS

Forse il più importante contributo di Chalmers (1996) al dibattito sulla coscienza consiste nella nuova formulazione del problema, che sposta il centro dell'attenzione proprio sulla distinzione fra coscienza psicologica e coscienza fenomenica. Seguendo questa impostazione, Chalmers parla di *problema facile* e *problema difficile* della coscienza. Per gli scopi di questa indagine, tuttavia, è opportuno distinguere chiaramente fra due diverse versioni del problema difficile di Chalmers, che ho deciso di chiamare, rispettivamente, *problema difficile* e *problema superdifficile*.

3.1 IL PROBLEMA FACILE DELLA COSCIENZA

Esso consiste nel *trovare e sviluppare una teoria adeguata della coscienza neuro-psicologica*. Per quanto riguarda le sue diverse formulazioni specifiche, questo problema è affrontato e risolto con crescente successo dalla psicologia scientifica, dalle neuroscienze, o dalle scienze cognitive. Secondo Chalmers, al di là dei diversi livelli di descrizione a cui esse si collocano, tali soluzioni particolari si basano tutte sul *paradigma funzionalista*, secondo cui i fatti esterni della coscienza possono sempre essere spiegati scoprendo le opportune relazioni causali fra input e output esterni e stati mentali (o neurofisiologici) interni.

Anche dal punto di vista del problema mente-corpo, la coscienza psicologica non presenta per Chalmers una difficoltà speciale. Infatti, è in linea di principio chiaro come una teoria di tipo funzionale possa essere riduttivamente spiegata da una teoria fisica. Quindi, se i fatti esterni della coscienza sono *funzioni coscienti*, essi sono *riduttivamente spiegabili*, almeno in linea di principio, in termini di *fatti fisici*.

3.2 IL PROBLEMA DIFFICILE DELLA COSCIENZA

Esso consiste nel *trovare e sviluppare una teoria adeguata della coscienza fenomenica*. Questo problema, secondo Chalmers, è stato largamente trascurato dalla psicologia scientifica, dalle neuroscienze e dalle scienze cognitive. Ma, quel che forse è peggio, al momento non si ha la più pallida idea di come il paradigma funzionalista potrebbe essere utilizzato per arrivare a spiegare la coscienza fenomenica. E questo perché tutte le volte che una spiegazione funzionalista è stata tentata, tale spiegazione, invece di gettare luce sull'*explanandum*, sembra inevitabilmente arrivare a cogliere soltanto i suoi correlati esterni e cioè, in ultima analisi, soltanto la coscienza psicologica.

Conseguentemente, non si vede neanche come sarebbe possibile una spiegazione riduttiva dei fatti interni della coscienza in termini di fatti fisici. Infatti, tale spiegazione dovrebbe prima passare per una spiegazione funzionalista della coscienza fenomenica; ma, essendo questa un mistero, una spiegazione riduttiva delle esperienze coscienti in termini di fatti fisici risulta un mistero ancora maggiore.

3.3 IL PROBLEMA SUPERDIFFICILE DELLA COSCIENZA

In realtà, per Chalmers, i fatti interni della coscienza *non sono spiegabili riduttivamente* nei termini di quelli fisici e quindi neanche in base a funzioni

coscienti. Questo spiega perché fino ad oggi non c'è stato alcun progresso verso la costruzione di una teoria scientifica delle esperienze coscienti. Tale teoria, infatti, *non può* basarsi sul paradigma funzionalista, perché sappiamo che ogni teoria di quel tipo è in principio spiegabile riduttivamente per mezzo di teorie fisiche.

Ne segue che, per poter affrontare il problema difficile della coscienza, è prima necessario costruire un nuovo paradigma, che risulti appropriato per lo studio dei fatti interni della coscienza, come quello funzionalista lo è per i fatti esterni. In particolare si tratta di:

1. costruire un paradigma scientifico radicalmente nuovo, che permetta poi l'elaborazione di una teoria adeguata della coscienza fenomenica;
2. tale paradigma deve essere sviluppato all'interno di un quadro metafisico antiriduzionistico in senso forte¹² e quindi dualistico; ci sono cioè due tipi di fatti fondamentali: *esperienze coscienti* e *fatti fisici*;
3. si pone quindi, in generale, il problema dei rapporti fra questi due tipi di fatti fondamentali;
4. in particolare, per quanto riguarda il rapporto di causalità, si pone il problema dell'efficacia causale delle esperienze coscienti. Come è noto, la soluzione di Chalmers consiste in una forma di epifenomenalismo;
5. si devono anche trovare nuovi metodi adeguati allo studio della coscienza fenomenica. Bisogna quindi considerare, oltre ai metodi intersoggettivi o in terza persona tipici delle scienze naturali e della psicologia scientifica, anche *metodi soggettivi* o *in prima persona*. Ne consegue un rinnovato interesse per il filone storicista-ermeneutico-fenomenologico-esistenzialista della filosofia europea del '900.

Considerando tutti i punti 1-5, sembra però che il problema prospettato da Chalmers¹³ non risulti soltanto *estremamente* difficile, ma quasi intrattabile. Ma allora, è forse più opportuno e interessante:

- affrontare, non negare, il problema della costruzione di una teoria della coscienza fenomenica;
- ma condurre tale costruzione all'interno di un quadro rigorosamente fisicalista.

Tuttavia, se Chalmers ha ragione nel sostenere che la coscienza fenomenica non è riduttivamente spiegabile in termini di fatti fisici, questa linea di ricerca non può essere perseguita. Occorre quindi prima di tutto analizzare l'argomento chiave su cui Chalmers basa tale convinzione. Come è noto, si tratta dell'argomento degli zombie fenomenici. Per averne una comprensione sufficientemente profonda, però, è necessario premettere un minimo quadro formale.

4 L'ARGOMENTO DEGLI ZOMBIE FENOMENICI E LA SFIDA DI CHALMERS

La struttura matematica di cui abbiamo bisogno per formulare e capire l'argomento degli zombie è una quintupla $(\mathbf{D}, \mathbf{P}, \mathbf{Q}, \mathbf{W}, \mathbf{w}^*)$ tale che:

1. \mathbf{D} è un insieme non vuoto, detto *dominio*;
2. \mathbf{P} è un insieme non vuoto; ciascun elemento di \mathbf{P} è una coppia (P, n) tale che n è un numero naturale e $1 \leq n$; per comodità, ciascuna di tali coppie viene indicata con la notazione P^n ; ogni elemento di \mathbf{P} è detto una *proprietà* e il numero n è detto la sua *arietà* (o il suo *numero di posti*). Si noti che la precisa natura del primo elemento di ciascuna proprietà non è specificata, e tale natura non è in alcun modo rilevante in tutto ciò che segue. Inoltre, quando l'arietà n di una proprietà P^n risulti chiara dal contesto, tale proprietà può essere indicata con la più semplice notazione P ;
3. \mathbf{Q} è un sottoinsieme di \mathbf{P} ; \mathbf{Q} è detto *l'insieme delle proprietà fisiche*;
4. \mathbf{W} è un insieme di funzioni; ciascuna di esse, a ciascuna proprietà P^n appartenente a \mathbf{P} , associa un insieme di n -uple $\{(x_1, x_2, \dots, x_n)\}$ tale che, per ogni n -upla di tale insieme, ciascun suo elemento, x_i , appartiene a \mathbf{D} ; ogni funzione w appartenente a \mathbf{W} è detta un *mondo possibile*; l'insieme di n -uple assegnato da w a P^n è indicato con il simbolo P^n_w ed è detto *l'estensione di P^n in mondo w* ;
5. $\mathbf{w}^* \in \mathbf{W}$; \mathbf{w}^* è detto *il nostro mondo o il mondo attuale*.

Questa struttura ci permette di dare una serie di definizioni:

- [1] f è un *fatto possibile* (o, più brevemente, *un fatto*) sse:
 f è una coppia ordinata del tipo $(P^n, (x_1, x_2, \dots, x_n))$, dove $P^n \in \mathbf{P}$ e (x_1, x_2, \dots, x_n) è una n -upla ($n \geq 1$) di elementi di \mathbf{D} ;
la proprietà P^n è detta il *predicato di f* , e la n -upla (x_1, x_2, \dots, x_n) è detta il *soggetto di f* ; si tenga presente che, nel caso in cui $n = 1$, il soggetto si riduce ad una n -upla di un solo elemento, (x_1) , che, per definizione, è identica all'elemento stesso x_1 ; inoltre, per semplificare la notazione, conveniamo di scrivere $P^n(x_1, x_2, \dots, x_n)$ al posto di $(P^n, (x_1, x_2, \dots, x_n))$.
- [2] f è un *fatto fisico* sse:
 f è un fatto e il predicato di f appartiene a \mathbf{Q} .
- [3] f *vale in (mondo) w* sse:
 $f = P^n(x_1, x_2, \dots, x_n)$ è un fatto, $w \in \mathbf{W}$, e $(x_1, x_2, \dots, x_n) \in P^n_w$.

Possiamo adesso definire la relazione di conseguenza logica, indicata dal simbolo \models , che può valere fra due qualsiasi insiemi di fatti F_1 e F_2 . La

notazione $F_1 \models F_2$ si legge: da F_1 segue logicamente F_2 (o anche: F_2 è conseguenza logica di F_1).

[4] $F_1 \models F_2$ sse:

per ogni mondo possibile w , se per ogni f_1 appartenente a F_1 , f_1 vale in w , allora per ogni f_2 appartenente a F_2 , f_2 vale in w ;

si intende inoltre che la relazione di conseguenza logica è definita in modo analogo anche nei casi: $F \models f$ (da un insieme di fatti segue un fatto); $f \models F$ (da un fatto segue un insieme di fatti); $f_1 \models f_2$ (da un fatto segue un fatto).

Abbiamo così costruito il quadro formale che ci permette di formulare e discutere l'argomento degli zombie. Indichiamo con F^* l'insieme di tutti i fatti fisici che valgono nel nostro mondo, con CF^* la coscienza fenomenica del nostro mondo (cioè, l'insieme di tutti i fatti interni della coscienza che valgono nel nostro mondo) e con CP^* la coscienza psicologica del nostro mondo (cioè, l'insieme di tutti i fatti esterni della coscienza che valgono nel nostro mondo).¹⁴

La tesi dell'argomento degli zombie è allora

[Z] $\neg(F^* \models CF^*)$;

inoltre, siccome Chalmers ammette

[R] $F^* \models CP^*$,

da [Z] e dall'assunzione [R], segue anche il corollario

[Z₁] $\neg(CP^* \models CF^*)$.

Prima di procedere oltre, si tenga presente che l'assunzione [R] esprime un'ipotesi physicalista *minimale*, in quanto essa afferma soltanto che la coscienza psicologica è riduttivamente spiegabile in base ai fatti fisici,¹⁵ mentre lascia aperta la questione della spiegabilità riduttiva della coscienza fenomenica.

Vediamo adesso l'argomento vero e proprio:

1. si consideri un mondo (*mondo zombie*) in cui valgano tutti e soli i fatti fisici che valgono nel nostro mondo, ma in cui non valga alcun fatto interno della coscienza;
2. tale mondo, in quanto ben definito, è ovviamente concepibile; ma è anche uno dei mondi possibili in \mathbf{W} ?
3. Lo è certamente, perché l'affermazione dell'esistenza di un tale mondo possibile sembra sicuramente non contraddittoria;

4. ma allora, se il mondo zombie è uno dei mondi possibili, dalla def. [4] segue che $\neg(F^* \models CF^*)$ sse CF^* non è vuoto;¹⁶
 5. nel nostro mondo w^* ci sono fatti interni di coscienza;
 6. conseguentemente, da 5 e 4 segue che $\neg(F^* \models CF^*)$.
- Q.E.D.

La forza di questa forma dell'argomento degli zombie è che esso rovescia l'onere della prova su chiunque volesse sostenere $F^* \models CF^*$. Vediamo esattamente perché.

Notiamo prima di tutto che questo modo di porre l'argomento evita qualsiasi discussione sulla *concepibilità* del mondo zombie.¹⁷ Che tale mondo sia concepibile è assicurato dal fatto che il concetto di tale mondo può essere definito in modo adeguato all'interno del quadro formale assunto. Per rendere completamente chiaro questo punto è forse opportuno formulare esplicitamente tale definizione:

- [5] w^z è una copia zombie di w sse:
 sia w che w^z sono mondi possibili e, per ogni f_1 , per ogni f_2 , se f_1 è un fatto fisico e f_2 è un fatto interno di coscienza, allora f_2 non vale in w^z e (f_1 vale in w^z sse f_1 vale in w).

La questione cruciale non è quindi se una copia zombie del nostro mondo sia o meno concepibile ma, piuttosto, se tale copia esista.¹⁸ Ma, trattandosi qui di pura esistenza di tipo logico-matematico, tale questione si riduce a verificare che l'asserzione dell'esistenza di un mondo zombie sia consistente con i principi formali assunti.¹⁹ E, in effetti, almeno da un punto di vista intuitivo, non si vede come dalla congiunzione di:

1. tale asserzione;
2. le cinque condizioni assunte sull'ontologia (**D, P, Q, W, w***);
3. eventuali altre condizioni implicitamente o esplicitamente assunte su tale ontologia;²⁰
4. gli usuali principi di teoria degli insiemi;

possa seguire una contraddizione.

In ogni caso, anche se la congiunzione di 1, 2, 3 e 4 risultasse contraddittoria, ciò non è affatto immediatamente evidente. Perciò, chiunque volesse sostenere la spiegabilità riduttiva della coscienza in base ai fatti fisici (e cioè: $F^* \models CF^*$) deve prima mostrare che l'apparenza di non contraddittorietà di tale congiunzione è, appunto, solo un'apparenza.

L'argomento degli zombie pone quindi una sfida a chiunque si proponga di:

- costruire una teoria della coscienza fenomenica,
- che permetta di sostenere la spiegabilità riduttiva della coscienza fenomenica in base ai fatti fisici (e cioè: $F^* \models CF^*$).

La *sfida di Chalmers* è:

- mostrare che l'asserzione dell'esistenza di una copia zombie del nostro mondo è in contraddizione con la congiunzione dei principi dell'ontologia e della teoria degli insiemi.

L'intento di questo lavoro è proprio quello di accettare la sfida di Chalmers. Cominciamo quindi dal primo punto, e cioè dal gettare le fondamenta di una possibile teoria della coscienza fenomenica.

5 CHE COS'È UN'ESPERIENZA IN GENERALE?

Se il nostro obiettivo è quello di spiegare che cos'è un *fatto interno della coscienza*, o un'*esperienza cosciente*, è opportuno affrontare prima il problema più semplice di che cosa sia un *fatto interno* in generale, o un'*esperienza*.²¹

Partiamo dall'analisi di un esempio tipico, domandandoci:

sotto quali condizioni il fatto che
Giorgio corre = $C(g)$
 può essere considerato un'esperienza per qualcuno?

Sembra piuttosto chiaro che questa domanda non ha una risposta univoca, ma che invece essa dipende dal punto di vista, o dal sistema di riferimento, che consideriamo: tale fatto è un'esperienza per Giorgio, ma non lo è per me, o per chiunque altri che non sia Giorgio.

Ma che cos'è che Giorgio ha di esclusivo nei confronti del suo correre? Non è forse, questo qualcosa proprio di Giorgio e di nessun altro, *il suo essere collocato internamente a tale fatto come soggetto*? In effetti, sembra proprio di sì. Decidiamo quindi di definire:

[6] *f* è un'esperienza (o un fatto interno) per *x* sse:
f è un fatto e *x* è il soggetto di *f*.

[7] *f* è un fatto esterno per *x* sse:
f è un fatto e *x* non è il soggetto di *f* e, per qualche *n* tale che $1 \leq n$, *x* è una *n*-upla di elementi di **D**.²²

Notiamo subito che dalla def. [6] segue immediatamente:

Teorema 1 [di privatezza dell'esperienza]

Per ogni x , per ogni y , per ogni f , se f è un'esperienza per x e f è un'esperienza per y , allora $x = y$.

Dimostrazione

La tesi segue dalle def. [6] e [1].

Q.E.D.

Definiamo inoltre:

[8] f è un fatto interno (o un'esperienza) sse:
esiste x tale che f è un fatto interno per x .

[9] f è un fatto esterno sse:
esiste x tale che f è un fatto esterno per x .

Domandiamoci adesso se esistano fatti che sono *intrinsecamente* o interni o esterni, e cioè che sono interni ma non esterni o esterni ma non interni. Un attimo di riflessione sulle definizioni appena date basterà a convincerci che la risposta è *no*. Infatti vale:

Teorema 2 [di relatività dell'esperienza]

Ogni fatto è sia interno che esterno.

Dimostrazione

Sia f un fatto arbitrario e x il suo soggetto. Allora, per la def. [6], f è un fatto interno per x ; quindi, per la def. [8], f è un fatto interno. D'altra parte, per la def. [7], f è un fatto esterno per la coppia (x, x) ; quindi, per la def. [9], f è un fatto esterno.

Q.E.D.

6 DUE DIFFICOLTÀ NELLA DEFINIZIONE DI ESPERIENZA

Nella sezione precedente ho proposto di analizzare il concetto generale di esperienza basandomi sull'idea che un fatto risulta o meno un'esperienza per una certa entità x a seconda che x sia o meno collocata internamente a tale fatto come soggetto (def. [6]).

Questa definizione sembra funzionare molto bene nel caso di fatti monadici, cioè fatti il cui predicato sia una proprietà ad un posto. Tuttavia, se consideriamo anche fatti relazionali (il cui predicato è una proprietà a

due o più posti) si presenta subito una *difficoltà*, che ho deciso di chiamare *delle esperienze assenti*.

Inoltre, considerando altri esempi, e applicando ad essi la def. [6], ci si trova di fronte ad alcuni casi di esperienze che, ad un primo esame, risultano intuitivamente poco convincenti. Questa è la *difficoltà delle esperienze strane*.

Questa sezione è dedicata all'esposizione e alla risoluzione di queste due difficoltà.

6.1 ESPERIENZE ASSENTI

Consideriamo il fatto relazionale:

$$\textit{Giorgio ama Maria} = A(g, m)$$

Secondo la def. [6], questo fatto è un'esperienza *esclusivamente* per la coppia (g, m) che è il soggetto del fatto. Ma sembra invece che sia Giorgio che Maria, in quanto *individui* collocati internamente al fatto, debbano, in un certo senso, esperirlo.

D'altra parte, è chiaro che Giorgio e Maria non possono ciascuno esperire quello stesso fatto, perché ciò è vietato dal teorema di privatezza dell'esperienza. Ma allora, che cosa esperiscono, individualmente, Giorgio e Maria?

La risposta intuitiva a questa difficoltà è che ciascuno di loro esperisce un *suo* fatto, diverso da quello dell'altro e diverso anche da quello relativo alla coppia Giorgio e Maria. Tuttavia, tutti e tre questi fatti risultano logicamente equivalenti.

Più in dettaglio, ciò significa che, in primo luogo, la *coppia Giorgio e Maria* esperisce il fatto che

la coppia Giorgio e Maria è nella relazione x ama y ;

inoltre, *Giorgio* esperisce il fatto che

Giorgio ama *Maria*;

e infine *Maria* esperisce il fatto che

Maria è amata da *Giorgio*.

Secondo questa analisi, ci sono quindi tre diversi fatti esperiti. Utilizzando la sottolineatura per mettere in evidenza il *soggetto logico* di tali fatti, li possiamo rappresentare in modo uniforme nel seguente modo:

1. Giorgio ama Maria
2. Giorgio ama Maria
3. Giorgio ama Maria

È inoltre ovvio che tutti e tre questi fatti risultano, da un punto di vista intuitivo, logicamente equivalenti.

Tuttavia, il problema che si pone rispetto a questa analisi è che il quadro formale sviluppato finora non permette di rappresentare adeguatamente la *specificata* forma logica di ciascuno dei fatti 1, 2, 3 e le loro relazioni reciproche. Si pone quindi la necessità di ampliare tale quadro formale in modo da poter tradurre al suo interno la soluzione proposta.

Abbastanza sorprendentemente, lo strumento formale che permette di operare tale traduzione in modo del tutto naturale è il λ -calcolo. Infatti, utilizzando il formalismo del λ -calcolo possiamo generare i fatti 2 e 3 a partire da 1. Vediamo come.

Prima di tutto, rappresentiamo formalmente il fatto 1:

1. Giorgio ama Maria = $A(g, m)$

Questo è il fatto di partenza, da cui si ottengono gli altri due utilizzando le due operazioni di λ -astrazione e istanziazione.

Infatti, λ -astraendo su g si ottiene la proprietà:

$$x \text{ ama } Maria = [\lambda(g)A(g, m)]$$

e quindi, istanziano tale proprietà con g stesso, si ha:

2. Giorgio ama Maria = $[\lambda(g)A(g, m)](g)$

Analogamente, λ -astraendo su m si ottiene la proprietà:

$$Giorgio \text{ ama } y = [\lambda(m)A(g, m)]$$

e istanziano tale proprietà con m stessa si ha:

3. Giorgio ama Maria = $[\lambda(m)A(g, m)](m)$

Gli usuali assiomi del λ -calcolo ci assicurano proprio che (i) l'operazione di λ -astrazione è lecita e che (ii) 1, 2 e 3 sono tutti logicamente equivalenti.

Per adattare tali assiomi al particolare formalismo utilizzato qui, estendiamo prima di tutto l'ontologia, introducendo in essa λ come nuova entità primitiva. Dal punto di vista insiemistico, l'operatore λ può essere pensato come una funzione parziale a due posti, il cui codominio è \mathbf{P} , il cui primo dominio è l'unione di tutti i prodotti cartesiani k -esimi di \mathbf{D} (per $1 \leq k$), ed il cui secondo dominio è l'insieme di tutti i fatti possibili della struttura. Assumiamo quindi due nuovi assiomi sull'ontologia estesa (\mathbf{D} , \mathbf{P} , \mathbf{Q} , \mathbf{W} , \mathbf{w}^* , λ):

6. [chiusura di \mathbf{P} rispetto alla λ -astrazione di proprietà da fatti]

λ è una funzione parziale a due posti, il cui codominio è \mathbf{P} , il cui primo dominio è l'unione di tutti i prodotti cartesiani k -esimi di \mathbf{D} (per $1 \leq k$), ed il cui secondo dominio è l'insieme di tutti i fatti possibili della struttura; $\lambda((x_{i,1}, x_{i,2}, \dots, x_{i,k}), P^n(x_1, x_2, \dots, x_n))$ è definito sse $P^n(x_1, x_2, \dots, x_n)$ è un qualunque fatto e $(x_{i,1}, x_{i,2}, \dots, x_{i,k})$ è una k -upla di esattamente k elementi diversi di \mathbf{D} , ciascuno dei quali occorre nella n -upla (x_1, x_2, \dots, x_n) ; se $\lambda((x_{i,1}, x_{i,2}, \dots, x_{i,k}), P^n(x_1, x_2, \dots, x_n))$ è definito, esso è una proprietà k -aria appartenente a \mathbf{P} ; conveniamo inoltre di abbreviare $\lambda((x_{i,1}, x_{i,2}, \dots, x_{i,k}), P^n(x_1, x_2, \dots, x_n))$ con $[\lambda(x_{i,1}, x_{i,2}, \dots, x_{i,k})P^n(x_1, x_2, \dots, x_n)]$;

Prima di enunciare il secondo assioma relativo a λ , premettiamo la seguente definizione:

[10] / è una funzione parziale a tre posti, il cui codominio ed i cui tre domini sono, ciascuno, l'unione di tutti i prodotti cartesiani n -esimi di \mathbf{D} (per $1 \leq n$); i valori di / sono definiti come segue:

se (x_1, x_2, \dots, x_n) è una n -upla ($1 \leq n$) di elementi di \mathbf{D} , $(x_{i,1}, x_{i,2}, \dots, x_{i,k})$ è una k -upla ($1 \leq k \leq n$) di esattamente k elementi distinti di \mathbf{D} ciascuno dei quali occorre in (x_1, x_2, \dots, x_n) , e (y_1, y_2, \dots, y_k) è una k -upla di elementi (non necessariamente distinti) di \mathbf{D} , allora $/((x_1, x_2, \dots, x_n), (x_{i,1}, x_{i,2}, \dots, x_{i,k}), (y_1, y_2, \dots, y_k))$ è definito, ed esso è uguale alla n -upla di elementi di \mathbf{D} ottenuta sostituendo l'elemento y_j ($1 \leq j \leq k$) nella n -upla (x_1, x_2, \dots, x_n) a tutte le occorrenze in essa dell'elemento $x_{i,j}$; altrimenti, il valore di / non è definito; conveniamo inoltre di scrivere $[(x_1, x_2, \dots, x_n) / (x_{i,1} / y_1, x_{i,2} / y_2, \dots, x_{i,k} / y_k)]$ al posto di $/((x_1, x_2, \dots, x_n), (x_{i,1}, x_{i,2}, \dots, x_{i,k}), (y_1, y_2, \dots, y_k))$.

Possiamo adesso enunciare il secondo assioma relativo a λ :

7. [condizioni di validità dell'instanziazione di una proprietà ottenuta per λ -astrazione]

se (x_1, x_2, \dots, x_n) è una n -upla ($1 \leq n$) di elementi di \mathbf{D} , $(x_{i,1}, x_{i,2}, \dots, x_{i,k})$ è una k -upla ($1 \leq k \leq n$) di esattamente k elementi distinti di \mathbf{D} ciascuno dei quali occorre in (x_1, x_2, \dots, x_n) , e (y_1, y_2, \dots, y_k) è una k -upla di elementi (non necessariamente distinti) di \mathbf{D} , allora, per ogni mondo possibile w , $[\lambda(x_{i,1}, x_{i,2}, \dots, x_{i,k})P^n(x_1, x_2, \dots, x_n)](y_1, y_2, \dots, y_k)$ vale in w sse $P^n[(x_1, x_2, \dots, x_n) / (x_{i,1} / y_1, x_{i,2} / y_2, \dots, x_{i,k} / y_k)]$ vale in w .

Tenendo presente quanto mostrato sopra, notiamo infine che i due nuovi assiomi sull'ontologia (assiomi 6 e 7) permettono di risolvere in modo elegante e del tutto naturale la difficoltà delle esperienze assenti.

6.2 ESPERIENZE STRANE

Consideriamo il fatto:

$$\textit{il treno corre} = C(t)$$

In forza della def. [6], risulta che tale fatto è un'esperienza per il treno. Ma come si può veramente dire che un *treno* esperisce il suo correre?

La stessa difficoltà sorge anche al riguardo di un qualsiasi fatto relazionale. Per esempio, si consideri di nuovo il fatto:

$$\textit{Giorgio ama Maria} = A(g, m)$$

Secondo la def. [6], tale fatto è un'esperienza per la coppia Giorgio e Maria. Ma, anche in questo caso, sembra a prima vista piuttosto strano affermare che una *coppia* esperisce la propria relazione d'amore.

La mia risposta a questo tipo di difficoltà è che essa è soltanto apparente. Che il treno esperisca il suo correre sembra paradossale perché di solito noi non distinguiamo fra un'esperienza e l'esperienza di esser coscienti di quell'esperienza.

Ora, che $C(t)$ sia un fatto direttamente esperito dal treno non è affatto paradossale perché, in accordo con la def. [6], ciò significa soltanto che il treno è collocato internamente a tale fatto nel ruolo di soggetto. Sarebbe però certamente paradossale affermare che il treno esperisca tale fatto e che *inoltre* sia anche cosciente di tale esperienza. Il punto da tenere presente è cioè che *qualsiasi cosa ha esperienze*, ma solo *poche cose hanno esperienze coscienti* (e fra queste, ovviamente, né i treni né le coppie sono compresi).²³

Quindi, prima di tutto, dobbiamo fare un'importante distinzione: esperienza e coscienza non sono la stessa cosa, nel senso che ci sono esperienze che non sono oggetto di coscienza per il loro soggetto (es_1 : il treno non è cosciente del *suo correre*). Analogamente, ci sono anche fatti esterni di cui non si ha coscienza (es_2 : siccome Giorgio dorme, Giorgio non è cosciente che *il semaforo è rosso*). E, infine, ci possono essere sia esperienze che fatti esterni di cui si ha coscienza (es_3 : Giorgio è cosciente del *suo correre*; es_4 : Giorgio è cosciente che *Maria parla*).

Da un altro punto di vista, tuttavia, coscienza ed esperienza non possono essere disgiunte: è questo il caso di tutti i fatti interni di coscienza, cioè le *esperienze* coscienti, che sono tutte un tipo particolare di esperienza. Ma, esattamente, che tipo di esperienza sono le esperienze coscienti?

7 CHE COS'È UN'ESPERIENZA COSCIENTE?

La mia risposta a questa domanda è che le esperienze coscienti, essendo esperienze, sono prima di tutto, per la def. [6], fatti. Tuttavia, esse sono un tipo particolare di fatti, che chiamo *fatti di coscienza*. E che cos'è un fatto di coscienza?

Per rispondere a questa seconda domanda, consideriamo alcuni esempi tipici. I seguenti sono tutti fatti di coscienza:

1. Giorgio è cosciente del suo correre
2. Giorgio è cosciente che Maria parla
3. Maria è cosciente del suo parlare
4. Maria è cosciente che Giorgio corre

Osservando gli esempi notiamo:

- (a) in ciascuno di questi fatti è presente la relazione binaria x è *cosciente di* f_0 che sussiste fra un elemento x del dominio e un fatto f_0 ;
- (b) il *soggetto* di ciascun fatto di coscienza è x (sottolineato negli esempi);
- (c) il fatto f_0 costituisce l'*oggetto* o il *contenuto* del fatto di coscienza;
- (d) i fatti 1 e 3 sono simili, in quanto in essi il soggetto del fatto di coscienza ed il soggetto del suo contenuto sono identici; fatti come questi possono quindi essere detti *fatti di coscienza interna*, o *autocoscienza*.
- (e) nei fatti 2 e 4, invece, il soggetto del fatto ed il soggetto del suo contenuto sono diversi; fatti come questi possono quindi essere detti *fatti di coscienza esterna*.

Tenendo presente l'osservazione (a) possiamo prima di tutto estendere ulteriormente l'ontologia, aggiungendo una speciale relazione binaria, **C**, che può valere soltanto fra un elemento del dominio e un fatto. Intuitivamente **C** è da identificarsi con la relazione *essere coscienti di*. Assumiamo quindi il seguente assioma sull'ontologia (**D**, **P**, **Q**, **W**, **w***, **λ**, **C**):

8. [assioma della coscienza]

per ogni fatto f , $f \in \mathbf{D}$,²⁴ **C** appartiene a **P**, **C** è una proprietà a due posti e, per ogni w , per ogni x , per ogni f_0 , se **C**(x , f_0) vale in w , allora f_0 è un fatto.

Tenendo poi presenti le osservazioni (b) e (c), possiamo, utilizzando la λ -astrazione e l'assioma **8**, definire un fatto di coscienza nel seguente modo:

[11] f è un fatto di coscienza sse:

esiste x , esiste f_0 tale che $f = [\lambda(x)\mathbf{C}(x, f_0)](x)$;

f_0 è detto l'oggetto o il contenuto del fatto di coscienza f .

Inoltre, tenendo presenti le osservazioni (d) ed (e), definiamo anche:

[12] f è un fatto di coscienza interna (o f è un fatto di autocoscienza) sse:

f è un fatto di coscienza ed il soggetto di f è identico a quello del suo contenuto.

[13] f è un fatto di coscienza esterna sse:

f è un fatto di coscienza ed il soggetto di f è diverso da quello del suo contenuto.

Utilizzando adesso le definizioni [6] e [11] possiamo finalmente definire le esperienze coscienti per x :

[14] f è un fatto interno di coscienza per x (o f è un'esperienza cosciente per x) sse:

f è un'esperienza per x e f è un fatto di coscienza.²⁵

Analogamente, utilizzando [7] e [11], definiamo le funzioni coscienti per x :

[15] f è un fatto esterno di coscienza per x (o f è una funzione cosciente per x) sse:

f è un fatto esterno per x e f è un fatto di coscienza.

Incrociando infine prima la def. [14] e poi la [15] con le due def. [12] e [13], otteniamo anche le definizioni dei seguenti concetti:

[16] *f* è un fatto interno di coscienza interna per *x* (o *f* è un'esperienza autocosciente per *x*) sse:

f è un fatto interno di coscienza per *x* e *f* è un fatto di coscienza interna;

[17] *f* è un fatto interno di coscienza esterna per *x* (o *f* è un'esperienza cosciente esterna per *x*) sse:

f è un fatto interno di coscienza per *x* e *f* è un fatto di coscienza esterna;

[18] *f* è un fatto esterno di coscienza interna per *x* (o *f* è una funzione autocosciente per *x*) sse:

f è un fatto esterno di coscienza per *x* e *f* è un fatto di coscienza interna;

[19] *f* è un fatto esterno di coscienza esterna per *x* (o *f* è una funzione cosciente esterna per *x*) sse:

f è un fatto esterno di coscienza per *x* e *f* è un fatto di coscienza esterna.

Esempi

(1°) gli esempi 1 e 3 sopra sono fatti di coscienza interna;

(2°) gli esempi 2 e 4 sopra sono fatti di coscienza esterna;

(3°) l'esempio 1 è un fatto interno di coscienza interna per Giorgio e un fatto esterno di coscienza interna per Maria;

(4°) l'esempio 3 è un fatto interno di coscienza interna per Maria e un fatto esterno di coscienza interna per Giorgio;

(5°) l'esempio 2 è un fatto interno di coscienza esterna per Giorgio e un fatto esterno di coscienza esterna per Maria;

(6°) l'esempio 4 è un fatto interno di coscienza esterna per Maria e un fatto esterno di coscienza esterna per Giorgio.

8 PROPOSTA DI ANALISI DEL CASO PARADIGMATICO DI ESPERIENZA COSCIENTE: L'ESPERIENZA DEI QUALIA

Per saggiare se le definizioni date nella sezione precedente siano adeguate, vediamo adesso se esse ci permettono di analizzare il caso paradigmatico di esperienza cosciente, e cioè le esperienze dei *qualia* o delle qualità immediatamente vissute. Consideriamo quindi la seguente esperienza di una qualità vissuta:

l'esperienza cosciente di Maria che prova il dolore di una martellata sul pollice = ?

La mia proposta di analisi è la seguente. La martellata sul pollice causa il fatto che

Maria è in un particolare stato neurofisiologico $d = [\lambda(m)S(m, d)](m) = f_d$

Il soggetto di tale fatto è Maria stessa e quindi, per la def. [6], esso è un'esperienza per Maria.

Il fatto che Maria sia in stato neurofisiologico d è a sua volta causa, sotto condizioni normali,²⁶ del fatto che

Maria è conscia del suo essere in stato neurofisiologico d

Il soggetto di questo fatto è ancora Maria. Tale fatto può quindi essere analizzato come:

Maria è conscia del suo essere in stato neurofisiologico $d =$
 $= [\lambda(m)C(m, f_d)](m) = f$

In base alla def. [16], f risulta quindi un'esperienza autocosciente per Maria, il cui contenuto è il suo essere nello stato d , causato dalla martellata sul pollice (cioè, il contenuto dell'esperienza cosciente di Maria f è la sua esperienza f_d , causata dalla martellata sul pollice). Propongo quindi di identificare f con ?²⁷

Si noti infine che lo stesso tipo di analisi può essere condotta per qualunque altra esperienza di qualia descritta dallo schema (Q) (vedi sez. 2).

9 TEORIA DELLA COSCIENZA FENOMENICA E FISCALISMO

Nelle sezioni precedenti (sez. 5-8) ho costruito le basi di una possibile teoria della coscienza fenomenica. Lo scopo della sez 9.1 è duplice: (i) inserire tale teoria in un quadro fiscalista minimale (espresso dall'assioma **9**, o **9.1**, o **9.2**) e quindi (ii) raccogliere la sfida di Chalmers (vedi sez. 4), dimostrando la spiegabilità riduttiva fisica delle esperienze coscienti (teorema 4 o 4.1) e, conseguentemente, l'impossibilità logica dei mondi zombie (teorema 5 o 5.1). Infine, dopo aver introdotto il concetto di *copia angelica*²⁸ di un mondo possibile, dimostrerò anche l'impossibilità logica dei mondi angelici (teorema 6.2).

Nella sez. 9.2, invece, farò vedere come, assumendo un'ipotesi fiscalista più forte (assioma **9.3**), si possano ottenere tutti i risultati della sez. 9.1 come casi particolari.

9.1 LA SPIEGABILITÀ RIDUTTIVA FISICA DELLE ESPERIENZE COSCIENTI, L'IMPOSSIBILITÀ LOGICA DEI MONDI ZOMBIE E DI QUELLI ANGELICI

Per cominciare ad affrontare queste questioni, definiamo prima di tutto anche i concetti assoluti di esperienza cosciente e funzione cosciente:

[20] *f* è un fatto interno di coscienza (o *f* è un'esperienza cosciente) sse:
esiste *x* tale che *f* è un fatto interno di coscienza per *x*.

[21] *f* è un fatto esterno di coscienza (o *f* è una funzione cosciente) sse:
esiste *x* tale che *f* è un fatto esterno di coscienza per *x*.

Notiamo adesso che, in base alle def. [20] e [21], il teorema di relatività delle esperienze (teorema 2) si applica anche alle esperienze coscienti. In altre parole, vale il seguente teorema:

Teorema 3 [di identità reale fra esperienze coscienti e funzioni coscienti]
Per ogni *f*, *f* è un'esperienza cosciente sse *f* è una funzione cosciente.

Dimostrazione

La tesi segue dalle def. [20], [21], [14], [15], [8], [9] e dal teorema 2.
Q.E.D.

Questo teorema ci permette intanto di risolvere la questione se la distinzione fra esperienze coscienti e funzioni coscienti sia o meno reale (vedi sez. 2). La risposta è no: esperienze coscienti e funzioni coscienti sono due concetti diversi, ma essi tuttavia descrivono lo stesso insieme di fatti (cioè, i fatti di coscienza).

Conseguentemente, qualsiasi argomento che affermi di provare il contrario deve essere o invalido o fondarsi su qualche premessa falsa. Ciò, in particolare, vale per tutti gli argomenti classici citati sopra (vedi sez. 2).

Vediamo adesso quali sono le conseguenze di questo fatto per l'argomento degli zombie. Se l'argomento di Chalmers è dimostrativo (valido e con tutte le premesse vere), esso prova che le esperienze coscienti del nostro mondo non sono spiegabili riduttivamente dai fatti fisici del nostro mondo (tesi [Z], sez. 2). Assumendo l'ulteriore premessa che le funzioni coscienti del nostro mondo siano spiegabili riduttivamente dai fatti fisici del nostro mondo (assunzione [R], sez. 2), da [Z] segue anche che le esperienze coscienti del nostro mondo non sono spiegabili riduttivamente per mezzo delle funzioni coscienti del nostro mondo (corollario [Z₁], sez. 2). Ma questo contraddice il teorema 3. Perciò, per Chalmers rimangono soltanto due alternative:

Dilemma di Chalmers (I formulazione)

1. se l'argomento di Chalmers è dimostrativo, allora, contrariamente a quanto da lui affermato, le funzioni coscienti del nostro mondo non sono spiegabili riduttivamente dai fatti fisici del nostro mondo (cioè, l'assunzione $[R]$ è falsa);
2. se invece le funzioni coscienti del nostro mondo (e quindi, in forza del teorema 3, anche le esperienze coscienti del nostro mondo) sono spiegabili riduttivamente dai fatti fisici del nostro mondo, allora l'argomento degli zombie non è dimostrativo (cioè, o è invalido, o si fonda su qualche premessa falsa).

In effetti, il dilemma di Chalmers può essere messo in una forma ancora più cogente. Abbiamo infatti visto che l'argomento degli zombie si fonda soltanto su due premesse:

$[E_1]$ esiste w^{*z} tale che w^{*z} è una copia zombie del nostro mondo w^* ;

$[E_2]$ CF^* non è vuoto.

La tesi $[Z]$ segue quindi da $[E_1]$, $[E_2]$ e dalla definizione della relazione di conseguenza logica fra insiemi di fatti (def. [3]). Perciò, l'argomento degli zombie è certamente valido. Inoltre, la verità della premessa $[E_2]$ non è presumibilmente in discussione. Da ciò segue che il dilemma di Chalmers ammette una seconda formulazione, equivalente alla prima:

Dilemma di Chalmers (II formulazione)

1. se esiste una copia zombie del nostro mondo, allora, contrariamente a quanto affermato da Chalmers, le funzioni coscienti del nostro mondo non sono spiegabili riduttivamente dai fatti fisici del nostro mondo (cioè, l'assunzione $[R]$ è falsa);
2. se invece le funzioni coscienti del nostro mondo (e quindi, in forza del teorema 3, anche le esperienze coscienti del nostro mondo) sono spiegabili riduttivamente dai fatti fisici del nostro mondo, allora non esiste una copia zombie del nostro mondo.

Abbiamo visto (sez. 2) che Chalmers ammette la spiegabilità riduttiva delle funzioni coscienti dai fatti fisici del nostro mondo (assunzione $[Z]$). Ma allora, assumendo tale ipotesi come un ulteriore assioma della nostra ontologia, la II formulazione del dilemma di Chalmers ci fornisce una vera e propria *dimostrazione* della non esistenza di una copia zombie del nostro mondo. Vediamo i dettagli.

Trasformiamo l'assunzione [Z] in un nuovo assioma sull'ontologia (**D, P, Q, W, w*, λ, C**):

9. [*ipotesi fiscalista minimale nel nostro mondo*]

$$F^* \models CP^*.$$

Il teorema 3, in congiunzione con l'assioma **9**, ci permette subito di dimostrare la spiegabilità riduttiva fisica della coscienza *fenomenica* nel nostro mondo:

Teorema 4 [di spiegabilità riduttiva fisica della coscienza fenomenica nel nostro mondo]

$$F^* \models CF^*.$$

Dimostrazione

Per l'assioma **9**, $F^* \models CP^*$ e, per il teorema 3, $CP^* = CF^*$. Perciò, $F^* \models CF^*$.

Q.E.D.

Da ciò segue quindi:

Teorema 5 [di impossibilità logica di una copia zombie del nostro mondo]

Non esiste w^{*z} tale che w^{*z} è una copia zombie del nostro mondo w^* sse esiste f tale che $f \in CF^*$.

Dimostrazione

Dimostriamo prima da sinistra a destra e poi da destra a sinistra.

Tesi da sinistra a destra

Proviamo la contrapposta. Assumiamo quindi che non esiste f tale che $f \in CF^*$, e dimostriamo che esiste w^{*z} tale che w^{*z} è una copia zombie di w^* . Dall'ipotesi assunta e per la def. [5], w^* è una copia zombie di w^* . Ma allora la tesi vale per $w^{*z} = w^*$.

Tesi da destra a sinistra

Assumiamo che CF^* non sia vuoto. Supponiamo per assurdo che esista w^{*z} tale che w^{*z} sia una copia zombie di w^* . Per l'ipotesi di riduzione all'assurdo, in quanto CF^* non è vuoto, e per le def. [5] e [4] di copia zombie e conseguenza logica, $\neg(F^* \models CF^*)$. Ma, per il teorema 4, $F^* \models CF^*$. Dall'ipotesi di riduzione all'assurdo segue quindi una contraddizione.

Q.E.D.

Si noti che potremmo sostituire l'assioma **9** con un altro assioma relativo non al nostro mondo w^* , ma ad un altro mondo possibile w fissato; o, addirittura, potremmo aggiungere un corrispondente assioma per tutti quei mondi possibili per cui riteniamo vera l'ipotesi della spiegabilità riduttiva fisica della coscienza psicologica. Anche in tal caso, potremmo

ovviamente dimostrare gli analoghi dei teoremi 4 e 5 relativi a ciascuno di tali mondi.

Indichiamo con F_w l'insieme di tutti i fatti fisici che valgono in un mondo w fissato, con CP_w la coscienza psicologica di w (cioè, l'insieme di tutti i fatti esterni di coscienza che valgono in w) e con CF_w la coscienza fenomenica di w (ovvero, l'insieme di tutti i fatti interni di coscienza che valgono in w). La versione dell'assioma 9 per un qualunque mondo w fissato è quindi:

9.1 [*ipotesi fisicalista minimale locale*]

$$F_w \models CP_w.$$

Da cui otteniamo:

Teorema 4.1 [*di spiegabilità riduttiva fisica della coscienza fenomenica in un mondo possibile fissato*]

$$F_w \models CF_w.$$

Dimostrazione

Identica a quella del teorema 4, previa sostituzione di assioma 9 con assioma 9.1, F^* con F_w e CF^* con CF_w .

Q.E.D.

Analogamente, dimostriamo il corrispondente del teorema 5 per un mondo possibile fissato w .

Teorema 5.1 [*di impossibilità logica di una copia zombie di un mondo possibile fissato*]

Non esiste w^z tale che w^z è una copia zombie di w sse esiste f tale che $f \in CF_w$.

Dimostrazione

Identica a quella del teorema 5, previa sostituzione di w^* con w , w^{*z} con w^z , teorema 4 con teorema 4.1, F^* con F_w e CF^* con CF_w .

Q.E.D.

E' stato notato che l'idea di un mondo zombie è in pieno accordo con una visione cartesiana, secondo cui esistono due regni ontologicamente indipendenti: il pensiero e la materia. La copia zombie di un mondo w fissato può in effetti essere vista come la pura materia, separata dal pensiero, la cui esistenza, anche se non attuale, sarebbe pur sempre una possibilità logica. Ma abbiamo appena visto che, se assumiamo un'ipotesi fisicalista alquanto debole relativa a tale mondo, e cioè che i fatti della coscienza *psicologica* di w siano spiegabili riduttivamente in base alla sua fisica, questa possibilità deve essere esclusa.

Che cosa ne è dell'altro lato del dualismo cartesiano, e cioè del puro spirito? Così come abbiamo immaginato una copia zombie, possiamo immaginare una *copia angelica* di un mondo w fissato, cioè un mondo che sia del tutto identico a w dal punto di vista dei fatti interni di coscienza, ma in cui manchino del tutto i fatti fisici. Esiste fra i mondi possibili della nostra ontologia una copia angelica di w ? Sembra che, per una risposta negativa a questa domanda, l'ipotesi fisicalista minimale *locale* (assioma **9.1**) non sia sufficiente. Tuttavia, introducendo una versione *globale* (assioma **9.2**, sotto) di tale ipotesi, cioè generalizzata a tutti i mondi possibili, possiamo dimostrare che non esiste una copia angelica di w se in CF_w c'è almeno un fatto contingente, ovvero, che non vale in tutti i mondi possibili.

Definiamo:

[22] w^a è una copia angelica di w sse:

sia w che w^a sono mondi possibili e, per ogni f_1 , per ogni f_2 , se f_1 è un fatto fisico e f_2 è un fatto interno di coscienza, allora f_1 non vale in w^a e (f_2 vale in w^a sse f_2 vale in w).

Al posto dell'assioma **9.1** (o dell'assioma **9**) assumiamo la seguente condizione sull'ontologia (**D, P, Q, W, w***, λ , **C**):

9.2 [ipotesi fisicalista minimale globale]

per ogni mondo possibile w , $F_w \models CP_w$.

Sia w un mondo possibile fissato; dimostriamo:

Teorema 6.2 [di impossibilità logica di una copia angelica di un mondo possibile fissato]

Se esiste f tale che $f \in CF_w$ e f non vale in tutti i mondi possibili,²⁹ allora non esiste w^a tale che w^a è una copia angelica di w .

Dimostrazione

Assumiamo che esista $f_{\#}$ tale che $f_{\#} \in CF_w$ e $f_{\#}$ non valga in tutti i mondi possibili. Supponiamo per assurdo che esista w^a tale che w^a è una copia angelica di w . Siccome $f_{\#} \in CF_w$ e w^a è una copia angelica di w , $f_{\#} \in CF_w^a$; quindi, per il teorema 3, $f_{\#} \in CP_w^a$. Essendo w^a una copia angelica di w , F_w^a è vuoto e quindi, banalmente, per ogni w , per ogni $f \in F_w^a$, f vale in w ; da ciò, per la def. [4], segue che $F_w^a \models CP_w^a$ sse per ogni w , per ogni $f \in CP_w^a$, f vale in w . Ma $f_{\#} \in CP_w^a$ e $f_{\#}$ non vale in tutti i mondi possibili. Quindi, $\neg(F_w^a \models CP_w^a)$. Ma, per l'assioma **9.2**, $F_w^a \models CP_w^a$. Dall'ipotesi di riduzione all'assurdo segue quindi una contraddizione.
Q.E.D.

Lo scopo della prossima sezione è far vedere che, assumendo un ulteriore assioma che esprime in modo molto naturale un'ipotesi fisicalista forte,³⁰ si

può dimostrare l'ipotesi fiscalista minimale globale (assioma **9.2**); da essa seguono poi ovviamente tutti i risultati precedenti, e cioè i teoremi 4, 4.1, 5, 5.1 e 6.2.

9.2 L'IPOTESI FISCALISTA FORTE E IL TEOREMA DI RIDUCIBILITÀ FISICA

Il risultato menzionato sopra sarà in effetti ottenuto come conseguenza di un teorema più generale. Tale teorema afferma che un qualunque fatto ha un corrispondente fatto fisico che risulta logicamente equivalente ad esso (teorema 7.3, sotto). Questo teorema può quindi essere interpretato come un vero e proprio teorema di *riducibilità* di un fatto arbitrario al proprio equivalente fisico.

L'assioma che permette la dimostrazione del teorema 7.3 esprime l'idea che, nella nostra ontologia (**D, P, Q, W, w*, λ, C**) ciascuna proprietà o è una proprietà fisica, oppure è ottenibile da una proprietà fisica attraverso applicazioni ripetute delle operazioni di istanziazione e λ -astrazione. Questa assunzione (assioma **9.3**, sotto), secondo cui in ultima analisi tutte le proprietà sono proprietà fisiche, risulta un'ipotesi fiscalista forte. Infatti, dal teorema 7.3 segue ovviamente che qualunque fatto è spiegabile riduttivamente dall'insieme dei fatti fisici.

Al posto dell'assioma **9.2** (o dell'assioma **9.1**, o **9**) assumiamo la seguente condizione sull'ontologia (**D, P, Q, W, w*, λ, C**):

9.3 [*ipotesi fiscalista forte*]

per ogni proprietà P^k appartenente a **P**, o $P^k \in \mathbf{Q}$, o P^k è ottenuta a partire da una proprietà $P^n \in \mathbf{Q}$ attraverso un numero finito $m \geq 1$ di applicazioni successive delle due operazioni di istanziazione e λ -astrazione (in quest'ordine).

Diamo inoltre la seguente definizione:

[23] f_1 è logicamente equivalente a f_2 sse:

per ogni mondo possibile w , f_1 vale in w sse f_2 vale in w .

Data questa definizione, l'assioma **9.3** permette di dimostrare che ogni fatto è logicamente equivalente ad un opportuno fatto fisico. Vale infatti il seguente teorema:

Teorema 7.3 [*di riducibilità fisica*]

Per ogni fatto f , esiste un fatto fisico $f_{\#}$ tale che f è logicamente equivalente a $f_{\#}$.

Dimostrazione

Per l'assioma **9.3**, possiamo indicare qualunque proprietà $P^k \in \mathbf{P}$ con un simbolo del tipo $P^{n(p)}$, dove $p \geq 0$, $n(p)$ è l'arietà della proprietà P^k (quindi, $n(p) = k$) e l'indice p ha il seguente significato:

- se $p = 0$, $P^{n(p)} \in \mathbf{Q}$;
- se $p \geq 1$, la proprietà $P^{n(p)}$ è ottenuta a partire da una proprietà $P^n \in \mathbf{Q}$ attraverso p applicazioni successive delle operazioni di istanziazione e λ -astrazione (in quest'ordine).

Sia $P^k = P^{n(p)}$ il predicato di f ; dimostriamo il teorema per induzione su p .

Base dell'induzione

supponiamo:

1. $p = 0$;

dall'ipotesi di induzione 1 e per la definizione di $P^{n(p)}$:

2. $P^{n(0)} \in \mathbf{Q}$;

da 2 e per la def. [2]:

3. f è un fatto fisico;

quindi, da 3, scegliendo $f_{\#} = f$, la tesi vale.

Passo dell'induzione

supponiamo che:

1. la tesi valga per ogni fatto f il cui predicato sia rappresentabile come $P^{n(p)}$;

dimostriamo la tesi per ogni fatto il cui predicato sia rappresentabile come $P^{n(p+1)}$. Per la definizione di $P^{n(p+1)}$:

2. $P^{n(p+1)} = [\lambda(x_{i,1}, x_{i,2}, \dots, x_{i,n(p+1)})P^{n(p)}(x_1, x_2, \dots, x_{n(p)})]$;

da 2 e per l'assioma 5:

3. $P^{n(p+1)}(y_1, y_2, \dots, y_{n(p+1)})$ vale in w sse $P^{n(p)}[(x_1, x_2, \dots, x_{n(p)}) / (x_{i,1} / y_1, x_{i,2} / y_2, \dots, x_{i,n(p+1)} / y_{n(p+1)})]$ vale in w ;

da 3, dall'ipotesi di induzione 1 e dalla def. [23]:

4. esiste un fatto fisico $f_{\#}$ tale che $P^{n(p)}[(x_1, x_2, \dots, x_{n(p)}) / (x_{i,1} / y_1, x_{i,2} / y_2, \dots, x_{i,n(p+1)} / y_{n(p+1)})]$ vale in w sse $f_{\#}$ vale in w ;

da 4 e 3:

5. $P^{n(p+1)}(y_1, y_2, \dots, y_{n(p+1)})$ vale in w sse $f_{\#}$ vale in w ;

la tesi è dunque dimostrata per ogni fatto il cui predicato sia rappresentabile come $P^{n(p+1)}$.

Q.E.D.

Dal teorema 7.3 segue facilmente l'ipotesi fiscalista minimale globale (cioè, l'assioma **9.2**):

Corollario 1.7.3 [ipotesi fiscalista minimale globale]

Per ogni w , $F_w \models CP_w$.

Dimostrazione

Se CP_w è vuoto, la tesi segue immediatamente dalla def. [4] di conseguenza logica. Assumiamo quindi che f sia un arbitrario fatto

appartenente a CP_w . Allora, per il teorema 7.3 e la def. [23], esiste un fatto fisico $f_{\#}$ tale che, per ogni w , f vale in w sse $f_{\#}$ vale in w . Da ciò, in quanto $f \in CP_w$, e per la def. [4] di conseguenza logica, segue che $f_{\#} \in F_w$ e $f_{\#} \models f$. Perciò, $F_w \models f_{\#}$, e quindi, per transitività, $F_w \models f$.

Q.E.D.

Si noti che, dal corollario 1.7.3 segue ovviamente il teorema 6.2. Inoltre, segue anche l'ipotesi physicalista minimale locale relativa ad un mondo w fissato arbitrario (assioma **9.1**). Perciò, seguono poi anche i teoremi 4, 4.1, 5 e 5.1.

Notiamo infine che il teorema 7.3 ci permette anche di provare una versione leggermente più forte del teorema 6.2. In questa versione, infatti, non si richiede più che almeno un fatto interno di coscienza di w sia contingente.

Corollario 2.7.3 [teorema forte di impossibilità logica di una copia angelica di un mondo possibile fissato]

Se esiste f tale che $f \in CF_w$,³¹ allora non esiste w^a tale che w^a è una copia angelica di w .

Dimostrazione

Assumiamo che esista f tale che $f \in CF_w$, e supponiamo per assurdo che esista w ed esista w^a tale che w^a sia una copia angelica di w . Per l'ipotesi di riduzione all'assurdo e la def. [22] di copia angelica, f vale in w^a . Per il teorema 7.3, esiste un fatto fisico $f_{\#}$ tale che, per ogni w_1 , f vale in w_1 sse $f_{\#}$ vale in w_1 . Quindi, siccome f vale in w^a , $f_{\#}$ vale in w^a . Ma, essendo $f_{\#}$ un fatto fisico, per la def. [22], $f_{\#}$ non vale in w^a . Dall'ipotesi di riduzione all'assurdo segue quindi una contraddizione.

Q.E.D.

10 CONCLUSIONE: IL PROBLEMA DIFFICILE DELLA COSCIENZA NON È UN PROBLEMA SCIENTIFICO, MA FILOSOFICO

L'intuizione che guida Chalmers nel trattare il problema della coscienza è che il suo lato fenomenico o soggettivo andrà irrimediabilmente perduto una volta che la questione sia posta in una prospettiva di tipo physicalista. La fisica, in quanto scienza oggettivante per eccellenza, ci permetterebbe di cogliere soltanto gli aspetti esterni, o in terza persona, della coscienza, ma il suo aspetto più misterioso e affascinante, cioè il lato soggettivo, interno, o esperienziale, sarebbe assolutamente al di là della portata di qualsiasi teoria fisica.

La conclusione che possiamo trarre dalle analisi condotte fin qui è che questa intuizione è sostanzialmente priva di fondamento. Infatti, ho mostrato che: (i) è possibile sviluppare una teoria della coscienza fenomenica (sez. 5-8); (ii) la spiegazione della natura della coscienza fenomenica fornita da tale teoria non è affatto incompatibile con una prospettiva fisicalista ma, al contrario, una volta inserita in un quadro fisicalista minimale (espresso dall'assioma **9**, o **9.1**, o **9.2**, sez. 9.2), essa permette di provare la spiegabilità riduttiva fisica delle esperienze coscienti (teorema 4 o 4.1) e, conseguentemente, l'impossibilità logica dei mondi zombie (teorema 5 o 5.1).

Bisogna tuttavia riconoscere che tale conclusione non esaurisce tutti gli aspetti importanti del problema della coscienza fenomenica. In particolare, c'è una domanda alquanto sottile, ma assolutamente cruciale, che siamo in grado di porre con la necessaria chiarezza soltanto adesso, una volta giunti alla fine del nostro percorso.

Abbiamo visto che, secondo Chalmers, il problema difficile della coscienza, consiste nel *trovare e sviluppare una teoria adeguata della coscienza fenomenica*. Tuttavia, la domanda cruciale che ci dovremmo porre rispetto a tale problema è la seguente: la teoria che stiamo cercando di che tipo è? Più precisamente, stiamo cercando una teoria che ci permetta di dare una *spiegazione scientifica* dei fatti della coscienza fenomenica, oppure una teoria che ci dia una spiegazione filosofica, o un'*esplicazione*, del concetto di coscienza fenomenica?

Sebbene Chalmers non si sia mai posto questa domanda in modo esplicito, è comunque chiaro che egli propende per la prima alternativa. La mia convinzione, al contrario, è che il problema difficile della coscienza sia, in primo luogo, un *problema filosofico*, cioè un problema di chiarificazione e delimitazione concettuale, e che risolvere questo problema sia condizione necessaria per porre successivamente nei giusti termini anche il problema di una spiegazione scientifica dei fatti della coscienza fenomenica.

Ma allora, se le cose stanno così, possiamo anche capire come mai la questione della coscienza fenomenica appaia a Chalmers intrattabile dal punto di vista della fisica e, più in generale, di qualsiasi altra disciplina che ne condivida la stessa impostazione scientifica. Infatti, se il problema difficile della coscienza è un problema filosofico di *esplicazione concettuale*, e non un problema scientifico di *spiegazione di un tipo di fatti*, non è affatto sorprendente che la fisica, o qualsiasi altra scienza, risulti del tutto inadeguata a risolverlo. La teoria richiesta non è infatti una teoria scientifica, ma una teoria filosofica.

Questo lavoro ha tentato di fornire una prima versione di tale teoria filosofica. Se essa dovesse risultare adeguata, sarebbero tracciati i limiti entro i quali poi costruire le spiegazioni scientifiche di specifici fatti della coscienza.

Marco Giunti
 Università di Cagliari
 Dipartimento di Scienze Pedagogiche e Filosofiche
 Via Is Mirrionis 1
 09123 Cagliari

APPENDICE: CONSISTENZA DELLA TEORIA ASSIOMATICA

La prima clausola dell'assioma **8**, richiedendo che un qualsiasi fatto f appartenga al dominio, introduce una certa autoreferenzialità nella teoria assiomatica. Infatti, per la def. [1], f risulta elemento del soggetto di altri fatti, anch'essi appartenenti al dominio. Quindi, per ogni f , si pone la questione se f sia elemento del suo stesso soggetto. Si potrebbe allora sospettare che questa situazione permetta la costruzione di un'antinomia del tipo di quella di Russell.³² Perciò diventa importante dimostrare che la teoria costituita dagli assiomi **1 - 9.3** ammette un modello insiemistico, e risulta quindi consistente con la teoria degli insiemi.³³

I modelli della teoria assiomatica sono tutte le strutture **(D, P, Q, W, w*, λ, C)** che soddisfano gli assiomi **1 - 9.3**. La struttura definita sotto risulta un modello della teoria.

Sia $\mathbf{P} = \{P^1, C^2\}$, dove $P^1 = (\emptyset, 1)$ e $C^2 = (\emptyset, 2)$;

sia $\mathbf{Q} = \{C^2\}$;

sia $\mathbf{W} = \{w^*\}$, dove w^* è la funzione su \mathbf{P} definita da: $w^*(P^1) = \emptyset$ e $w^*(C^2) = \emptyset$;

sia $\mathbf{D} =$ l'unione di tutti i domini n -esimi Δ_n , dove, per ogni $n \geq 0$, il dominio Δ_n è definito induttivamente come segue:

$\Delta_0 = \{\emptyset\}$;

supposto di avere definito Δ_n , definiamo Δ_{n+1} :

$\Phi_n = \{f: f \text{ è una coppia ordinata del tipo } (P^n, (x_1, x_2, \dots, x_n)), \text{ dove } P^n \in \mathbf{P} \text{ e } (x_1, x_2, \dots, x_n) \text{ è una } n\text{-upla } (n \geq 1) \text{ di elementi di } \Delta_n\}$; conveniamo di scrivere $P^n(x_1, x_2, \dots, x_n)$ al posto di $(P^n, (x_1, x_2, \dots, x_n))$;

$\Delta_{n+1} =$ l'unione di Δ_n e Φ_n .

Notiamo subito che, in forza delle definizioni di $\mathbf{D}, \mathbf{P}, \mathbf{Q}, \mathbf{W}$ e w^* gli assiomi **1, 2, 3, 4 e 5** sono ovviamente soddisfatti.

Ricordiamo che, dal punto di vista insiemistico, l'operatore λ può essere pensato come una funzione parziale a due posti, il cui codominio è \mathbf{P} , il cui primo dominio è l'unione di tutti i prodotti cartesiani k -esimi di \mathbf{D} (per $1 \leq k$), ed il cui secondo dominio è l'insieme di tutti i fatti possibili della struttura (vedi def. [1]). Definiamo tale funzione come segue:

se $P^n(x_1, x_2, \dots, x_n)$ è un qualunque fatto della struttura definita sopra (quindi, $1 \leq n \leq 2$), e $(x_{i,1}, x_{i,2}, \dots, x_{i,k})$ è una k -upla di esattamente k elementi diversi, ciascuno dei quali occorre nella n -upla (x_1, x_2, \dots, x_n) , allora:

se $k = 1$, $[\lambda(x_{i,1}, x_{i,2}, \dots, x_{i,k})P^n(x_1, x_2, \dots, x_n)] = P^1$;
 se $k = 2$, $[\lambda(x_{i,1}, x_{i,2}, \dots, x_{i,k})P^n(x_1, x_2, \dots, x_n)] = C^2$.

Per la definizione di λ , esso soddisfa ovviamente l'assioma **6**. Inoltre, siccome nella struttura che stiamo considerando c'è un solo mondo possibile (cioè \mathbf{w}^*) e, per la definizione di \mathbf{w}^* , nessun fatto vale in \mathbf{w}^* , anche l'assioma **7** è banalmente soddisfatto.

Per quanto riguarda la proprietà binaria \mathbf{C} , stabiliamo $\mathbf{C} = C^2$. In base a questa definizione, tutte le clausole dell'assioma **8** valgono banalmente, eccetto la prima. Dimostriamo quindi che:

Per ogni fatto f , $f \in \mathbf{D}$.

Dimostrazione

Sia f un fatto arbitrario. Allora, f è una coppia ordinata del tipo $P^m(x_1, x_2, \dots, x_m)$, dove $P^m \in \mathbf{P}$ e (x_1, x_2, \dots, x_m) è una m -upla ($m \geq 1$) di elementi di \mathbf{D} . Essendo \mathbf{D} l'unione di tutti i domini n -simi Δ_n , per ogni elemento x_i di (x_1, x_2, \dots, x_m) , esiste $n(i)$ tale che $x_i \in \Delta_{n(i)}$. Sia $\text{MAX} = \max[n(1), n(2), \dots, n(m)]$. Allora, per la definizione di dominio n -esimo, per ogni i , $\Delta_{n(i)}$ è incluso in Δ_{MAX} . Quindi, $f \in \Phi_{\text{MAX}}$. Ma, per la definizione di \mathbf{D} , per ogni n , Φ_n è incluso in \mathbf{D} . Ne segue che $f \in \mathbf{D}$.

Q.E.D.

Infine, anche l'assioma **9.3**, risulta soddisfatto. Infatti, essendo $\mathbf{Q} = \{C^2\}$, dobbiamo soltanto verificare che P^1 è ottenibile da C^2 mediante un numero finito di applicazioni successive delle due operazioni di istanziazione e λ -astrazione. Ma ciò è ovvio: istanziamo dapprima C^2 con (\emptyset, \emptyset) , ottenendo il fatto $C^2(\emptyset, \emptyset)$. Applichiamo quindi la λ -astrazione a \emptyset e $C^2(\emptyset, \emptyset)$; per la definizione di λ , otteniamo $[\lambda(\emptyset)C^2(\emptyset, \emptyset)] = P^1$.

RIFERIMENTI BIBLIOGRAFICI

Armstrong, David M. 1968. *A materialist theory of mind*. London:

- Routledge and Kegan Paul.
- 1981a. *The nature of mind and other essays*. Ithaca, NY: Cornell University Press.
- 1981b. "The causal theory of the mind." In Armstrong (1981a). [Ristampato in Lycan a cura di (1990), pp. 37-47.]
- Block, Ned. 1978. "Troubles with functionalism." In *Perception and cognition: Issues in the foundations of psychology*, a cura di C. W. Savage. Minneapolis: University of Minnesota Press, pp. 261-325. [Ristampato senza la sezione su Lewis in Block a cura di (1980), vol. 1, pp. 268-305. La stessa versione, col titolo, "An excerpt from *Troubles with functionalism*" si trova anche in Lycan a cura di (1990), pp. 444-468.]
- Block, Ned, a cura di. 1980. *Readings in the philosophy of psychology*. 2 voll. Cambridge, MA: Harvard University Press.
- Block, Ned e Robert Stalnaker. 1999. Conceptual analysis, dualism and the explanatory gap. *The Philosophical Review* 108:1-46.
- Carnap, Rudolf. 1950. *Logical foundations of probability*. Chicago: The University of Chicago Press.
- Chalmers, David J. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2:200-220. [Ristampato in Hameroff, Kaszniak e Scott a cura di (1996), pp. 5-28.]
- 1996. *The conscious mind: In search of a fundamental theory*. New York: Oxford University Press. [Chalmers, David J. 1999. *La mente cosciente*. Trad. it. di Alfredo Paternoster e Cristina Meini. Introduzione di Michele di Francesco. Milano: Mc-Graw Hill.]
- 1997. Forward on the problem of consciousness. *Journal of Consciousness Studies* 4:3-47.
- 2002. "Does conceivability entail possibility?" In *Conceivability and possibility*, a cura di T. Gendler e J. Hawthorne. New York: Oxford University Press, pp. 145-200.
- Chalmers, David J. e Frank Jackson. 2001. Conceptual analysis and reductive explanation. *Philosophical Review* 110:315-361.
- Cottrell, Allin. 1999. Sniffing the camambert: On the conceivability of zombies. *Journal of Consciousness Studies* 6:4-12.
- Hameroff, Stuart R., Alfred W. Kaszniak e Alwyn C. Scott, a cura di. 1996. *Toward a science of consciousness: The first Tucson discussions and debates*. Cambridge, MA: The MIT Press.
- 1998. *Toward a science of consciousness II: The second Tucson*

- discussions and debates*. Cambridge, MA: The MIT Press.
- Hameroff, Stuart R., Alfred W. Kaszniak e David J. Chalmers, a cura di. 1999. *Toward a science of consciousness III: The third Tucson discussions and debates*. Cambridge, MA: The MIT Press.
- Hempel, Carl G. 1952. *Fundamentals of concept formation in empirical science*. Chicago: The University of Chicago Press. [Hempel, Carl G. 1961. *La formazione dei concetti e delle teorie nelle scienze empiriche*. Trad. it. di Alberto Pasquinelli. Milano: Feltrinelli.]
- Dennett, Daniel. 1991. *Consciousness explained*. Boston: Little Brown.
- Descartes, René. 1637. *Discours de la méthode*. [Descartes, René. 1977. *Discorso sul metodo*. Trad. it. di Enza Carrara. Introduzione e note a cura di Étienne Gilson. XXIII edizione. Firenze: La Nuova Italia.]
- Jackson, Frank. 1982. Epiphenomenal qualia. *Philosophical Quarterly* 32:127-136. [Ristampato in Lycan a cura di (1990), pp. 469-477.]
- Kirk, Robert. 1974. Zombies versus materialists. *Aristotelian Society* 48(suppl.):135-152.
- Lewis, David K. 1966. An argument for the identity theory. *Journal of Philosophy* LXIII, 1:17-25.
 ——— 1972. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy* 50:249-258. [Ristampato in Block a cura di (1980), vol. 1, pp. 232-233.]
- Locke, John. 1690. *An essay concerning human understanding*. [Locke, John. 1994. *Saggio sull'intelligenza umana*. Trad. it. di C. Pellizzi, rivista da G. Farina. Roma e Bari: Laterza.]
- Lycan, William. G. , a cura di. 1990. *Mind and Cognition: A reader*. Oxford, UK: Blackwell.
- Mathieson, Chris. 2000. Reining in Chalmers: On the logical possibility of zombies. *Canadian Undergraduate Journal of Philosophy* 1:10-18.
- Moody, Todd C. 1994. Conversations with zombies. *Journal of Consciousness Studies* 1:196-200.
- Nagel, Thomas. 1974. What is it like to be a bat? *Philosophical review* 4:435-450. [Trad. it. "Che effetto fa essere un pipistrello?", in Nagel (1986), pp. 162-176.]
 ——— 1986. *Questioni mortali*. Milano: Il Saggiatore.

- Place, Ullin T. 1956. Is consciousness a brain process? *British Journal of Psychology* 47:44-50.
- Putnam, Hilary. 1960. "Minds and machines." In *Dimensions of mind*, a cura di Sidney Hook. New York: New York University Press, pp. 148-179. [Trad. it. "Menti e macchine", in Putnam (1987), pp. 392-416].
- 1964. Robots: machines or artificially created life? *Journal of Philosophy* LXI, 21:688-691. [Trad. it. "I robot: macchine o vita creata artificialmente?", in Putnam (1987), pp. 416-439.]
- 1965. "Brains and Behavior." In *Analytical philosophy, Second Series*, a cura di R. J. Butler. Oxford: Basil Blackwell, pp. 211-235. [Trad. it. "Cervelli e comportamento", in Putnam (1987), pp. 354-372.]
- 1967a. "The mental life of some machines." In *Intentionality, minds, and perception*, a cura di Hector-Neri Castañeda. Detroit: Wayne State University Press, pp. 177-200.
- 1967b. "Psychological Predicates." In *Art, mind, and religion*, a cura di W. H. Capitan e D. D. Merrill. Pittsburgh: University of Pittsburgh Press, pp. 37-48. [Ristampato col titolo "The nature of mental states", in Rosenthal a cura di (1987), pp. 150-161. Con lo stesso titolo anche in Lycan a cura di (1990), pp. 47-56].
- 1987. *Mente linguaggio e realtà*. Milano: Adelphi.
- Rogers, Robert. 1971. *Mathematical logic and formalized theories*. Amsterdam: North Holland Publishing Company.
- Rosenthal, David M., 1987. *Materialism and the mind-body problem*. Indianapolis: Hackett Publishing Company, Inc.
- Ryle, Gilbert. 1949. *The Concept of mind*. London: Hutchinson.
- Seager, William. 2002. "Are zombies logically possible?" Presentato al *CPA Meeting*, Toronto, maggio 2002.
<http://www.scar.utoronto.ca/~seager/zombie.html>
- Smart, John Jamieson C. 1959. Sensations and brain processes. *Philosophical Review* 68:141-156. [Ristampato in Rosenthal a cura di (1987), pp. 53-66].
- Suppes, Patrick. 1957. *Introduction to logic*. New York: D. Van Nostrand Company.

NOTE

¹ Si tenga presente che, per il funzionalismo, il carattere interno degli stati mentali si riduce in ultima analisi al loro status di entità teoriche, quindi non immediatamente osservabili, le quali risultano implicitamente definite dal ruolo occupato nella rete dei rapporti causali che esse permettono di descrivere.

² Essi sono: l'argomento del *What is it like to be?* (Nagel 1974); l'argomento dei *qualia assenti* (Block 1978); l'argomento dello *spettro invertito* (Block 1978; già presente in Locke 1690); l'argomento della *conoscenza* o di *Mary* (Jackson 1982) e, infine, l'argomento degli *zombie fenomenici* (Chalmers 1996; già presente in forme diverse in Kirk 1974 e Dennett 1991. L'articolo di Moody 1994 ha dato luogo ad un forum di discussione, "Conversations with Zombies", costituito da tredici articoli, in *Journal of Consciousness Studies*, vol. 2, n. 4, 1995).

³ Fino ad oggi si sono tenuti cinque convegni, denominati *Biennial Tucson Consciousness Conferences*, dal 1994 al 2002. MIT Press ha pubblicato gli atti dei primi tre convegni; i primi due a cura di Hameroff, Kaszniak e Scott (1996, 1998) ed il terzo a cura di Hameroff, Kaszniak e Chalmers (1999).

⁴ Si sono tenuti fino ad oggi sette congressi, denominati *Annual ASSC Conferences*, dal 1997 al 2003. L'ultimo si è svolto a Memphis dal 30 maggio al 2 giugno del 2003.

⁵ Edita nel Regno Unito da Keith Sutherland, è forse la più importante fra le riviste espressamente dedicate agli studi sulla coscienza. Fondata nel 1994 e coinvolta nell'organizzazione delle *Biennial Tucson Consciousness Conferences*, ha ospitato gran parte del dibattito sul cosiddetto *problema difficile della coscienza (hard problem)*. Tale dibattito, raccolto in cinque forum di discussione, si è aperto e chiuso con due articoli di Chalmers (1995, 1997). Il primo articolo era stato da lui presentato al primo convegno di Tucson del 1994.

⁶ Fondata nel 1995, è una delle due riviste ufficiali della ASSC. *Psyche* è esclusivamente elettronica e completamente gratuita, e si trova all'indirizzo: <http://psyche.cs.monash.edu.au> .

⁷ Fondata nel 1992, è la seconda rivista ufficiale della ASSC.

⁸ Fondata nel 2000, ospita principalmente studi sulle emozioni.

⁹ Il termine *corrisponde* è volutamente ambiguo in quanto, a questo livello così iniziale dell'analisi, la precisa natura di tale corrispondenza non può essere presupposta. In particolare, quindi, non è esclusa la possibilità che fatti interni ed esterni della coscienza siano identici, e che quindi tale distinzione non sia reale ma soltanto concettuale, cioè relativa al modo di descrivere, apprendere o conoscere tali fatti. Altre due possibilità da considerare sono una relazione di conseguenza logica fra i due tipi di fatti (o di sopravvenienza logica fra le proprietà corrispondenti), o una relazione più debole che comunque sia descrivibile come una qualche forma di correlazione o congiunzione costante. In questi ultimi due casi la distinzione fra fatti interni ed esterni della coscienza sarebbe comunque reale.

¹⁰ La questione della corrispondenza fra fatti interni ed esterni della coscienza non deve essere confusa con quella del loro carattere intenzionale. Supposto che fra i fatti interni di coscienza possiamo annoverare l'esperienza cosciente di Maria che è in preda all'ansia, non è chiaro se esista un corrispondente fatto esterno che abbia più o meno lo stesso

ruolo che ha la martellata sul pollice rispetto all'esperienza cosciente di Maria che prova il dolore di tale colpo. L'asimmetria fra questi due tipi di fatti interni di coscienza viene di solito descritta come la mancanza di un contenuto determinato nell'esperienza di uno stato ansioso, e può essere considerata come una buona ragione per sostenere la non intenzionalità di tali esperienze. Comunque vogliamo risolvere questa questione, rimane tuttavia il fatto che all'esperienza cosciente di Maria che è in preda all'ansia si accompagna un particolare stato di attivazione del suo sistema nervoso centrale e/o periferico. Ed è proprio questo stato di attivazione, descrivibile esternamente o in terza persona, che costituisce il fatto esterno, o la funzione cosciente, corrispondente all'esperienza dell'ansia vissuta internamente da Maria.

¹¹ Questa affermazione di una corrispondenza biunivoca fra fatti interni ed esterni della coscienza non è ovviamente da prendere in senso categorico. Infatti, essendo stata ottenuta per generalizzazione a partire da una serie di osservazioni assolutamente non sistematiche, la sua verità è al più limitata al nostro mondo; inoltre, alla luce di più approfondite considerazioni di tipo empirico o teorico, essa potrebbe addirittura risultare falsa anche in tale mondo. Si noti quindi che sottoscrivere questa affermazione non pregiudica in alcun modo la questione della possibilità logica di un mondo zombie (vedi sez. 4).

¹² Per *antiriduzionismo forte* intendo qui la posizione di Chalmers, secondo cui i fatti interni della coscienza, o le esperienze fenomeniche, non sono riduttivamente spiegabili in termini di fatti fisici. Altri modi equivalenti per dire la stessa cosa sono: le proprietà fenomeniche non sopravvivono logicamente su quelle fisiche; i fatti interni della coscienza (o le esperienze coscienti) non seguono logicamente dalla totalità dei fatti fisici.

¹³ Mi riferisco qui a quello che ho chiamato *problema superdifficile della coscienza*, cioè il problema di costruire una teoria della coscienza fenomenica che soddisfi i requisiti 1-5. E' questo tipo di problema che Chalmers (1996) tenta di risolvere. Il problema superdifficile della coscienza è un *tipo particolare di problema difficile della coscienza* in quanto tale problema più generale può essere definito come la ricerca di una teoria adeguata della coscienza fenomenica (dove le condizioni di adeguatezza non necessariamente sono specificate dai requisiti 1-5).

¹⁴ Per il momento, trattiamo l'insieme dei fatti interni di coscienza **CF** e quello dei fatti esterni di coscienza **CP** come due ulteriori entità primitive della teoria, con il relativo assioma che tutti gli elementi di questi due insiemi siano fatti. Su questa base, definiamo quindi CF^* (o CP^*) come l'insieme dei fatti interni (esterni) di coscienza che valgono in w^* . Vedremo tuttavia che sia **CF** che **CP** risultano definibili all'interno di un'opportuna estensione della teoria assiomatica (def. [20] e [21], sez. 9.1).

¹⁵ Ciò, secondo Chalmers, è equivalente ad affermare che i fatti della coscienza psicologica seguono logicamente dai fatti fisici o, il che è lo stesso, che le proprietà della coscienza psicologica sono logicamente sopravvenienti sulle proprietà fisiche.

¹⁶ Se CF^* non è vuoto, esiste un mondo possibile, cioè il mondo zombie, in cui valgono tutti i fatti fisici del nostro mondo ma in cui nessun fatto appartenente a CF^* vale. Perciò, dalla def. [4], $\neg(F^* \models CF^*)$. Viceversa, se CF^* è vuoto, allora il conseguente della def. [4] è banalmente vero e quindi $F^* \models CF^*$.

¹⁷ Dubbi a tal riguardo sono per esempio stati espressi da Cottrell (1999). La questione della concepibilità dei mondi zombie è connessa a quella del rapporto fra concepibilità, non contraddittorietà e possibilità logica. Su questo punto si vedano Chalmers (2002) e

Mathieson (2000). Anche Block e Stalnaker (1999) e Chalmers e Jackson (2001) affrontano questa problematica nell'ambito della questione più ampia che riguarda le condizioni di una spiegazione riduttiva della coscienza fenomenica.

¹⁸ La questione dei rapporti fra concepibilità, non contraddittorietà e possibilità logica è alquanto controversa, ma essa è resa ancora più confusa dal persistere di una serie di veri e propri fraintendimenti ed errori di tipo logico. E' quindi opportuno, prima di tutto, sgombrare il campo da tali errori. Intanto, è bene precisare che, in questo contesto, si sta parlando di *concetti*, e non di *enunciati* o asserzioni. Chiarito ciò, si dà spesso per scontato che dalla concepibilità di un concetto segua la sua non contraddittorietà, e da questa la possibilità logica di un oggetto che soddisfi tale concetto. La prima implicazione ha sicuramente un fondamento logico certo. Se un concetto C è definibile in una qualche teoria non contraddittoria T , allora C è sicuramente concepibile. Inoltre, per il principio di non creatività delle definizioni, anche la teoria T^* , costituita dalla congiunzione di T con la definizione di C , è non contraddittoria; in tal caso si dice quindi che C è *relativamente non contraddittorio* (data, cioè, la non contraddittorietà di T). Possiamo perciò affermare che, se un concetto C è concepibile nel senso che è definibile in una teoria non contraddittoria T , allora esso è relativamente non contraddittorio rispetto a quella teoria. Ma ciò implica anche che un oggetto che soddisfa C è logicamente possibile? Per rispondere a questa domanda bisogna prima di tutto chiarire che cosa si intenda per possibilità logica di un oggetto che soddisfa C . Anche a questa domanda c'è una risposta standard, cioè in linea con l'uso logico-matematico della nozione di possibilità logica. Secondo questa risposta, la possibilità logica di un oggetto che soddisfa C si identifica con la non contraddittorietà relativa dell'*enunciato* che afferma l'esistenza di un tale oggetto. In altre parole, un oggetto del tipo C è logicamente possibile se, e solo se, la teoria $T^{\#}$, costituita dalla congiunzione di T^* con l'enunciato $E_C = \text{"esiste } x \text{ tale che } x \text{ è } C\text{"}$ è non contraddittoria. Ma allora è evidente che la non contraddittorietà relativa del concetto C non implica affatto la possibilità logica di un oggetto di tipo C . Infatti, affermare che C è relativamente non contraddittorio rispetto a T significa soltanto affermare che T^* è non contraddittoria. Ma ciò non implica la non contraddittorietà di $T^{\#}$, che è un'estensione di T^* . In effetti, T^* potrebbe benissimo implicare la negazione di E_C ; in tal caso, $T^{\#}$ sarebbe ovviamente contraddittoria, e quindi un oggetto di tipo C risulterebbe logicamente impossibile.

¹⁹ Se l'asserzione dell'esistenza del mondo zombie non è consistente con i principi assunti (cioè, per la definizione della nota 18, se il mondo zombie è logicamente impossibile), allora essi escludono l'esistenza del mondo zombie. Se invece i principi assunti sono consistenti con l'asserzione dell'esistenza del mondo zombie (cioè, per la definizione della nota 18, se il mondo zombie è logicamente possibile), allora tale asserzione può essere aggiunta ai principi, e l'esistenza, nel senso logico-matematico, del mondo zombie è così garantita.

²⁰ E' stato notato (Seager 2002) che, a meno di incorrere in una petizione di principio, fra tali assunzioni ausiliari non può certamente essere inclusa un'ipotesi fiscalista forte, cioè un'ipotesi da cui segua che *qualsunque* fatto (del nostro mondo) è spiegabile riduttivamente dall'insieme dei fatti fisici (del nostro mondo). Tale ipotesi ha come ovvia conseguenza la spiegabilità riduttiva fisica della coscienza fenomenica, e quindi essa è inconsistente con l'asserzione dell'esistenza di una copia zombie del nostro mondo. Tuttavia, una "prova" di questo genere non potrebbe essere accettata da Chalmers, perché, dal suo punto di vista, essa non sarebbe nient'altro che una riduzione all'assurdo dell'ipotesi fiscalista forte. L'unico tipo di ipotesi fiscalista che può essere inclusa fra le ipotesi ausiliari è una che lo stesso Chalmers non avrebbe problemi ad accettare, e cioè l'ipotesi fiscalista minimale [R], che afferma *soltanto* la spiegabilità riduttiva fisica della coscienza psicologica.

²¹ È interessante notare che il problema che ci stiamo ponendo è un tipico problema di esplicazione concettuale, in cui si richiede una definizione di un concetto per genere (in questo caso: le esperienze per un soggetto x arbitrario) e differenza specifica (in questo caso: ciò che rende tali esperienze coscienti). Tradizionalmente, tali definizioni sono dette *definizioni reali*. Per la teoria della definizione in generale si veda Suppes (1957, cap. 8), Rogers (1971, sez. 4.4) e Hempel (1952, cap. 1); per le definizioni reali ed il loro rapporto con l'esplicazione si veda Hempel (1952, cap. 1, par. 3); per l'esplicazione, si veda Carnap (1950, cap. 1). Tutte le volte che un problema di esplicazione si pone nella forma della richiesta di una definizione reale, esso si riduce a due sottoproblemi più semplici: (1) definire il genere; (2) definire la differenza specifica. Questo primo momento della risoluzione di un problema di definizione reale corrisponde alla seconda regola del metodo di Descartes (regola della divisione), secondo cui bisogna "dividere ciascuna delle difficoltà da esaminare in tutte le parti in cui fosse possibile e di cui ci fosse bisogno per meglio risolverle" (Descartes 1637). Una volta che si sia riusciti a risolvere i sottoproblemi (1) e (2), anche il problema di partenza è risolto, perché la definizione cercata non è nient'altro che la congiunzione della definizione del genere e di quella della differenza specifica. Si noti infine che questo secondo momento della risoluzione di un problema di definizione reale corrisponde alla terza regola del metodo di Descartes (regola dell'ordine) secondo cui, quando una difficoltà sia scomposta in parti semplici, si deve risalire da esse alla soluzione del problema più complesso. A questo proposito, si veda anche la nota 25.

²² L'ultima clausola della definizione assicura che un fatto può risultare esterno soltanto rispetto ad entità (le n -uple di elementi del dominio) che possono essere soggetto di altri fatti. Si evitano così sia casi irrilevanti, che intuitivamente non persuasivi.

²³ L'affermazione che *tutte le cose hanno esperienze* non ha niente a che vedere con una posizione panpsichista. La mia posizione è che avere un'esperienza di un fatto si identifica con una particolare relazione *ontologica* che sussiste esclusivamente fra il fatto che vale, f , ed un sistema, s , in una posizione metafisica privilegiata rispetto ad esso. Siccome il fatto f è costituito da una proprietà P e da un *unico* soggetto a cui la proprietà P inerisce, la relazione in questione non è nient'altro che l'essere il soggetto del fatto f . Il sussistere o meno di questa relazione fra f ed s è sufficiente a spiegare la dualità *interno - esterno*, *prima persona - terza persona*, *soggettivo - oggettivo*; ma non è affatto sufficiente per attribuire una qualche forma di vita psichica al sistema s . Il carattere distintivo della vita psichica, infatti, non è la soggettività ma, piuttosto, la presenza o meno della *coscienza*, ovvero, il valere o meno di un particolare tipo di fatti, che possiamo chiamare *fatti di coscienza*. Conseguentemente, un sistema s per cui non valga alcun fatto di coscienza di cui s è soggetto ha soltanto esperienze non coscienti (o, se vogliamo, *inconsce*) e non è quindi un soggetto psicologico (s è soltanto un soggetto in senso metafisico o ontologico). Perché s sia un soggetto psicologico occorre invece che s sia soggetto di un tipo appropriato di fatti, ovvero, di fatti di coscienza, e che almeno alcuni di tali fatti *valgano*.

²⁴ Questa condizione introduce una certa autoreferenzialità nella teoria assiomatica, in quanto i fatti, appartenendo al dominio, risultano elementi del soggetto di altri fatti, anch'essi appartenenti al dominio. Si potrebbe sospettare che questa situazione permetta la costruzione di un'antinomia del tipo di quella di Russell. Ma non è così, perché si dimostra che la teoria costituita dagli assiomi **1 - 9.3** ammette un modello insiemistico, ed è quindi consistente con la teoria degli insiemi. Per i dettagli, si veda l'appendice.

²⁵ Si noti che la def. [14], della relazione f è un'esperienza cosciente per x , è una definizione per genere e differenza specifica, in cui il genere è dato dalla def. [6] e la differenza specifica dalla def. [11]. La def. [14] risolve il problema di esplicazione del

concetto di esperienza cosciente per un soggetto x arbitrario. A questo proposito, si veda anche la nota 21.

²⁶ Per esempio, se Maria fosse sotto anestesia, il suo essere nello stato neurofisiologico d non causerebbe il suo esser cosciente di trovarsi in tale stato (ovvero, la coscienza del dolore non sarebbe presente).

²⁷ Secondo la mia proposta, quindi, la descrizione determinata $D = \text{"l'esperienza cosciente di Maria che prova il dolore di una martellata sul pollice"}$ denota esattamente il fatto f . Questa identificazione è giustificata dall'analisi condotta, dal significato intuitivo della descrizione D , e dalla def. [16], secondo cui:

[D*] f è un'esperienza autocosciente per Maria, il cui contenuto è il suo essere nello stato d , causato dalla martellata sul pollice (cioè, il contenuto dell'esperienza cosciente di Maria f è la sua esperienza f_d , causata dalla martellata sul pollice);

confrontando il significato intuitivo di D con le caratteristiche di f asserite da [D*], risulta in effetti del tutto naturale identificare l'oggetto descritto da D con f .

²⁸ Una copia angelica di un mondo w può essere pensata come il concetto duale di una copia zombie. Infatti, una copia angelica w^a di un mondo possibile w è un mondo in cui valgono tutti i fatti interni di coscienza che valgono in w , ma in cui non vale alcun fatto fisico.

²⁹ La condizione che esista f tale che $f \in CF_w$ e f non valga in tutti i mondi possibili non è molto limitativa, perché le uniche copie angeliche che il teorema 6.2 non esclude o sono quelle, banali, di un mondo in cui non vale alcun fatto interno di coscienza, o quelle di mondi in cui nessun fatto interno di coscienza è contingente. In particolare, una copia angelica del nostro mondo w^* è esclusa dal teorema 6.2, perché, presumibilmente, almeno alcune esperienze coscienti di w^* sono contingenti.

³⁰ Un'ipotesi fiscalista forte è un principio da cui segua che *qualunque* fatto è riduttivamente spiegabile dall'insieme dei fatti fisici.

³¹ La condizione che esista f tale che $f \in CF_w$ non è realmente limitativa, perché le uniche copie angeliche che il corollario 2.7.3 non esclude sono quelle, banali, di un mondo in cui non esiste alcun fatto interno di coscienza. In effetti, potremmo addirittura aggiungere tale condizione alla definizione stessa di copia angelica, eliminandola così dal corollario 2.7.3. Tuttavia, ciò non sembra opportuno, perché in tal modo verrebbe meno la perfetta simmetria fra la def. [22] di copia angelica e la def. [5] di copia zombie. Si potrebbe restaurare la simmetria aggiungendo alla def. [5] la condizione che F_w non sia vuoto. Ma allora nel teorema 5.1 non varrebbe più il bicondizionale, ma soltanto l'implicazione da destra a sinistra.

³² Ringrazio Francesco Paoli per aver sollecitato la mia attenzione ad obiezioni di questo tipo.

³³ Per essere precisi, gli assiomi 1 - 9.3 definiscono un tipo di struttura matematica \mathcal{O} , seguendo la terminologia di Suppes (1957), *un predicato insiemistico (a set theoretical predicate)*. Chiamiamo tale predicato $C = \text{essere un'ontologia della coscienza}$, e definiamo quindi: una struttura matematica S è un'ontologia della coscienza sse esistono $\mathbf{D}, \mathbf{P}, \mathbf{Q}, \mathbf{W}, \mathbf{w}^*, \lambda, \mathbf{C}$, tali che $S = (\mathbf{D}, \mathbf{P}, \mathbf{Q}, \mathbf{W}, \mathbf{w}^*, \lambda, \mathbf{C})$ e gli assiomi 1 - 9.3 sono soddisfatti. In generale diciamo che una teoria assiomatizzata T è *consistente* sse esiste un modello della teoria, ovvero, sse esiste una struttura matematica che soddisfa il

predicato insiemistico C_T da essa definito. È interessante notare che questa definizione di consistenza per le teorie assiomatizzate è equivalente a quella di non contraddittorietà relativa alla teoria degli insiemi dell'enunciato $E_{C_T} = \text{"esiste } S \text{ tale che } S \text{ è } C_T\text{"}$ (si veda la nota 18). Infatti, secondo tale definizione, E_{C_T} è *non contraddittorio relativamente alla teoria degli insiemi* sse la congiunzione della teoria degli insiemi con E_{C_T} è non contraddittoria. Inoltre, se E_{C_T} è non contraddittorio relativamente alla teoria degli insiemi, possiamo aggiungere E_{C_T} a tale teoria e renderlo quindi un suo teorema. D'altra parte, per la definizione di consistenza data sopra, se E_{C_T} è dimostrabile all'interno della teoria degli insiemi, allora T è consistente. Conseguentemente, se E_{C_T} è non contraddittorio relativamente alla teoria degli insiemi, allora T è consistente. Ma, in effetti, vale anche l'inverso. Assumiamo che T sia consistente. Allora, per la definizione di consistenza data sopra, esiste S tale che S è C_T . Conseguentemente, E_{C_T} è vera. Ma, siccome tutti i teoremi della teoria degli insiemi sono veri, la congiunzione di tale teoria con E_{C_T} (che è anch'essa vera) è non contraddittoria. Ne segue quindi che, se T è consistente, allora E_{C_T} è non contraddittorio relativamente alla teoria degli insiemi.