

L'emergenza fenomenica*

Achim Stephan

achim.stephan@uos.de
Institute of Cognitive Science University of Osnabrück (Germany)

1 Introduzione

Qualche anno fa Dave Chalmers iniziava un libro sulla coscienza con le seguenti parole:

La coscienza è il più grande dei misteri. Può essere il più grande ed eccezionale ostacolo nella ricerca di una spiegazione scientifica dell'universo. La fisica non è ancora completa, ma è ben compresa; la biologia ha svelato molti antichi misteri intorno alla natura della vita. Vi sono lacune nella nostra comprensione di questi campi, ma essi non sembrano intrattabili. Abbiamo un'idea di ciò a cui può somigliare una soluzione a questi problemi.

Anche nella scienza della mente c'è stato un notevole progresso. Il recente lavoro in scienza cognitiva e neuroscienza ci sta conducendo verso una migliore comprensione del comportamento umano e dei processi che lo guidano...

La coscienza, tuttavia, è altrettanto misteriosa di quanto lo fosse una volta. Sembra ancora del tutto misterioso che la causazione di un comportamento debba essere accompagnata da una vita interiore soggettiva. Abbiamo buone ragioni per credere che la coscienza sorga da sistemi fisici come i cervelli, ma abbiamo poche idee sul come sorga, o semplicemente sul perché essa esista. Come può un sistema fisico come il cervello essere anche un *soggetto di esperienza (experiencer)*? Perché dovrebbe esserci *qualcosa che si provi ad essere* tale sistema? Le attuali teorie scientifiche difficilmente sfiorano le questioni davvero difficili circa la coscienza. Non ci manca semplicemente la teoria giusta; siamo del tutto all'oscuro su come la coscienza si inserisca nell'ordine naturale... (1996, xi).

Tra gli scienziati cognitivi e i filosofi, sono comuni due reazioni opposte alla diagnosi di Chalmers: gli ottimisti sostengono che alla fine anche il problema della coscienza fenomenica troverà una spiegazione scientifica adeguata. La scienza ha sempre avuto successo nel trovare soluzioni, e lo stesso avverrà anche per la coscienza, o così dicono. Al contrario, la reazione pessimista riconosce un fallimento in linea di principio – non riusciremo mai a dare una spiegazione riduzionistica della coscienza.

* Traduzione di Alessandro Dell'Anna

L'emergenza fenomenica

In ciò che segue mi concentrerò sulla visione pessimistica. Questa visione interpreta la coscienza (e le qualità fenomeniche) come fenomeni *emergenti*; specificamente, come fenomeni che sono *irriducibili*, e perciò emergenti in un senso molto forte.

Dal momento che i criteri che dovrebbero distinguere le proprietà emergenti da quelle non-emergenti sono molto controversi, esaminerò diverse nozioni di emergenza, che hanno forza diversa. In particolare, distinguerò teorie dell'emergenza deboli, diacroniche e sincroniche. Queste distinzioni aiutano a chiarire il tipo di emergenza che andrebbe attribuito alla coscienza fenomenica, di contro, per esempio, ai tipi di emergenza che potrebbero essere attribuiti a certe proprietà delle reti connessioniste e dei robots. Inoltre vedremo quali conseguenze accompagnano ciascuna nozione di emergenza. In tal modo diventerà chiaro che certe nozioni deboli di emergenza sono compatibili col riduzionismo della proprietà, mentre le versioni forti non lo sono¹.

Emergentismo debole, diacronico e sincronico

Tra le diverse teorie dell'emergentismo, tre meritano particolare interesse: l'emergentismo *sincronico*, l'emergentismo *diacronico* e una forma di emergentismo *debole*. Nell'emergentismo sincronico è di centrale importanza la relazione tra le proprietà di un sistema e la sua microstruttura, cioè la disposizione (arrangement) e le proprietà delle parti del sistema. Una proprietà di un sistema è considerata emergente se è *irriducibile* o, e ritengo sia la stessa cosa, se *non* è spiegabile riduzionisticamente. Al contrario, l'emergentismo diacronico è interessato soprattutto nella *prevedibilità* di nuove proprietà. Sono considerate emergenti quelle proprietà che non possono essere predette, in linea di principio, prima della loro prima occorrenza. Le forme di emergentismo diacronico e sincronico non sono indipendenti l'una dall'altra, dal momento che le proprietà irriducibili sono *eo ipso* non prevedibili, prima della loro prima apparizione. Quindi le proprietà emergenti sincronicamente sono anche diacronicamente emergenti, ma non viceversa.

Entrambe queste versioni forti di emergentismo sono basate su una comune teoria debole, dalla quale si possono sviluppare aggiungendo ulteriori tesi. Attualmente *l'emergentismo debole* pervade le teorizzazioni emergentiste sia nelle teorie dell'auto-organizzazione che in vari approcci alla scienza cognitiva, ad esempio il connessionismo, la robotica e la vita artificiale. Esso è compatibile con vari approcci riduzionistici; e qualche proponente dell'emergentismo debole cita tale compatibilità come uno dei suoi meriti, comparato alle versioni più forti di emergentismo.

2 Emergentismo debole

La prima tesi delle attuali teorie dell'emergenza – la tesi del *monismo fisico* – riguarda la natura dei sistemi dotati di proprietà o strutture emergenti. Essa dice che i portatori di proprietà emergenti sono solo entità fisiche. Nega l'esistenza di qualsiasi componente soprannaturale che sia responsabile delle proprietà emergenti del sistema. Quindi rifiuta ogni dualismo sostanziale (substance-dualistic positions); in quanto questo basa proprietà quali essere vivo o avere stati cognitivi su dei portatori sopran-

¹ Per ulteriori dettagli sull'emergenza e l'emergentismo, si veda il mio libro (Stephan 1999).

Achim Stephan

naturali, quali un'entelechia o una *res cogitans*, rispettivamente. Invece sostiene che gli esseri viventi e i sistemi cognitivi consistono delle stesse unità di base della natura inanimata: non è altro che la specifica sequenza di stati fisico-chimici altamente complessi a realizzare il loro comportamento vitale o i loro stati mentali.

- (1) *Monismo fisico*. Le entità esistenti o che vengono all'esistenza nell'universo consistono esclusivamente di componenti fisiche. Allo stesso modo, proprietà, disposizioni, comportamenti o strutture classificati come emergenti sono istanziati da sistemi che consistono esclusivamente di entità fisiche.

Mentre la prima tesi pone le proprietà e le strutture emergenti nel quadro di un naturalismo fisicalistico, la seconda tesi – la tesi delle *proprietà sistemiche* – delimita il tipo di proprietà possibili candidate ad essere emergenti. Essa si basa sull'idea che proprietà generali di sistemi complessi cadono in classi diverse, quelle che hanno anche alcune parti dei sistemi e quelle che nessuna parte dei sistemi ha. Queste ultime proprietà sono chiamate proprietà *sistemiche* o *collettive*. Le proprietà generali sono proprietà di un tipo generale, come avere un peso; non sono proprietà specifiche, come avere un peso di 85,4 chilogrammi.

- (2) *Proprietà sistemiche*. Le proprietà emergenti sono proprietà sistemiche. Una proprietà di un sistema è sistemica se e solo se il sistema la possiede, ma non la possiede alcuna sua parte.

Dovrebbe essere chiaro che esistono sistemi con proprietà sistemiche sia naturali che artificiali. Chi negasse la loro esistenza dovrebbe sostenere che tutte le proprietà di un sistema sono già istanziate da una qualche parte del sistema. Infiniti esempi confutano tale affermazione; ad esempio, tra le proprietà di una rana vi è il saltare e il gracchiare, ma nessuna sua parte (testa, zampa, cuore, né alcun gruppo cellulare) può saltare o gracchiare; e tra le proprietà di una rete concessionista c'è il generalizzare o il riconoscere pattern, ma nessuna singola proprietà di essa (unità, ecc.) ha tali proprietà.

Mentre il primo tipo di tesi circoscrive il tipo di parti dalle quali può essere costruito il sistema avente proprietà emergenti, e mentre la seconda tesi caratterizza in maggiore dettaglio il tipo di proprietà che può essere emergente, la terza tesi specifica il tipo di relazione che si instaura tra la microstruttura di un sistema e le sue proprietà emergenti:

- (3) *Determinazione sincronica*. Le proprietà e le disposizioni al comportamento di un sistema dipendono nomologicamente dalla sua struttura, vale a dire, dalle proprietà e dalla disposizione delle sue parti. Non vi può essere alcuna differenza nelle proprietà sistemiche di un sistema senza qualche differenza nelle proprietà o nella disposizione delle sue parti.

Nel dibattito recente la tesi della determinazione sincronica è sostenuta talvolta in forma meno forte come tesi della *sopravvenienza mereologica*, la quale afferma che le proprietà intrinseche del sistema sopravvivono sulle proprietà delle sue parti e sulla loro disposizione. Quindi, non c'è neanche differenza nelle proprietà sistemiche senza differenza nelle proprietà delle parti o nella loro disposizione. Comunque, la tesi della

L'emergenza fenomenica

soppravvenienza mereologica è più debole della tesi della determinazione sincronica, poiché non implica, strettamente parlando, la *dipendenza* delle proprietà del sistema dalla sua microstruttura, implica solo la loro *covarianza*.

Chiunque neghi la tesi della determinazione sincronica o deve ammettere proprietà di un sistema che non sono legate alle proprietà e alla disposizione delle sue parti, o deve supporre che qualche altro fattore, in questo caso fattori non naturali, sono responsabili delle diverse disposizioni di sistemi che sono identici nella loro microstruttura. Ci si potrebbe trovare ad ammettere, ad esempio, che vi sono oggetti che hanno le stesse parti nell'identica disposizione dei diamanti, ma che non hanno la stessa durezza dei diamanti. Ciò sembra implausibile. Altrettanto implausibile è l'idea che possano esistere due organismi fisicamente identici, l'uno in grado di sopravvivere e l'altro no. Nel caso dei fenomeni mentali le opinioni possono differire ancora di più; ma una cosa sembra essere certa: chiunque creda, ad esempio, che due creature fisicamente identiche possano essere tali che una sia daltonica e l'altra no, non è un fiscalista. Considerazioni simili valgono per gli atteggiamenti proposizionali solo nella misura in cui non si sottoscrive l'esternismo, vale a dire, se non si sostiene, ad esempio, che il contenuto di una credenza dipenda essenzialmente dalla natura dei riferimenti dei pensieri e dei concetti dell'osservatore.

Come delineato fin qui, l'emergentismo debole specifica i criteri minimi per le proprietà emergenti. Esso è *debole* in quanto compatibile con pressoché tutti gli approcci riduzionistici. Esso è anche la base comune per tutte le teorie dell'emergenza più forti. In più – e questo è un motivo per distinguerlo come teoria di per sé – esso è abbracciato sia da filosofi (Bunge 1977) che da scienziati cognitivi (Rumelhart e McClelland 1986, Varela et al. 1991) proprio nella sua forma debole.

3 Emergentismo diacronico

Tutte le teorie diacroniche dell'emergenza su una tesi sull'occorrenza di novità genuine nell'evoluzione. Questa tesi esclude tutte le posizioni performazioniste (performationist).

- (4) *Novità*. Durante l'evoluzione si verificano di continuo esempi di genuine novità. Entità già esistenti formano nuove costellazioni che producono nuove strutture che possono costituire nuove entità con proprietà e comportamenti nuovi.

Ad ogni modo, la tesi della novità non trasforma di per sé una teoria dell'emergenza debole in una forte, dato che il fiscalismo riduzionista resta compatibile con tale variante di emergentismo. Solo l'aggiunta della tesi della *imprevedibilità* (unpredictability), in linea di principio, porterà a forme più forti di emergentismo *diacronico*.

La prima occorrenza di una proprietà sistemica può essere imprevedibile per almeno due ragioni diverse. Innanzitutto, in linea di principio, è imprevedibile se è irriducibile. Ciò non significa comunque che ulteriori occorrenze potrebbero non essere adeguatamente predette. In secondo luogo, può essere imprevedibile poiché la microstruttura del sistema che esemplifica per la prima volta la proprietà nell'evoluzione è imprevedibile. Poiché, se la microstruttura di un nuovo sistema emergente è impreve-

dibile, tali sono le proprietà che dipendono nomologicamente da essa. Dal momento che nel primo caso i criteri per essere imprevedibile sono identici a quelli per essere irriducibile, pospongo l'analisi di quelle proprietà alla prossima sezione, e mi limito qui al secondo caso, *l'imprevedibilità di struttura*. Questa versione guadagna un significato notevole rispetto al forte interesse nei sistemi dinamici e nei processi caotici.

La struttura di un sistema appena formato può essa stessa essere imprevedibile per due ragioni. Se l'universo è non deterministico, allora le sue nuove strutture saranno imprevedibili. Comunque, da un punto di vista emergentista, non è di alcun interesse se la comparsa di una struttura è imprevedibile solo come conseguenza della sua indeterminazione – molti emergentisti sostengono che lo sviluppo di nuove strutture è governato da leggi deterministiche. Ciononostante, il formarsi deterministico di alcune strutture può essere *in linea di principio imprevedibile* se esse sono governate da leggi del caos deterministico.

Un risultato essenziale della teoria del caos è che esistono funzioni matematiche – anche molto semplici – il cui stesso comportamento non può essere predetto. Solo il sorgere della “matematica sperimentale” su computer molto efficienti ha rivelato, ad esempio, le proprietà di varie funzioni logistiche. La loro imprevedibilità intramatematica ha a che fare con il comportamento aperiodico di queste funzioni, dal quale valori iniziali marginalmente diversi di qualche variabile possono portare a traiettorie radicalmente differenti delle funzioni.

Un esempio standard è la funzione logistica $f(x) = \mu x(1-x)$ per $0 \leq x \leq 1$. Per un parametro μ con $0 \leq \mu \leq 4$ la funzione logistica mappa l'intervallo $[0,1]$ su se stessa. Di particolare interesse è come il parametro μ esercita un'influenza sul comportamento a lungo termine della funzione, quando iterato ripetutamente. Per $0 \leq \mu \leq 1$ la situazione è ovvia. Tutti i valori iniziali della variabile x lasciano la funzione $f(x)$ approssimata al valore 0 dopo un numero sufficientemente grande di iterazioni, quindi l'origine è l'attrattore. Per $1 \leq \mu \leq 3$ esiste esattamente un attrattore A di valore $A = 1 - 1/\mu$: la funzione si equilibra su un valore stabile. Se $\mu = 3$, il punto fisso della funzione è “marginalmente stabile”; la convergenza è decisamente lenta – un'indicazione per un cambiamento fondamentale nel comportamento della funzione. Per valori più grandi la dinamica diventa notevolmente complessa. Nel caso di $3 < \mu < 1 + \sqrt{6}$ i valori oscillano tra due punti fissi. Aumentando il valore di μ anche gli attrattori del “period” due diventano instabili. Abbiamo un ciclo di “period” quattro (cioè, dopo quattro iterazioni i valori della funzione approssimano in ciascun caso i quattro punti fissati). A 3.56 il “periodo” raddoppia di nuovo e diventa otto, a 3.567 diventa sedici, e quindi otteniamo una sequenza di “periodi” che cresce rapidamente fino a 32, 64, 128, ecc. – significativamente si parla di cascate. A circa 3.58 questa sequenza finisce. Il “period” ha raddoppiato se stesso un infinito numero di volte. In futuro non sembrano essere possibili previsioni. Valori iniziali x marginalmente diversi portano a traiettorie radicalmente diverse della funzione iterata. I valori saltano “caoticamente”, convergenza e divergenza non sono distinguibili: domina il caos.

In questo modo, sembra che proprio la scienza più esatta di tutte ci abbia portato indietro a una dei punti di partenza dell'emergentismo. Mentre – dopo i successi pionieristici in fisica e in chimica – oggi non annoveriamo più le proprietà e le disposizioni dei composti chimici tra i fenomeni emergenti, le analisi del caos deterministico suggeriscono l'esistenza di sistemi che possono sviluppare strutture in linea

L'emergenza fenomenica

di principio imprevedibili, e così possono mostrare un comportamento *emergente strutturale* (structure-emergent).

Ovviamente si potrebbe argomentare che un calcolatore laplaciano potrebbe predire correttamente anche processi caotici. Non è ancora stato stabilito se sarebbe “davvero” così. Dipende soprattutto da che tipo di informazione permettiamo che abbia una tale creatura. Ad esempio, nelle osservazioni di Alexander (1920, ii, 72 f., 328) il calcolatore di Laplace conosce molti stati precedenti dell'intero mondo e, in più, tutte le leggi naturali che governano i cambiamenti nel mondo. Egli sembra in grado di estrapolare dalla sua conoscenza di tutti gli eventi capitati nell'universo fino ad ora anche il corso dei processi caotici. Ma su quali basi può farlo? Dal momento che i processi caotici sono aperiodici, dai processi capitati fino a un certo momento non si può determinare con certezza la formula esatta che descriverebbe il loro corso successivo. Anche se il corso successivo del mondo è governato da leggi deterministiche, non segue solo dagli eventi e dagli stati precedenti, da *quali* leggi esso è governato. Continuazioni del tutto differenti sembrano essere compatibili con il precedente corso del mondo. Pertanto anche un calcolatore laplaciano può fallire nelle sue previsioni. Comunque, se si concede che egli conosca *tutti* i dettagli degli stati precedenti del mondo – fino a un numero infinito di cifre -, e se si concede che egli conosca *a priori* quali processi sono governati da quali *specifiche* leggi caotiche, allora certo egli sarebbe in grado di predire il formarsi di strutture governate da queste leggi. Lascero aperta la questione se sia plausibile o meno ascrivere tale conoscenza a tale favolosa creatura. Come minimo si può essere certi che creature con le nostre capacità mentali non hanno queste capacità di previsione, e così possiamo legittimamente supporre che, laddove il caos esiste, esistono strutture in linea di principio imprevedibili.

- (5) *Imprevedibilità strutturale*. L'emergere di una nuova struttura è in linea di principio imprevedibile se la sua formazione è governata da leggi del caos deterministico. Allo stesso modo, ogni proprietà istanziata dalla nuova struttura è in linea di principio imprevedibile.

Sebbene l'emergentismo diacronico con la tesi dell'imprevedibilità strutturale implichi l'imprevedibilità di tutte le proprietà istanziate da sistemi che emergono da processi caotici, non implica per questo la loro irriducibilità. L'imprevedibilità, in linea di principio, delle proprietà sistemiche è interamente compatibile con il loro essere riducibili alla microstruttura del sistema che le istanzia. Comunque, anche le proprietà sistemiche sono imprevedibili in linea di principio, se sono irriducibili. Questa caratteristica è al centro dell'emergentismo sincronico, la dottrina più importante per l'interpretazione della coscienza fenomenica, e che esamino nel prossimo paragrafo.

4 Emergentismo sincronico

La nozione di *irriducibilità* è al centro di tutte le versioni forti di emergentismo. Sebbene potrei introdurlo in contrasto col concetto di *spiegazione riduzionistica* come definito da Levine e Kim, è tuttora affascinante seguire prima il tentativo di C. D. Broad di spiegare una nozione (sincronica) forte di emergenza. Nel suo libro *La men-*

Achim Stephan

te e il suo posto nella natura, troviamo un passaggio che oggi giorno può contare come perfettamente classico; esso dice:

Messa in termini astratti la teoria emergente asserisce che ci sono certi interi, composti (mettiamo) dei costituenti A, B e C in relazione R tra loro; che tutti gli insiemi composti da costituenti dello stesso tipo di A, B, e C in relazioni dello stesso tipo di R hanno certe proprietà caratteristiche; che A, B e C sono capaci di capitare in un altro tipo di complesso dove la relazione non è dello stesso tipo di R; e che le proprietà caratteristiche dell'intero R(A, B, C) non possono, neanche in teoria, essere *dedotte* dalla conoscenza più completa delle proprietà di A, B e C in isolamento o in altri interi che non hanno la forma R(A, B, C) (1925, 61).

Secondo la definizione di Broad, una proprietà sistemica che si suppone sia nomicamente dipendente dalla microstruttura del suo sistema (per la tesi della determinazione sincronica) viene detta *irriducibile*, e per questo *emergente*, se e solo se non può essere dedotta dalla disposizione delle parti del suo sistema e dalle proprietà che esse hanno "in isolamento" o in altri sistemi (più semplici).

Sebbene, a prima vista, può sembrare che la definizione di Broad ci fornisca una spiegazione chiara e distinta di che cosa significhi per una proprietà sistemica essere irriducibile (o non deducibile), uno sguardo ulteriore rivela che sono nascosti due tipi diversi di irriducibilità, i quali hanno conseguenze piuttosto diverse. Come vedremo, un tipo di irriducibilità sembra implicare una *causazione verso il basso* (downward), mentre l'altra sembra implicare l'*epifenomenismo*. Il fallimento nel tenere separate le due specie di irriducibilità ha confuso il dibattito recente sull'emergenza delle proprietà.

Per chiarire le cose, prima discuterò quando una proprietà sistemica è *riducibile*. Perché ciò avvenga devono essere soddisfatte due condizioni: la prima è che dal comportamento delle sole parti del sistema deve seguire che il sistema ha una qualche proprietà P. La seconda richiede che il comportamento esibito dalle parti del sistema quando esse sono parti del sistema segue dal comportamento che esse mostrano in isolamento o in sistemi più semplici del sistema in questione. Se entrambe le condizioni sono soddisfatte, il comportamento delle parti del sistema in *altri* contesti rivela quali proprietà sistemiche ha il sistema reale. In altre parole, quelle proprietà sono riducibili. Dal momento che entrambe le condizioni sono indipendenti l'una dall'altra, si hanno due diverse possibilità per l'occorrenza di proprietà sistemiche *irriducibili*: (a) una proprietà sistemica P di un sistema S è *irriducibile se non segue*, neanche in linea di principio, dal comportamento delle parti del sistema che S ha la proprietà P; e (b) una proprietà P di un sistema S è *irriducibile se non segue*, neanche in linea di principio, dal comportamento delle parti del sistema in costellazioni più semplici di S come esse si comportano in S.

Così, una condizione necessaria affinché una proprietà sistemica sia riducibile è che la sua istanziazione deve seguire dal comportamento delle parti del suo portatore. In altri termini: dal comportamento delle parti del sistema dovrebbe seguire che il sistema ha tutte le caratteristiche essenziali per avere la proprietà sistemica. Broad, ad esempio, ritiene che questa condizione, racchiusa nel primo criterio per la riducibilità,

L'emergenza fenomenica

sia sempre soddisfatta nel caso delle proprietà caratteristiche dei composti chimici e degli organismi viventi. Le loro proprietà potrebbero essere irriducibili solo violando il secondo criterio, il che significa che dal comportamento delle parti del sistema in altri sistemi (più semplici), non seguirebbe come esse si comportano nel sistema reale.

Al contrario, egli sostiene che l'irriducibilità delle qualità secondarie e delle qualità fenomeniche risulta già dalla violazione della prima condizione, poiché esse non erano adeguatamente caratterizzabili né dal comportamento macroscopico né da quello microscopico delle parti del sistema, nemmeno in linea di principio. Dal momento che, quando diciamo che un certo oggetto è rosso o una sostanza chimica ha l'odore dell'ammoniaca, non intendiamo che le parti corrispondenti del sistema si *comportano* o si *muovono* in un certo modo. Nessun progresso scientifico potrebbe in alcun modo cambiare questo stato di cose². Broad ha illustrato la distinzione fondamentale tra proprietà *analizzabili* e *inalizzabili* (dal punto di vista comportamentale) riferendosi rispettivamente a proprietà caratteristiche dell'organismo e a qualità secondarie³. Se le qualità secondarie e le qualità fenomeniche non sono analizzabili, neanche in linea di principio, non c'è alcuna speranza che una crescita della conoscenza scientifica chiuderà il gap tra i processi fisici e le qualità secondarie o, rispettivamente, tra i processi fisiologici e gli stati fenomenici della coscienza.

Ora possiamo precisare in maggiore dettaglio il carattere della irriducibilità centrale per l'emergenza sincronica. La sua prima variante è basata sull'inalizzabilità comportamentale delle proprietà sistemiche:

- (6) *Inalizzabilità*. Le proprietà sistemiche che non sono analizzabili comportamentalmente – sia macro che microscopicamente – sono (necessariamente) irriducibili.

Comunque anche se le qualità secondarie e fenomeniche sono proprietà inalizzabili non segue da ciò che il comportamento specifico delle parti del sistema sulle quali sopravvengono sia esso stesso non deducibile dal comportamento che queste parti esibiscono in sistemi semplici. L'irriducibilità che risulta dalla violazione del primo criterio di riducibilità non implica di per sé la violazione del secondo principio di riducibilità.

D'altro canto, comunque, anche le proprietà sistemiche analizzabili possono essere irriducibili e per questo emergenti. Questo è il caso che si presenta quando è violato il secondo criterio di riducibilità, cioè quando il comportamento delle parti del sistema non segue dal loro comportamento in altre costellazioni (più semplici). Broad ritiene che tali esempi di comportamento irriducibile potrebbero verificarsi nei composti chimici e anche negli organismi⁴. La sua idea centrale è che le parti di una struttura

² Ad ogni modo, si dibatte se il riferimento agli usi linguistici potrebbe rispondere definitivamente a domande riguardanti la riducibilità. In particolare, P. Churchland ha opposto argomenti nello stile di Broad (1988, 29 ss).

³ Le proprietà che qui per semplicità chiamo "inalizzabili", potrebbero essere analizzabili in modi diversi dai caratteri comportamentali. Ad esempio, un certo odore potrebbe essere analizzato come una miscela di odori di Chanel n. 5 e farina di pesce. Questa, comunque, non sarebbe un'analisi basata sui concetti di movimento e di comportamento.

⁴ In un recente articolo (Boogerd et al., di prossima pubblicazione) argomentiamo che la biologia cellulare offre molti esempi di questo tipo di emergenza forte.

Achim Stephan

davvero nuova, come ad esempio un organismo rispetto a qualsiasi composto inorganico, potrebbe comportarsi in un modo non deducibile dal comportamento delle parti in altre strutture. Implicitamente, ciò significa che il comportamento reale delle parti che interagiscono negli interi non risulta dal loro comportamento in coppie (*pairs*). Se il comportamento di alcune parti del sistema è irriducibile su questo piano, allora anche tutte le proprietà che dipendono nomologicamente dal comportamento delle parti del sistema (la riproduzione, ad esempio) sono irriducibili.

Così possiamo specificare la seconda variante dell'irriducibilità delle proprietà sistemiche. Essa è basata sulla non deducibilità del comportamento delle parti del sistema:

- (7) *Irriducibilità del comportamento dei componenti.* Il comportamento specifico dei componenti del sistema all'interno del sistema è irriducibile se non segue dal comportamento dei componenti in isolamento, o in altre costellazioni (più semplici).

Una violazione del secondo criterio di riducibilità, evidente nell'irriducibilità del comportamento dei componenti, non implica comunque una violazione del primo criterio di riducibilità. Le proprietà sistemiche che non possono essere ridotte perché il comportamento delle parti del sistema è irriducibile potrebbero comunque essere comportamentalmente analizzabili. Quindi le due condizioni necessarie di riducibilità come quelle di irriducibilità basate sulla violazione di queste condizioni sono indipendenti l'una dall'altra. In sintesi, da (6) e (7) otteniamo la seguente versione modificata dell'irriducibilità di proprietà sistemica:

- (8) *Irriducibilità.* Una proprietà sistemica è irriducibile se (6) non è analizzabile sul piano del comportamento né macro né microscopicamente, o (7) se il comportamento specifico dei componenti del sistema, sui quali sopravvivono le proprietà del sistema, non segue dal comportamento dei componenti in isolamento o in altre costellazioni (più semplici).

Quindi, dobbiamo distinguere due tipi completamente diversi di irriducibilità delle proprietà sistemiche. Altrettanto diverse sembrano le conseguenze che derivano da essi. Se una proprietà sistemica è irriducibile a causa dell'irriducibilità del comportamento delle parti sulle quali sopravviene la proprietà, sembra che abbiamo un caso di "causazione verso il basso". Poiché, se il comportamento dei componenti non è riducibile alla loro disposizione e al comportamento che mostrano in sistemi più semplici o in isolamento, sembra sussistere una qualche influenza causale "verso il basso" esercitata dal sistema stesso (o dalla sua struttura) sul comportamento delle parti del sistema. Sicuramente, se un tale esempio di "causazione verso il basso" esistesse, non porterebbe alla violazione di qualche assunzione largamente condivisa, come ad esempio il principio della chiusura causale del mondo fisico. All'interno del mondo fisico dovremmo semplicemente accettare ulteriori tipi di influenze causali, oltre ai già noti tipi base di mutue interazioni.

Al contrario, l'occorrenza di proprietà inanalizzabili non implica alcun tipo di causazione verso il basso. I sistemi dotati di proprietà inanalizzabili, le quali dipendono nomologicamente dalla microstruttura del loro portatore, non hanno bisogno di essere costituiti in modo da essere irriducibili al comportamento dei loro componenti. Ciò

L'emergenza fenomenica

non implica che la struttura del sistema abbia un'influenza causale verso il basso sulle parti del sistema, né c'è alcuna ragione di assumere che le stesse proprietà inanalizzabili esercitino un'influenza causale sulle parti del sistema. Piuttosto si possono avere dei dubbi sul come le proprietà inanalizzabili possano giocare un qualsiasi ruolo causale. Dal momento che non sono analizzabili sul piano comportamentale – cioè, non sembrano corrispondere ad alcun “meccanismo” né sembrano risultare da alcun “meccanismo” – è difficile capire come esse stesse possano essere efficaci casualmente. E se non si riesce a capire *come* le proprietà in analizzabili possano giocare un ruolo causale, è difficile concepirle altro che come epifenomeni⁵.

È istruttivo vedere come le nozioni di *spiegazione riduzionistica* di Kim e Levine (che possono anche servire a definire la nozione di emergenza sincronica) si relazionino all'analisi precedente. Entrambi gli autori notano che, come primo passo, bisogna elaborare il concetto di proprietà. Kim inoltre la chiama “procedura primaria”, nella quale dobbiamo costruire o ricostruire la proprietà da ridurre relazionalmente o estrinsecamente. Egli considera un dominio di proprietà (o fenomeni, fatti, ecc.) B che serve come base di riduzione, e una proprietà E che sarà ridotta (1999, 10-12)⁶:

Passo 1. E deve essere *funzionalizzata* – ovvero, E deve essere costruita o ricostruita come una proprietà definita dalle sue relazioni nomico/causali ad altre proprietà, specificamente le proprietà nella base di riduzione B.

Ora, questa condizione, che possiamo chiamare la *condizione di funzionalizzazione*, non è altro che una nuova versione della condizione di analizzabilità di Broad. Se il concetto di proprietà non può essere elaborato “in forma” per la riduzione, ovvero, se la proprietà non può essere funzionalizzata, allora necessariamente la riduzione fallisce. La proprietà si rivelerà essere irriducibile, quindi sincronicamente emergente. Ma, come si è appena visto, bisogna notare che se le proprietà irriducibili esercitano “causazione verso il basso”, esse non fanno ciò *in quanto* inanalizzabili (o “non funzionalizzabili”).

Come secondo passo sia Kim che Levin menzionano il compito empirico di trovare i soggetti di una proprietà che potrebbero essere elaborati in forma per la riduzione: “lo stadio 2 implica il lavoro empirico di scoprire esattamente che cosa sono tali meccanismi sottostanti” (Levine 1993, 132). Questo compito non è menzionato esplicitamente da Broad, poiché egli discute situazioni in cui si hanno già correlazioni nomologiche tra la microstruttura di un sistema e alcune proprietà sistemiche, ma non si può ridurre la proprietà dalla base di riduzione, cioè dal comportamento delle parti del sistema in sistemi più semplici o in isolamento.

Passo 2. Trova soggetti di E in B. Se si vuole la riduzione, o la spiegazione riduzionistica, di un esempio particolare di E in un dato sistema, trovare la particolare proprietà realizzante P in virtù della quale E è istanziata in questo sistema in

⁵ In Stephan (1997) ho provato a trovare un modo in cui si riesca comunque attribuire poteri causali alle proprietà inanalizzabili.

⁶ Levine la pone in questi termini: “lo stadio 1 implica il processo (relativamente? quasi?) a priori di elaborare il concetto di proprietà da ridurre ‘in forma’ per la riduzione, identificando il ruolo causale del quale stiamo cercando i meccanismi sottostanti” (1993, 132).

Achim Stephan

questa occasione (similmente per classi di sistemi che appartengono alle stesse specie o tipi di strutture).

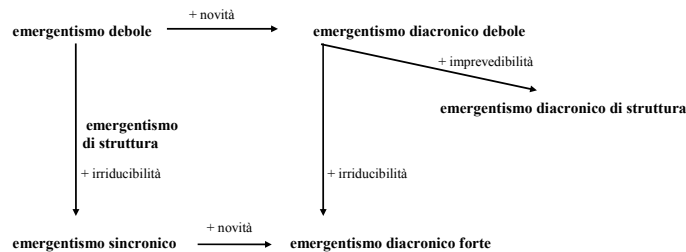
Kim, in opposizione a Levin, postula un terzo passo nel quale si riferisce allo sviluppo delle teorie che possono spiegare come i soggetti di E, la proprietà da ridurre, svolgono il particolare ruolo funzionale caratteristico di E.

Passo 3. Trova una teoria (al livello di B) che spieghi come i soggetti di E svolgono il compito causale che è costitutivo di E (cioè il ruolo causale specificato al passo 1). Tale teoria può anche spiegare altre significative relazioni nomico/causali nelle quali E gioca un ruolo.

Sebbene *prima facie* il terzo passo di Kim sembra corrispondere al secondo criterio di Broad, uno sguardo più ravvicinato rivela importanti differenze. Kim considera solo teorie che connettono la base di riduzione della proprietà sistemica (tipicamente: la microstruttura del sistema) alla proprietà da ridurre. Così, nulla si può dire su alcuna influenza causale “verso il basso” del sistema o della sua struttura. Comunque, l’analisi dettagliata dell’approccio di Broad ci ha dato i mezzi per diagnosticare casi in cui dovremmo assumere una causazione verso il basso, cioè dove nessuna teoria è in grado di spiegare il comportamento delle parti del sistema all’interno di questo stesso sistema con riferimento al loro comportamento in sistemi più semplici.

5 Sinossi

La figura seguente mostra le relazioni logiche tra le diverse versioni di emergentismo:



L'emergentismo diacronico debole risulta dall'emergentismo debole aggiungendo una dimensione temporale sotto forma della tesi di novità. Entrambi le versioni sono compatibili col fiscalismo riduzionistico. Le teorie deboli dell'emergenza oggi sono usate soprattutto in scienza cognitiva, in particolare per la caratterizzazione delle proprietà sistemiche delle reti connessioniste, e nelle teorie dell'auto-organizzazione. *L'emergentismo sincronico* risulta dall'emergentismo debole con l'aggiunta della tesi di irriducibilità. Questa versione di emergentismo è importante per la filosofia della mente, in particolare nel dibattito sul fiscalismo non riduzionista e la coscienza fenomenica. Non è compatibile col fiscalismo riduzionista. *L'emergentismo diacronico*

L'emergenza fenomenica

forte differisce dall'emergentismo sincronico a causa della dimensione temporale nella tesi della novità. *L'emergentismo strutturale* è interamente dipendente dall'emergentismo sincronico. Esso risulta dall'emergentismo diacronico debole con l'aggiunta della tesi dell'imprevedibilità di struttura. Sebbene l'emergentismo strutturale sottolinei i confini della predizione negli approcci fisicalistici, esso è compatibile col fiscalismo riduzionista, e quindi è più debole dell'emergentismo sincronico. Le teorie del caos deterministico possono essere considerate un tipo di emergentismo strutturale. Questa prospettiva è importante nella ricerca evoluzionistica.

Coda

Un punto che concerne i fenomeni mentali irriducibili è la coscienza fenomenica. Tra gli altri, Levine e Chalmers sostengono che non possiamo spiegarla riduzionisticamente, e non per ragioni empiriche. Apparentemente, le proprietà fenomeniche non possono essere individuate dal loro ruolo causale o, come dice Levine: "Ciò che sembra responsabile del gap esplicativo, quindi, è il fatto che i nostri concetti di carattere qualitativo non rappresentano, almeno nei termini del loro contenuto psicologico, ruoli causali [...] Così, nella misura in cui nel nostro concetto di carattere qualitativo c'è un elemento che non è catturato dal suo ruolo causale, esso sfuggirà alla rete esplicativa di una riduzione fisicalista" (1993, 134). Kim sembra condividere questa valutazione quando dice: "Per cogliere il punto, senza troppe storie, mi sembra che le qualità fenomeniche, sentite, dell'esperienza, o qualia, sono proprietà intrinseche, se ne è qualcuna" (1998, 102). Quindi egli ritiene che la funzionalizzazione dei qualia non funzionerà. Per la stessa ragione già Broad le ritenne inanalizzabili. Così, i qualia sembrano i candidati perfetti per le proprietà irriducibili, cioè sincronicamente emergenti – proprietà, comunque, che non esercitano poteri causali verso il basso *in quanto* inanalizzabili.

6 Riferimenti

- Alexander, S. (1920), *Space, Time, and Deity*. Two Volumes. London: Macmillan.
- Boogerd, Fred C., Frank J. Bruggeman, Robert C. Richardson, Achim Stephan, and Hans V. Westerhoff (forthcoming), "Emergence and its place in nature. A case study of biochemical networks". *Synthese*.
- Broad, Charles D. (1925), *The mind and its place in nature*. London: Kegan Paul, Trench, Trubner & Co.
- Bunge, Mario (1977), "Emergence and the mind". *Neuroscience* 2, 501-509.
- Chalmers, David J. (1996), *The conscious mind*. Oxford: Oxford University Press.
- Churchland, P. (1988), *Matter and consciousness*. Revised Edition. Cambridge, Ma., London: MIT Press.
- Levine, J. (1993), "On leaving out what it's like". In: M. Davies and G. W. Humphreys (eds.) *Consciousness*. Oxford: Blackwell, 121-136.
- Kim, J. (1998), *Mind in a physical world*. Cambridge: MIT Press.

Achim Stephan

- Kim, J. (1999), "Making sense of emergence". *Philosophical Studies* 95, 3-36.
- Rumelhart, David E. and James L. McClelland (1986), "PDP models and general issues in cognitive science". In: D. E. Rumelhart, J. L. McClelland and the PDP Research Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. Vol. 1. Cambridge: MIT Press, 110-146.
- Stephan, A. (1997), "Armchair arguments against emergentism". *Erkenntnis* 46, 305-314.
- Stephan, A. (1999), *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*. Dresden, München: Dresden University Press.
- Varela, Francisco J., Evan Thompson, and Eleanor Rosch (1991), *The embodied mind*. Cambridge: MIT Press.

L'emergenza fenomenica