

L&PS – Logic and Philosophy of Science

Vol. VII, No. 1, 2009

LA RAZIONALITÀ IN MEDICINA

Numero monografico a cura di Vincenzo Crupi e Gian Franco Gensini

VINCENZO CRUPI, GIAN FRANCO GENSINI, <i>La razionalità in medicina. Introduzione</i>	p. 3
STEFANO CANALI, <i>Sviluppi, problemi e prospettive della medicina scientifica</i>	9
ROBERTO FESTA, VINCENZO CRUPI, PIERANIELE GIARETTA, <i>Deduzione, induzione e abduzione nelle scienze mediche</i>	41
VINCENZO CRUPI, ROBERTO FESTA, <i>Dai giochi d'azzardo ai test diagnostici: la teoria della decisione nella medicina clinica</i>	69
VARIA	
DANIELE CHIFFI, <i>A World in Time</i>	95
TYLER J. VANDERWEELE, <i>Criteria for the Characterization of Token Causation</i>	115
Information on the Journal	129

La razionalità in medicina. Introduzione

Vincenzo Crupi

Dipartimento di Area Critica Medico-chirurgica, Università di Firenze
e-mail: vincenzo.crupi@unifi.it

Gian Franco Gensini

Dipartimento di Area Critica Medico-chirurgica, Università di Firenze
e-mail: gianfranco.gensini@unifi.it

Questo numero di *L&PS – Logic & Philosophy of Science* ospita un inserto monografico che comprende, oltre a questa nota introduttiva dei curatori, tre saggi raccolti intorno al tema della razionalità in medicina. Prima di presentare brevemente i temi trattati nei tre contributi, riteniamo utile delinearne la collocazione disciplinare. Più precisamente, intendiamo suggerire che le tre parti dell'inserto, nel loro insieme, toccano molti dei principali problemi di interesse di un settore di studi relativamente unitario e ben definito che si può opportunamente chiamare *epistemologia della medicina*. Come si vedrà dalle osservazioni che seguono, l'epistemologia della medicina è una disciplina filosofica. Tuttavia, per i suoi contenuti, essa si rivolge a chiunque sia interessato ai fondamenti della medicina sul duplice versante della ricerca e della cura.

Da ogni buon dizionario di filosofia si apprende che uno dei significati principali del termine "epistemologia" (corrispondente all'uso dell'inglese *epistemology*) è quello di una riflessione sulla conoscenza, la sua acquisizione, la sua crescita e il suo impiego. L'epistemologia così intesa rappresenta un'area di studi vasta e articolata, i cui diversi settori hanno spesso proceduto interagendo poco o per nulla. Un caso istruttivo è rappresentato dalla filosofia della scienza. Poiché la scienza rappresenta una parte fondamentale delle nostre conoscenze, si ritiene solitamente che la filosofia della scienza sia una provincia dell'epistemologia. Nonostante questo diffuso riconoscimento, gli epistemologi hanno spesso ignorato i problemi della conoscenza scientifica, per concentrare l'attenzione sulla conoscenza ordinaria, relativa a oggetti ed eventi della vita quotidiana. Analogamente, i filosofi della scienza hanno dedicato quasi tutte le loro energie all'analisi dei problemi metodologici suscitati dalla pratica scientifica, senza occuparsi delle somiglianze e delle connessioni tra conoscenza scientifica e conoscenza ordinaria. Questo modo di procedere, in-

coraggiato anche dalla crescente specializzazione accademica, ha determinato una notevole differenza di problemi, concetti e linguaggi tra epistemologia e filosofia della scienza. In tempi recenti, tuttavia, ha avuto luogo un consapevole e diffuso tentativo di integrazione. Da una parte, i filosofi della scienza hanno reso maggiormente espliciti i presupposti e le implicazioni epistemologiche più generali delle loro riflessioni sul metodo scientifico. Dall'altra, molti aspetti del ragionamento comune sono stati riletti con l'ausilio di potenti sistemi formali – spesso riconducibili al quadro concettuale del cosiddetto “approccio bayesiano” (si veda Bovens e Hartmann 2003) – non lontani da strumenti di analisi tradizionali dei metodologi della scienza.

Questa tendenza all'integrazione fra epistemologia e filosofia della scienza trova nel campo della medicina un terreno particolarmente fertile, forse unico. Da una parte, infatti, la medicina clinica, insieme alla pratica giudiziaria (specie in ambito penale), è un esempio paradigmatico di quelle che si possono definire “pratiche esperte”, intese come procedure di indagine altamente strutturate nelle quali la conoscenza scientifica e il ragionamento comune inevitabilmente si intersecano.¹ L'attività degli ospedali, come quella dei tribunali, non è rivolta (o non lo è principalmente) all'acquisizione di conoscenze scientifiche. Tuttavia gli agenti che vi operano sono chiamati ad affrontare problemi epistemologicamente rilevanti e notevolmente complessi, anche perché le conoscenze acquisite attraverso la ricerca scientifica propriamente intesa vi svolgono un ruolo non secondario. Dall'altra parte (e qui, però, l'analogia con il diritto viene meno), la medicina è anche ambito relativamente autonomo di indagine scientifica nel senso tradizionale dell'interazione fra teorizzazione, osservazione e sperimentazione. Si profilano così due versanti di analisi: nell'*epistemologia della pratica clinica*, si tratta della questione “in che modo le informazioni disponibili (compresi i risultati della ricerca medica) possono essere impiegate in vista dell'elaborazione di diagnosi e scelte cliniche appropriate?”; nella *filosofia della scienza medica*, d'altra parte, della questione “in che modo la ricerca medica può acquisire conoscenze valide?”. Come si vede, i due problemi sono distinti ma connessi, e concorrono appunto a costituire il più ampio ambito dell'epistemologia della medicina.² (Si veda la fig. 1, alla pagina successiva.)

¹ L'assonanza di questa terminologia con la nozione di sistema esperto non è casuale. Tipicamente, la ricerca sui sistemi esperti ha mirato a fornire strumenti di sostegno alle valutazioni e decisioni di agenti impegnati in quelle qui definite “pratiche esperte”. Su questo tema, è utile rinviare a Festa (in corso di stampa) e ai riferimenti ivi contenuti.

² L'ambito qui indicato come epistemologia della pratica clinica è, per esempio, al centro di lavori come Shaffner (1985), Giaretta e Federspil (1998), Campaner e Festa (2005b). Rien-

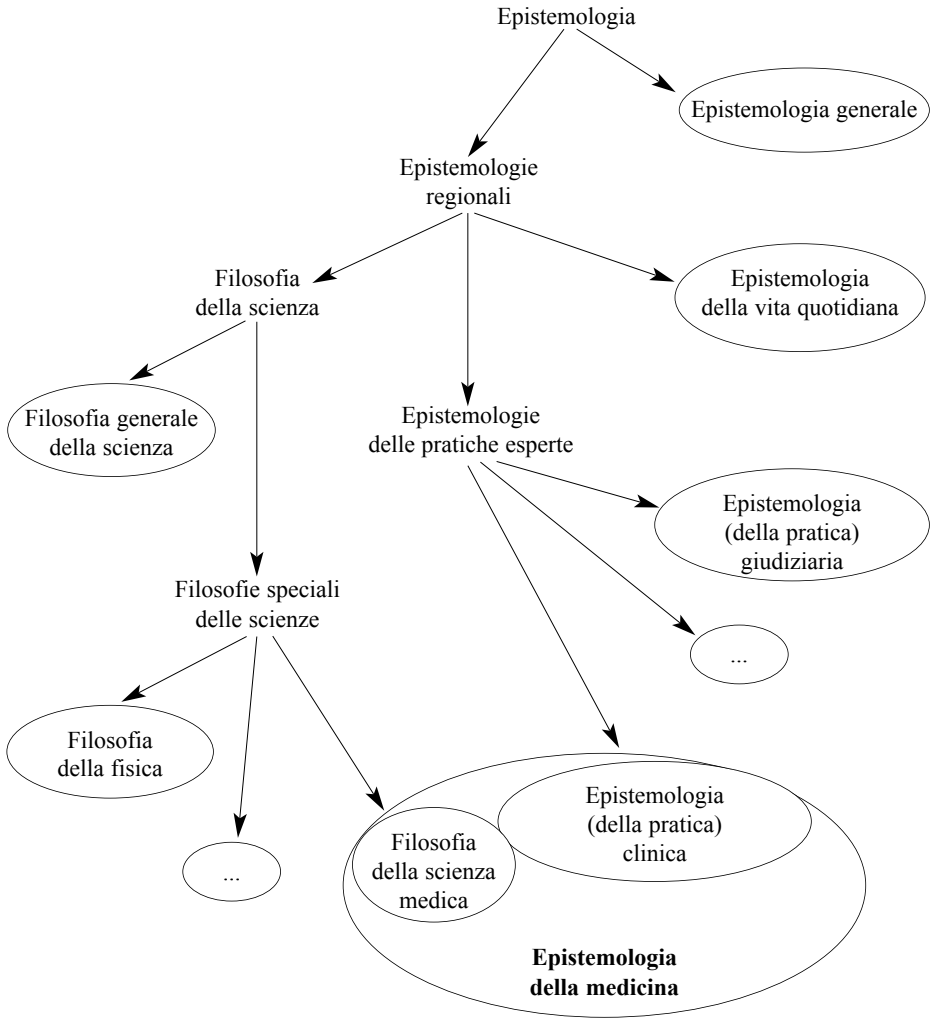


Fig. 1 - Una mappa della ricerca epistemologica.

trano invece prevalentemente nella filosofia della scienza medica i temi trattati da Vineis (1990), Shaffner (1993), Thagard (1999) e Campaner e Festa (2005a). Entrambe le componenti dell'epistemologia della medicina sono opportunamente segnalate e discusse da Corbellini (2003).

Per ricondurre questa digressione ai contributi presentati di seguito, sarà sufficiente segnalare gli stretti rapporti fra l'indagine epistemologica e la nozione di razionalità. Come le altre aree della ricerca epistemologica, infatti, anche l'epistemologia della medicina ha un orientamento essenzialmente prescrittivo (sia esso più o meno esplicitamente dichiarato). Ciò significa che, a partire da un'analisi di credenze, metodi e pratiche rilevanti nell'ambito di studio affrontato, si mira appunto a discuterne i fondamenti razionali (e gli eventuali limiti). In questa prospettiva, il saggio con cui si apre l'inserito qui presentato, di Stefano Canali (*Sviluppi, problemi e prospettive della medicina scientifica*), offre un'ampia ricostruzione e discussione storico-critica dei presupposti epistemologici e filosofici a fondamento della medicina scientifica moderna. Come si vedrà, Canali svolge sinergicamente l'analisi di questioni relative all'acquisizione delle conoscenze biomediche e degli orientamenti di fondo delle pratiche di cura, arrivando a delineare la possibilità di una nuova sintesi nei più recenti sviluppi della cosiddetta "medicina evolucionistica e genomica". I due successivi contributi dell'inserito – rispettivamente a firma di Roberto Festa, Vincenzo Crupi e Pierdaniele Giaretta, e di Vincenzo Crupi e Roberto Festa – riflettono una più consueta ripartizione tematica. Il primo (*Deduzione, induzione e abduzione nelle scienze mediche*) tratta di alcuni dei più basilari principi di inferenza alla base del ragionamento scientifico, mettendoli in relazione a episodi illustrativi della storia della ricerca medica moderna e contemporanea. Il secondo (*Dai giochi d'azzardo ai test diagnostici: la teoria della decisione nella medicina clinica*) discute invece i fondamenti teorici del ragionamento clinico, concentrandosi sui principi e su alcune applicazioni rilevanti della teoria della scelta razionale.

RIFERIMENTI BIBLIOGRAFICI

- BOVENS, L. e HARTMANN, S. (2003): *Bayesian Epistemology*, Oxford: Oxford University Press.
- CAMPANER, R. e FESTA, R. (2005a) (a cura di): *Incertezza e metodo in medicina (I). La ricerca medica*, "Nuova Civiltà delle Macchine", 23, n. 3.
- (2005b) (a cura di): *Incertezza e metodo in medicina (II). La pratica clinica*, "Nuova Civiltà delle Macchine", 23, n. 4.
- CORBELLINI, G. (2003): "Filosofia della medicina", in N. Vassallo (a cura di), *Filosofie delle scienze*, Torino: Einaudi, pp. 213-248.
- FESTA, R. (in corso di stampa), "Epistemologia bayesiana della testimonianza e sue applicazioni nella pratica clinica e giudiziaria".
- GIARETTA, P. e FEDERSPIL, G. (1998): *Il procedimento clinico. Analisi logica di una diagnosi*, Padova: Piccin.

- SCHAFFNER, K. F. (1985) (a cura di): *Logic of Discovery and Diagnosis in Medicine*, Berkeley: University of California Press.
- (1993): *Discovery and Explanation in Biology and Medicine*, Chicago (Ill.): The University of Chicago Press.
- THAGARD, P. (1999): *How Scientists Explain Disease*, Princeton (N.J.): Princeton University Press (trad. it. *La spiegazione scientifica della malattia*, Milano: McGraw-Hill, 2001).
- VINEIS, P. (1990): *Modelli di rischio*, Torino: Einaudi.

Sviluppi, problemi e prospettive della medicina scientifica

Stefano Canali

Laboratorio interdisciplinare per le scienze naturali e umanistiche,
SISSA (Trieste)
e-mail: canali@sissa.it

1. Medicina scientifica: un'analisi introduttiva del concetto
2. Caratteristiche della medicina scientifica
3. Dalla medicina scientifica alla medicina evolucionistica e genetica
4. Dalla medicina evolucionistica alla medicina genomica: verso una nuova logica della malattia

SOMMARIO. Nel presente saggio si fornisce un'analisi della medicina come disciplina scientifica. Tenendo conto dei principali aspetti problematici di questa nozione, si presenta una ricostruzione dei presupposti filosofici della medicina scientifica nel suo sviluppo storico, con particolare attenzione al ruolo fondativo dell'orientamento meccanicistico e riduzionistico. Si discute inoltre del possibile impatto della medicina evolucionistica e della genomica sulla struttura e sui contenuti della medicina scientifica contemporanea sia dal punto di vista conoscitivo sia sul versante dell'attività clinica.

PAROLE CHIAVE: medicina scientifica, medicina evolucionistica, medicina genetica, medicina genomica, logica della malattia.

1. Medicina scientifica: un'analisi introduttiva del concetto

La nozione di medicina scientifica è assai popolare. Sfortunatamente l'uso di questo concetto è alquanto vago e risultano scarsi i tentativi di specificarne più chiaramente il significato. Il termine "medicina" è infatti storicamente relativo, oltre che indeterminato e polisemico. L'aggettivo "scientifico", d'altronde, dovrebbe fare riferimento a qualcosa, la scienza, che non è un oggetto stabile e monolitico, bensì un variegato complesso, storicamente mutevole, di forme

di conoscenza disciplinari caratterizzate da metodologie e oggetti di indagine molto diversi.

Riferendoci all'accezione di medicina scientifica storicamente delineatasi dalla seconda metà dell'Ottocento, che descrive sostanzialmente il senso che oggi attribuiamo a questa espressione, con "medicina scientifica" intendiamo la medicina fondata sulla conoscenza, sul metodo e sugli strumenti delle scienze naturali e in particolare sulle teorie, le spiegazioni, le tecniche, gli apparati di indagine delle scienze biologiche e sull'uso in clinica di tecnologie e prassi derivate dalle conoscenze delle scienze naturali. È questo il senso cui Georges Canguilhem (1988) alludeva quando suggeriva di guardare alla medicina come "scienza applicata o somma evolutiva di scienze applicate". Allo stesso modo Donald Seldin (1981) definisce la medicina scientifica come biologia applicata alla cura e alla prevenzione delle malattie.

Quest'idea di una medicina che applica pezzi di conoscenze di (e metodi tratti da) altre scienze discende dalla formazione medica stessa. Secondo Uexküll e Wesiack (1991), infatti, i futuri medici apprendono le nozioni sulle strutture e i processi del corpo umano ben prima di entrare in contatto con i pazienti, ossia prima di sperimentare i problemi dell'incontro clinico e della gestione della cura, che sono invece lo scopo ultimo della medicina in quanto tale (Pellegrino e Thomasma 1981). Ciò impedisce loro di interpretare in senso pragmatico il dominio teoretico della biomedicina, base concettuale del loro lavoro. Non solo: spesso li porta a non comprendere quanto il ragionamento e l'intervento in clinica siano "carichi, anzi saturi di teorie" – cioè condizionati da metodologie, epistemologie e persino metafisiche di scienze non mediche, che tuttavia alimentano e indirizzano l'evoluzione della medicina scientifica.

Viene così in luce un elemento di criticità presente in ogni definizione: la natura intrinsecamente duplice della medicina, in quanto scienza e in quanto pratica – duplicità che inevitabilmente genera tensioni tra finalità terapeutiche e teorie esplicative e si riflette sul carattere composito e problematico della sua epistemologia.

Nella pratica e nella ricerca, la medicina pone costantemente a contatto la conoscenza scientifica con la dimensione dei valori (normale/patologico; sano/malato; malessere/benessere ecc.); tenta di definire un campo e un metodo per la risoluzione della tensione fra astratto e concreto, tra conoscenza universale e comprensione dell'individuale; tenta di combinare la ricerca sperimentale a quella del benessere. La medicina scientifica è così percorsa da una costante tensione fra aspetti conoscitivi, teorie e finalità pratiche, attività curative. Questa reciproca codificazione di domini di significato e finalità si attualizza in clinica e quindi attraverso uno spazio multiforme, composito e denso di di-

namiche soggettive e interpersonali. Per di più, il distintivo carattere strumentale e finalizzato della conoscenza scientifica in medicina, la cura del paziente, un soggetto comunque capace di agire, fa sì che in essa non esista soltanto, come ha evidenziato Engelhardt (1990, p. 70), un rapporto fra *explanans* ed *explanandum*, ma anche una cospicua interazione fra *manipulans* e *manipulandum*.

Tutto ciò rende la medicina scientifica estremamente vulnerabile alle critiche epistemologiche e allo stesso modo esposta ad appelli irrazionali. Norbert Paul (1998) ha parlato di “*Hiatus theoreticus*”, per indicare la vaghezza e l’insufficienza delle riflessioni sui peculiari e intrinseci rapporti fra teorie scientifiche e attività clinica che contraddistinguono la medicina scientifica. Esisterebbe cioè un gap epistemologico tra insieme delle conoscenze biomediche da un lato e loro applicazione nel giudizio e nella pratica clinica dall’altro. Questo vuoto concettuale sarebbe all’origine di molte debolezze della medicina scientifica, come le croniche difficoltà nel conciliare, sul piano pratico ma anche a livello pedagogico, gli aspetti biologici con le dimensioni psicosociali della malattia.

Ora però, dato che il rapporto tra conoscenze delle scienze biomediche e azione clinica costituisce l’architettura logica della medicina scientifica, questa inconsistenza epistemologica si estenderebbe all’interno delle stesse conoscenze biomediche, generando incongruenze e opacità anche nei rapporti teorici fra le diverse discipline scientifiche di riferimento. Dunque, la natura e lo statuto epistemologico della medicina scientifica sarebbero controversi anche restando sul solo versante teoretico, mettendo cioè fra parentesi il rapporto e la tensione fra conoscenza e pratica medica. La base conoscitiva della medicina scientifica risulta assai variegata per la natura multiforme delle discipline su cui poggia. Si pensi alla diversità di metodi e matrici disciplinari che esiste, per esempio, tra la fisica e l’anatomia o tra la matematica e la biologia cellulare ovvero ancora tra la statistica e la fisiologia. Per non parlare della complessità del rapporto fra dizionari concettuali, leggi e regole di inferenza propri di ogni livello di spiegazione della conoscenza biomedica del fatto biologico normale o morboso, da quello molecolare a quello fisiopatologico.

D’altra parte, come ha fatto notare Herbert Weiner (1977), il modello biomedico non costituisce un sistema concettuale unico o unitario ma rappresenta piuttosto una pluralità di modelli di concettualizzazione e pratica talora teoricamente giustapposti o, più comunemente, usati in maniera indipendente: il paradigma eziologico con cui si spiegano e si trattano le malattie infettive, il modello funzionale cellulare e patologico, il modello diagnostico e infine il modello curativo forte. Questi sistemi concettuali, secondo Weiner, sarebbero co-

munque tutti lineari, restrittivi e ipersemplicati in quanto incapaci di confrontarsi con la malattia come fenomeno biologico complesso ed esteso dalla dimensione molecolare a quella evolutivistica.

Nederbragt (2000) ha sostenuto che è possibile descrivere la complessità della struttura della conoscenza medica come un'interazione di fatti e teorie dalle principali discipline biomediche: fisiopatologia, epidemiologia, diagnostica, trial clinici, ricerche farmacologiche. Ognuna di queste discipline produrrebbe fatti e teorie basati su elementi probabilistici. Ma l'interrelazione di argomenti, fatti e spiegazioni da ognuna delle discipline biomediche, pure mediata dalla tecnologia e dalle innovazioni tecnologiche, produrrebbe un sistema conoscitivo, una coerente struttura di conoscenze accostabile a ciò che Thomas Kuhn (1962/1970) ha definito "paradigma". Nederbragt immagina che le interrelazioni concettuali fra le diverse discipline producano un corpo epistemico compatto, un sistema di significati congruo, unitario, secondo quanto suggerito dalle tesi dell'olismo epistemologico di Quine (1951). Tuttavia, a nostro parere, questo non sembra il caso della medicina.

Nel sistema di conoscenze della medicina scientifica opererebbero almeno due epistemologie diverse. Si tratta di due modi di concettualizzare la conoscenza dei fatti biomedici – e quindi di pensare l'azione in medicina – che appaiono sostanzialmente inconciliabili all'interno del quadro generale di riferimenti metafisico ed epistemico che contraddistingue la medicina scientifica di oggi. Da un lato, nella fisiopatologia e nelle sperimentazioni farmacologiche, si avrà l'uso di modelli di spiegazione causali e metodologie sperimentali di ricerca, approcci fortemente riduzionistici. Dall'altro, nell'epidemiologia, nei trial clinici e nella diagnostica, le indagini utilizzerebbero soprattutto modelli di tipo statistico.

Negli approcci biosperimentali delle ricerche eziologia-fisiopatologiche e farmacologiche la focalizzazione cade sulla malattia come deviazione da una norma fisiologica. La strategia di ricerca è il modello sperimentale e l'obiettivo è l'individuazione della cause immediate o prossime del fatto patologico. L'approccio metodologico è il riduzionismo, la scomposizione dei fenomeni in gioco al fine di arrivare all'analisi dei singoli nessi causali. Questa natura epistemologica è evidentemente informata, subordinata e sostanzialmente confinata alle scienze di base.

Le scienze biomediche come l'epidemiologia e i trial clinici, caratterizzate da modelli di indagine di tipo statistico, spostano invece il fuoco della ricerca sulla salute e inquadrano una dimensione popolazionale. Esse escludono l'indagine sulla natura dei meccanismi causali e cercano di accertare se esistono associazioni o correlazioni tra fattori ambientali, comportamenti e patologie,

al fine di fornire indicazioni sui modi per ridurre i rischi di ammalarsi e promuovere la salute e la prevenzione delle malattie. Questo approccio si è dimostrato utile alle indagini sui fattori sociali della malattia, sulle variabili economiche, comportamentali e politiche che condizionano la salute a livello individuale e della popolazione.

Questo sostanziale dualismo epistemologico della medicina scientifica, come ha sottolineato Corbellini (2003), ha notevoli conseguenze teoriche e pratiche, in quanto implica due diverse concezioni della malattia e due differenti nozioni dell'idea di causalità. Esso finisce così per complicare la pedagogia medica e rende i medici poco attrezzati a percepire i numerosi elementi di continuità fra dimensione individuale e orizzonte populazionale delle malattie, tra sfera fisica e spazio psicosociale, e quindi limitatamente capaci di intervenire su di essi.

2. Caratteristiche della medicina scientifica

2.1. *Il presupposto metafisico del monismo materialistico*

La medicina scientifica è iscritta entro una prospettiva metafisica di tipo monistico. Secondo la medicina scientifica, cioè, il mondo, e conseguentemente l'insieme dei processi normali e patologici della vita, sono costituiti e dipendono da un'unica sostanza. Questa sostanza è la materia e le sue concomitanti manifestazioni di energia e delle forze. Dal punto di vista ontologico, questo monismo tende oggi a coincidere con il fisicalismo, con l'idea metafisica che tutto sia fondamentalmente fisico.¹

Il materialismo in medicina ha portato a privilegiare la focalizzazione dello sguardo sul corpo, sulla struttura organica e sui processi all'interno dell'organismo. Ciò ha reso più faticosa la percezione delle dinamiche e delle variabili di contesto, come i fattori sociali, culturali e psicologici, che contribuiscono a determinare la malattia ovvero anche a realizzare le finalità della pratica medica. Inoltre, anche per l'influenza della tradizione anatomo-clinica, la prospettiva materialistica nel corso del tempo ha spostato l'attenzione dai sintomi riportati ai segni oggettivabili, dal racconto al riscontro fisico, dall'esperienza soggettiva del paziente – imprecisa, volatile, distorta dalla sofferenza – alla descrizione obiettiva. In assenza di un'adeguata architettura epistemica e di rife-

¹ Per una definizione e una discussione generale sull'idea di fisicalismo si rimanda a Pettit (1993) oppure a Poland (1994), o ancora a Gillett e Loewer (2001).

rimenti teorici a partire dai quali concettualizzare positivamente le variabili psicosociali delle condizioni che interessano la medicina, questo approccio ha tuttavia favorito la disumanizzazione dell'intervento medico e determinato una progressiva divaricazione fra l'identità della malattia vissuta e quella della malattia indagata e trattata dalla medicina scientifica. Nel corso del tempo questa frattura ha di fatto portato all'emergere di due ontologie della malattia apparentemente prive di elementi di contatto.

Questo particolare orientamento è all'origine delle accuse più forti che oggi vengono mosse alla medicina scientifica e ha contribuito ad alimentare un irrazionale scetticismo e alla diffusione delle medicine alternative anche nei paesi occidentali.

La forte adesione al materialismo ha inoltre prodotto un'altra stridente incongruenza della medicina scientifica. Si tratta della sottovalutazione dell'effetto placebo, una delle più consolidate verità dell'esperienza clinica; il fatto cioè che il medico è un agente terapeutico, che la presenza attiva del medico fa parte integrante del trattamento e che certi modi di partecipare al fatto clinico o intervenire sulla malattia possono essere addirittura nocivi. Sebbene numerose indagini abbiano dimostrato la sua azione in clinica (si veda per esempio la *review* di Vallance 2006) e studi recenti, soprattutto di neuroimmagini (Diederich e Goetz 2008; Klosterhalfen e Enck 2008), inizino a delineare alcuni aspetti biologici di questo controverso prodotto dell'azione clinica e del contesto clinico, la medicina scientifica non contempla il placebo in clinica e non educa il medico a usare il placebo, a sollecitarne e indirizzarne l'eventuale azione. Ciò perché il modello biomedico sta tardando a innestare nel concetto di placebo le acquisizioni sui suoi correlati neurobiologici. Per questa ragione l'idea di placebo è ancora confinata al dominio dei fatti psicologici, emotivi, fenomeni inaccessibili alla quantificazione e allo studio oggettivo, e che quindi il modello biomedico tende a tenere fuori nel giudizio clinico e nell'azione in clinica come fattori confondenti.

Va inoltre osservato che il rilievo dato al corpo come oggetto del pensiero e della pratica medica rappresenta anche una delle forme con cui è stata espressa la cornice metafisica ed epistemologica dell'individualismo che da Aristotele in poi ha contraddistinto la gran parte degli orientamenti di studio di tipo scientifico, ma anche buona parte delle tradizioni culturali occidentali (Laurent 1994). In particolare va considerata l'influenza delle dottrine individualistiche legate al liberalismo politico che si affermano parallelamente alla costruzione della medicina moderna fra Settecento e Ottocento. La centralità morale, sociale e politica riconosciuta all'individuo si è riflessa nella tendenza della medicina scientifica a indirizzare la sua ricerca e la sua azione sul pa-

ziente singolo, trascurando la rete di variabili e cause operanti a livello dei processi sociali. Il rapporto che l'evoluzione della medicina ha intrattenuto con il mutamento della concettualizzazione dell'individuo e dei suoi rapporti con i piani fenomenici più vasti cui appartiene è testimoniato anche dal fatto che una sensibilità verso i fattori sociali correlati ai fenomeni di interesse medico si è risvegliata dalla metà dell'Ottocento in concomitanza con l'affermazione di sistemi di pensiero di natura olistica, come l'hegelismo e le dottrine politiche di impronta socialista.

2.2. *Il meccanicismo e la causalità medica*

Il meccanicismo rappresenta l'altra dimensione caratteristica del monismo cui aderisce la medicina scientifica. Questa prospettiva considera la materia e il moto come gli elementi a partire dai quali, nell'applicazione delle leggi meccaniche del moto, è possibile comprendere e determinare tutti i fenomeni del mondo. Il meccanicismo ammette esclusivamente l'esistenza di cause efficienti, la materia e il moto, e rifiuta così qualsiasi tipo di finalismo. La filosofia meccanicistica è modellata sul concetto di meccanismo. Questo a sua volta indica elementi materiali e le relazioni tra essi: un sistema di parti che operano o interagiscono come quelle di una macchina, trasmettendo al suo interno moto, forze ed energia (Thagard 1999, p. 106). Più specificamente, un meccanismo potrebbe essere considerato un'entità fatta di componenti interagenti o un processo organizzato in modo tale da produrre regolarmente un determinato cambiamento da una data condizione iniziale a una precisa condizione finale. In questo senso, il meccanicismo implica il determinismo ed è associato a una concezione analitica e riduzionistica della realtà.

La rappresentazione meccanica dei fenomeni dell'universo medico è ovviamente mutata in parallelo all'evoluzione degli apparati tecnologici che via via hanno rappresentato modelli di riferimento e di raffigurazione del corpo umano e delle sue funzioni. Ma in generale, l'assunzione della prospettiva meccanicistica da parte della medicina scientifica ha fatto emergere una posizione metafisica in cui il paziente, in analogia a una macchina, è un insieme di parti giustapposte a produrre specifiche funzioni. Il meccanicismo avrebbe così assecondato una ridefinizione del corpo umano come corpo frammentato.

Sebbene la pratica clinica dimostri che la causalità biomedica è raramente deterministica e generalmente probabilistica (Giere *et al.* 2006), la metafisica collettiva condivisa che fa da scenario all'educazione medica trasmette soprattutto l'idea che le condizioni morbose dipendano da cause singolari e spe-

cifiche (Whitbeck 1977). Ancorché dipendenti da una rete causale complessa (Mackie 1973), le malattie vengono in larga parte definite a partire dall'identificazione di una singola causa (Wulff 1999). Idealmente, il modello di causalità della medicina scientifica tende quindi a modellarsi su una visione meccanica, per la quale gli agenti eziologici sono tali solo se riconducibili a determinismi di natura fisico-chimica particolari, se operanti in quanto agenti eziologici necessari, prossimi nel tempo (non cause remote, quindi) e nello spazio (contigui o a diretto contatto con il corpo e i suoi organi interni). Ciò dipende in larga parte da fattori storici e può spiegarsi appellandosi al ruolo avuto dagli ideali forti del determinismo e del meccanicismo nel promuovere l'evoluzione della concezione scientifica della medicina dalla Rivoluzione scientifica sino alla definizione del modello biomedico nel secondo Ottocento.

La concezione della causalità nel meccanicismo – soprattutto nelle versioni che ne diedero Hobbes e Cartesio, come ciò che produce o elimina un certo stato di cose in un corpo, azione determinante, movimento che dà ragione di un effetto – ha radicato la tendenza del pensiero medico a concepire la malattia come effetto di una singola causa prossima ed efficiente. Questa prospettiva monocausale ha ostacolato nel tempo l'emergere di approcci esplicativi più articolati e meglio capaci di contemplare il complesso dei fattori determinanti e delle dinamiche delle condizioni morbose.

La spiegazione meccanica del mondo, fra gli altri, assegna alla causalità lo status di rapporto di deducibilità fra due stati di cose, un'idea estremizzata da Pierre Simon de Laplace (1749-1827). Questi aveva affermato che se un'intelligenza potesse conoscere tutte le forze della natura e la situazione di tutti i corpi che la compongono, allora essa potrebbe dedurre tutte le condizioni e i movimenti dell'universo del futuro e del passato (Laplace 1878-1912). È evidente che per una disciplina come la medicina, in cui le esigenze operative si affiancano alle ragioni conoscitive, il principio della previsione infallibile contemplato negli sviluppi del meccanicismo rappresenta un'aspirazione ideale ancora più sentita che per le scienze pure. Per questa ragione, mezzo secolo più tardi Claude Bernard, il padre della medicina sperimentale, affermava l'inderogabilità del determinismo assoluto dei fenomeni per la fondazione della medicina scientifica.

L'ideale della ricerca dei meccanismi, delle cause prossime, è stato fittamente incrociato a una concezione della riducibilità dei fattori macrocausali a fattori microcausali e a un'idea della specificità delle cause. Queste ultime due forme della causalità sono state esaltate dai trionfi conoscitivi e terapeutici della microbiologia medica, una disciplina biomedica che si afferma con le clamorose scoperte di Louis Pasteur e di Robert Koch, fra il 1860 e il 1880. La de-

scrizione dei contributi e del rilievo epistemologico delle loro ricerche e scoperte richiederebbe un saggio a parte. È comunque opportuno indicare sinteticamente alcuni dei tratti dottrinali della rivoluzione microbiologica, in quanto avranno un rilievo notevole nell'indirizzare l'evoluzione teorica e pratica del modello biomedico della medicina scientifica. Il modello batteriologico introdotto da Pasteur e Koch può essere descritto a partire da questi principi: 1) la specificità di una malattia viene definita dalla presenza di uno specifico agente infettivo e non dai sintomi; 2) la presenza dell'agente infettivo è la causa necessaria della malattia; 3) la malattia si definisce quindi attraverso la sua causa.

Con i suoi straordinari successi terapeutici, il modello microbiologico dimostrava esemplarmente in medicina la potenza euristica e l'efficacia clinica dell'idea della causalità necessaria e della specificità. Questo indirizzo epistemologico veniva ulteriormente rafforzato dall'avvio della chemioterapia antinfettiva. Nel 1891, Paul Ehrlich (1845-1915) dimostrava la capacità di certe sostanze farmacologiche di interferire selettivamente sui processi vitali dei microbi patogeni, senza cioè interessare le funzioni dell'organismo umano. In questo caso, la specificità delle cause veniva finalmente a corrispondere con la selettività dell'azione di un farmaco. Ehrlich stesso proponeva la metafora del "proiettile magico" per rappresentare tale formidabile proprietà. Come sostiene Golub (1997), da allora la medicina ha ragionato soprattutto nei termini del proiettile magico, invocando cause singole per ogni malattia e coltivando così soprattutto il modello terapeutico forte, l'ideale della cura, dell'intervento specifico. Ma le malattie non hanno mai una singola causa. La dimensione eziologica delle condizioni morbose è complessa perché è una manifestazione della complessità della vita.

2.3. Il meccanicismo, le tecnologie biomediche e l'idea di progresso

Una delle ragioni e allo stesso tempo degli effetti della diffusione della concezione del corpo come macchina è stato lo sviluppo della tecnica applicata all'indagine e all'intervento medico: la tecnologia medica. L'evoluzione delle macchine e delle tecnologie non solo ha fornito una serie di diversi modelli per la concettualizzazione del corpo umano e delle sue funzioni normali e patologiche, ma ha anche modellato la percezione del corpo, del paziente e gli approcci in clinica sulla base delle diverse modalità con cui le varie tecnologie biomediche si interfacciano alle strutture e alla funzioni somatiche.

Dal punto di vista conoscitivo, le tecnologie biomediche hanno reso possibili le indagini di tipo obiettivo e quantitativo impossibili all'esplorazione del corpo

condotta con i soli cinque sensi. Nel regno della tecnologia medica, questo processo tende a isolare il corpo del paziente, ad astrarlo dalla sua vita e dalle sue relazioni, favorisce cioè l'alienazione del corpo del paziente da se stesso in quanto persona: lo straniamento del corpo (Marcum 2008). La tecnologia biomedica ha inoltre reso il corpo trasparente, ha moltiplicato a dismisura gli elementi strutturali e funzionali da contemplare nell'esame clinico. Il corpo trasparente è quindi anche un corpo tecnologicamente complicato (Van Dijck 2005).

Come ha affermato McWhinney (1978), un tema costante della tecnologia medica è la tendenza per la medicina a essere dominata dal valore meccanicistico dell'obiettività, della precisione e della standardizzazione. Questo approccio spinge a percepire il paziente a partire dal corpo standardizzato, che è il corpo generico, tipizzato, a cui il corpo del paziente in quanto collezione di dati clinici viene comparato. In maniera apparentemente paradossale, come suggerisce Frank (2002), oltre a subire questa colonizzazione tecnologica e questa espropriazione del corpo da parte del medico, il paziente viene anche disincarnato. La persona nel corpo del paziente è espulsa, bandita, estromessa, confinata ad assistere all'atto medico e a contemplare la sua rappresentazione nei dati tecnologicamente accertati, in modo non dissimile da quello di qualunque altra persona abbia modo di seguire la sua vicenda clinica.

Il meccanicismo e l'uso della tecnologia in medicina ha, inoltre, matematizzato l'espressione e il racconto delle condizioni fisiologiche e patologiche. In questo modo il racconto del sintomo, inquinato di soggettività e di significati confondenti, è stato relegato sullo sfondo della rappresentazione numerica e quantitativa, più precisa, neutra e imparziale, meglio rispondente alle forme e alle relazioni concettuali della matrice epistemologica della nuova scienza del vivente.

Dalla metà dell'Ottocento il regno della tecnologia è stato poi associato all'idea di progresso del positivismo, che improntava il pensiero filosofico e scientifico del tempo. La medicina scientifica che si delineava in quel particolare frangente storico-culturale ha così incorporato l'idea di progresso di matrice positivista. Anche per questa ragione, l'immagine prevalente della medicina scientifica è stata quella di un'impresa scientifica e tecnica in continuo progresso verso l'acquisizione di nuove conoscenze e verso la messa a punto di operazioni terapeutiche sempre più efficaci e sempre più capaci di intervenire sulle malattie, di prevenire le condizioni patologiche, di garantire la salute e allontanare la morte. Per ironia della sorte, questa immagine si è fortemente rafforzata nella seconda metà del Novecento, dopo la rivoluzione terapeutica, nel momento in cui il progresso dell'efficacia della medicina ha iniziato a rallentare. Le meravigliose capacità progressive della medicina scientifica contempo-

reana sono state così alimentate soprattutto dall'espansione della tecnica nella pratica medica e del ruolo delle scienze di base nella costruzione della conoscenza medica a dispetto della loro limitata capacità di far fronte all'espansione delle patologie cronico-degenerative, l'aspetto più preoccupante della transizione epidemiologica che si è realizzata nel Novecento. In questi ultimi anni la discrepanza tra l'efficacia attesa della medicina e il profilarsi di fenomeni morbosi nuovi, epidemici e relativamente controllabili è emersa con chiarezza, sino a innescare una profonda crisi di fiducia verso la medicina scientifica.

D'altra parte, questa idea originaria e inarticolata di progresso e la sua caratterizzazione come progresso tecnologico hanno condizionato evidentemente la percezione dell'opportunità e dei limiti dell'azione non solo sulle dimensioni patologiche ma anche sulle condizioni umane normali e ineluttabili come la senescenza e la morte. Gli arti possono essere reimpiantati, gli organi trapiantati, la vecchiaia rallentata, la morte prevenuta o sospesa indefinitamente con apparati di rianimazione. Così, due delle più forti radici filosofiche della medicina scientifica condizionano la comprensione del valore della pratica medica, dell'esercizio della medicina, e rendono complessa la valutazione dei limiti dell'azione terapeutica ai medici e alla pubblica opinione. A testimonianza di ciò si possono rilevare la considerazione che viene riservata alla chirurgia – medicina di intervento per eccellenza – e la sfiducia con cui si guarda al medico di famiglia, il quale lavora soprattutto al controllo e alla prevenzione delle malattie.

2.4. *Il riduzionismo*

L'aspetto compositivo che contraddistingue una delle dimensioni del concetto di macchina intorno a cui si è sviluppata l'idea di meccanicismo (Glennan 1998; Machamer *et al.* 2000) è correlato a un altro presupposto metafisico peculiare della medicina scientifica: il riduzionismo (Nagel 1961; Sarkar 1992).

Il riduzionismo può essere ontologico, teoretico o metodologico. Nel primo caso viene supposto che la realtà di un fenomeno possa ridursi alla somma delle proprietà delle sue parti, ovvero che un fenomeno di ordine più elevato sia interamente determinato da un fenomeno o ordine meno elevato (Dupré 2000), e cioè che non esistano proprietà nuove ed emergenti per i fenomeni di livello superiore (Marcum e Verschuuren 1986). Questo tipo di riduzionismo impone alla medicina scientifica di vedere i processi funzionali e morbosi del corpo umano come espressione di una serie di dinamiche di livello inferiore. Se la malattia interessa cioè un certo apparato funzionale, le cause vere agiscono a un livello più

basso, a livello cellulare o a livello molecolare, a partire da un qualche difetto strutturale e funzionale di un network genetico o di un singolo gene.

Il riduzionismo di tipo teoretico prevede la riduzione di termini e concetti di determinate teorie ai termini e ai concetti di una singola teoria o più teorie di base. Ciò è esemplificato dalla tendenza – o dalla tentazione – della medicina scientifica attuale di ridurre le idee della patologia e della fisiopatologia alla biochimica e alla genetica molecolare.

Il riduzionismo di tipo metodologico si risolve invece nella definizione di strategie di ricerca volte a scomporre i fenomeni complessi e di ordine superiore nei loro componenti costitutivi di ordine inferiore al fine di ricostruire, dopo le singole indagini, la rappresentazione della dinamica complessa. Per esempio, la ricerca sui processi fisiopatologici viene condotta separatamente per ogni livello funzionale implicato, da quello biochimico a quello molecolare, nel tentativo di determinarne le componenti individuali. Quindi i processi elementari così individuati vengono interconnessi per elaborare la rappresentazione della dinamica fisiopatologica complessiva. Questa strategia ha garantito la possibilità di indagine e descrizione delle proprietà funzionali dipendenti dall'organizzazione, le caratteristiche espressioni delle attività dei sistemi viventi, come l'autoregolazione, l'adattabilità, i meccanismi omeostatici, tendenzialmente trascurate dagli approcci riduzionistici più radicali.

Questo approccio si delinea con le ricerche di Claude Bernard e quindi segna tutta la storia del modello biomedico. Bernard riconosceva la produttività del metodo analitico e riduzionistico. L'analisi e la scomposizione sono fondamentali perché, scriveva Bernard:

Come in condizioni normali la vita si manifesta attraverso l'attività normale degli elementi organizzati, così le malattie si rivelano con la manifestazione anormale degli stessi elementi; la terapia poi può agire su questi elementi attraverso l'ambiente interno variamente modificato da certe sostanze tossiche o medicamentose. Per risolvere questi diversi problemi bisogna smontare per così dire l'organismo pezzo per pezzo, come si smonta una macchina quando si voglia conoscerne e studiarne tutti gli ingranaggi; in altre parole, prima di arrivare all'esame delle parti elementari bisogna sperimentare sugli organi e sugli apparati. Lo studio dei fenomeni biologici deve essere perciò analitico e graduale e fatto con lo stesso metodo sperimentale che il fisico e il chimico impiegano nell'analisi dei fenomeni dei corpi inanimati. Le difficoltà che compaiono quando si applica il metodo sperimentale ai fenomeni biologici sono dovute esclusivamente alla complessità dei fenomeni stessi; il fine e i principi del metodo non cambiano (Bernard 1865, trad. it. p. 77).

Le caratterizzazioni ontologica e teoretica del riduzionismo sono state espresse storicamente in una sequenza di successive riduzioni delle spiegazioni mediche alle matrici epistemiche, agli elementi e ai principi esplicativi via via emergenti nelle scienze della vita. Questo percorso è esemplarmente rappresentato dall'evoluzione dell'idea di malattia a partire dalla sintesi anatomico-clinica. L'anatomia patologica aveva isolato la malattia nel singolo organo, nella lesione localizzata. Con la patologia tissulare di Xavier Bichat (1771-1802) la malattia è nei tessuti, nelle membrane, va colta, indagata e affrontata a quel livello di organizzazione più fondamentale (Bichat 1799; 1800). Con la patologia cellulare, formulata da Rudolph Virchow intorno alla metà dell'Ottocento, la strategia riduzionistica in medicina faceva emergere un nuovo orizzonte concettuale e un nuovo obiettivo di intervento: la cellula (Virchow 1855; 1858). Solo un paio di decenni più tardi, gli straordinari successi esplicativi e terapeutici della batteriologia di Pasteur e Koch determineranno l'affermazione dell'idea della malattia come microbo, come agente eziologico microscopico o malattia esso stesso.

Nei primi decenni del Novecento, la diffusione e il miglioramento delle rilevazioni chimiche e sierologiche mise in luce una serie di nuove malattie caratterizzate dalla deviazione di una particolarità chimica dell'organismo. Da qui si diffuse l'idea che tutte le malattie possano ridursi a una qualche deviazione chimica, siano cioè *lesioni biochimiche* (Peters 1936; 1969). L'avanzamento nella comprensione dei cicli metabolici e delle funzioni degli enzimi e le nuove acquisizioni sui substrati biochimici portavano quindi all'idea di lesione molecolare (Schade 1935) e successivamente alla definizione del concetto di malattia molecolare, quando Linus Pauling (1901-94) dimostrava con analisi elettroforetiche la diversità dell'emoglobina dei soggetti affetti da anemia falciforme da quella delle persone sane (Pauling *et al.* 1949). Il concetto di malattia molecolare era stato formulato entro una prospettiva biochimica e a livello del comportamento delle proteine. Nella seconda metà del Novecento, a seguito della scoperta del DNA, la caratterizzazione molecolare della malattia si è spostata dalle proteine agli acidi nucleici, al DNA come codice che incorpora le informazioni necessarie alla costruzione delle proteine dagli acidi nucleici.

A partire dagli scorsi anni Settanta, la messa a punto e lo sviluppo delle tecnologie del DNA ricombinante, come le endonucleasi di restrizione² e, dagli anni Ottanta, la reazione a catena della polimerasi,³ hanno determinato la ri-

² Sono enzimi che scindono il DNA in punti specifici e consentono di isolare dei frammenti più corti e resistenti facilitando l'identificazione del gene in esame.

³ La reazione a catena della polimerasi (*Polimerase Chain Reaction*), scoperta da Kary Mullis nel 1983, permette la generazione di un enorme numero di copie di una specifica sequenza di DNA.

voluzione molecolare in biologia e genetica e portato alla caratterizzazione delle basi genetiche di diverse centinaia di malattie. Valorizzati entro la dominante prospettiva metafisica del materialismo e all'interno del quadro epistemico del riduzionismo, questi risultati hanno finito per riformulare in profondità lo stesso significato del DNA e delle sue relazioni con gli ambiti di spiegazione biomedica. Gli avanzamenti del Progetto genoma umano nell'ultima decade del Novecento hanno sostenuto questa trasformazione concettuale. Ne è sorta una nuova e radicale ontologia, che taluni hanno definito "essenzialismo genetico" (Nelkin e Lindee 1995), nella quale il DNA è visto come una realtà relativamente indipendente dal corpo, che dà all'organismo forma, vita e funzioni (normali o patologiche che siano): una realtà che determina l'identità stessa di un essere umano.

Tutto ciò ha contribuito ad alimentare il sogno di ricondurre le condizioni morbose a un'entità fisicamente determinata, il gene o addirittura il singolo nucleotide. È stata così prefigurata l'imminente riduzione delle scienze cliniche alla genetica molecolare (Guttmacher e Collina 2002), trasformata in genomica dai rapidi progressi delle tecnologie di laboratorio. Allo stesso modo si è diffusa l'idea che il grado di risoluzione e scomposizione raggiunto dall'anatomia molecolare dei geni stesse ormai per dischiudere il piano fondamentale per la comprensione dei meccanismi e delle cause delle malattie (Plotz 1997). La continuità con gli indirizzi analitici, riduzionistici e monocausali che hanno caratterizzato la storia e le transizioni epistemiche della medicina scientifica da Vesalio in poi è stata peraltro esplicitamente richiamata da Victor McKusick (1921-2008), uno dei padri fondatori della genetica medica molecolare. Questi sosteneva che a partire dall'anatomia del genoma umano si potesse fondare la medicina neovesaliana del XXI secolo, più predittiva e preventiva di quella centrata sulla fisiopatologia (McKusick 2001). L'euforia della velocità con cui si sono bruciate le tappe del sequenziamento del genoma umano ha finito per spingere troppo in là lo sguardo riduzionistico, rendendolo cieco (Tauber e Sarkar 1992).

Questo breve resoconto dimostra peraltro come nell'evoluzione delle strategie analitiche, la prospettiva riduzionistica si sia tendenzialmente accompagnata alla visione monocausale e ontologica della malattia, all'idea cioè di poter spiegare la malattia come effetto di un solo "corpo" anomalo o in quanto corpo patologico *tout court*.

Negli ultimi anni, gli eccezionali risultati della genomica, in particolare le conoscenze ottenute con il completamento del sequenziamento del nostro genoma (Venter *et al.* 2001), stanno però avendo effetti imprevisti. Contrariamente alle aspettative, invece di sostenere l'idea forte del riduzionismo onto-

logico nelle scienze biomediche e la visione del determinismo genetico ad essa associata (Kitcher 2001; Oyama 2000), le nuove acquisizioni genomiche hanno messo in evidenza l'impossibilità di ridurre ai singoli geni le funzioni normali o morbose degli organismi (Fox Keller 2000). Anche per quanto riguarda i semplici disturbi monogenici, imputati cioè all'azione di un singolo gene (per esempio le talassemie), la genomica sta delineando dinamiche e orizzonti causali molteplici, interrelati e complessi (Chang *et al.* 1997; Wolf 1997). In questo modo, la ricerca genomica tende a prefigurare un modello di spiegazione delle malattie più fine e articolato (Summers 1996; Wheaterall 2000). Per riprendere il caso delle talassemie, la discriminazione sempre più minuta di manifestazioni cliniche diverse, il loro carattere singolare e la caratterizzazione molecolare dei fenotipi associati alle alterazioni molecolari di queste malattie genetiche del sangue indicano il coinvolgimento di numerosi fattori interconnessi e di diversi livelli causali, che vanno dalla dimensione puntiforme a livello dei nucleotidi alla relazione fra geni diversi, dalla storia evolutiva alle dinamiche popolazionali, sino ai fattori di tipo ecologico, etologico e culturale (Wheaterall 1999). In questo caso, la dissezione delle basi molecolari di uno dei più semplici e diffusi disturbi monogenici ha dimostrato conclusivamente la complessità della relazione causale tra fenotipo malato e genotipo che rende di fatto impossibile isolare un singolo fattore determinante come causa necessaria e sufficiente. La corsa alla riduzione dei fatti clinici alla genetica molecolare rende peraltro problematica la percezione dei piani di realtà e dei livelli di connessioni causali più alti, come i fattori ambientali, le abitudini e gli stili di vita, le dinamiche sociali e politiche e potrebbe attenuare gli sforzi sanitari tesi a contrastare o ridurre questi ambiti di realtà e di causazione della patologia, che appaiono comunque decisivi (Holtzman e Marteau 2000; Cooper e Psaty 2003).

3. Dalla medicina scientifica alla medicina evolutivistica e genetica

3.1. Medicina scientifica, medicina clinica e prassi medica

La medicina scientifica e il modello biomedico hanno valorizzato la dimensione delle scienze mediche di base e delle tecnoscienze ad esse collegate. L'enfasi è così caduta sulla conoscenza scientifica, sulla tensione all'oggettivazione, alla classificazione, all'astrazione, alla standardizzazione degli elementi rilevanti, alla ricerca delle cause e delle spiegazioni universali. Ma, ciò che è oggettivato non è unico (Pellegrino e Thomasma 1981, p. 109). Non esi-

ste una scienza dell'individuale e la medicina per questo soffre di una contraddizione fondamentale: la sua pratica ha a che fare con gli individui mentre la sua dimensione teorica rimanda a oggetti universali (Temkin 1963). Oggettivare un'entità come un essere vivente, un soggetto, una persona significa sacrificarne l'individualità per ricondurlo a schemi e tratti comuni, caratteristici. Gli imperativi epistemologici della medicina scientifica tendono inevitabilmente a entrare in conflitto con gli obblighi e gli scopi della medicina clinica e della concreta azione in medicina: la comprensione dei processi in corso nella realtà individuale del paziente in carne e ossa e la particolarizzazione delle conoscenze delle scienze di base e della medicina clinica nel contingente in-contro clinico.

Come ha sottolineato Nordin (1999), la conoscenza scientifica in quanto tale è insufficiente a formulare decisioni razionali in clinica, dato che queste decisioni richiedono anche la conoscenza del paziente come individuo, e la conoscenza dei suoi valori e dei suoi desideri, dell'ambiente sociale in cui si muove. Non esisterebbe quindi una connessione semplice e lineare tra le scienze di base e la loro applicazione in medicina. L'accento posto dalla medicina scientifica sulle conoscenze di base ha portato a trascurare il fatto che la medicina in quanto tale è un'attività che si concretizza nell'evento clinico. E l'evento clinico richiede che la conoscenza scientifica venga messa in relazione con la vivente realtà di un paziente particolare in vista del trattamento di una specifica e individuale condizione morbosa o del ripristino di condizioni giudicate funzionali per quel paziente e da quel paziente. La conoscenza scientifica in ambito clinico, in particolare quella confinata all'orizzonte fisiopatologico del modello biomedico, viene quindi inevitabilmente posta in una problematica correlazione con l'individualità della persona, con le sue continue trasformazioni psicofisiche, la sua soggettività, la sua storia e con il contesto di valori in cui quella specifica condizione morbosa si è determinata, manifestata e ha finito per essere trattata.

La difficoltà a concettualizzare l'individualità, la soggettività e il mondo dei valori ha contribuito alla crisi della medicina scientifica fondata sul modello biomedico che abbiamo sinora descritto. Ciò ha peraltro favorito la diffusione del ricorso alle cosiddette "medicines alternative" e il proliferare di appelli per nuovi modelli medici più articolati ovvero anche per una imprecisata e rarefatta medicina umanistica o umanizzata. Uno degli approcci più discussi – il cosiddetto "modello biopsicosociale" di George Engel (Engel 1977) – postula la necessità di tenere maggiormente in conto il paziente, la dimensione psicologica individuale entro cui egli vive la malattia, il contesto sociale. In questo senso Engel suggerisce di integrare la prospettiva psicosociale ai piani

molecolare e cellulare cui generalmente vengono riferite le condizioni patologiche. Engel tuttavia non chiarisce in quale modo è possibile realizzare questa integrazione all'interno di una cornice genuinamente biologica. Egli non è in grado di spiegare cioè i meccanismi con cui ogni dimensione fenomenica della malattia determina e condiziona le altre. Esplicitamente inscritto nella prospettiva olistica della teoria generale dei sistemi (Bertalanffy 1968), inoltre, il modello biopsicologico presenta tutti i limiti delle concezioni organicistiche, come l'incapacità di definire univocamente l'identità e i limiti del "sistema" paziente rispetto agli altri sistemi con cui interagisce e retroagisce. Paradossalmente, quindi, il modello biopsicosociale fatica a misurarsi proprio con la dimensione individuale che invece avrebbe dovuto finalmente valorizzare.

3.2. *La medicina evoluzionistica e genetica*

Abbiamo più volte definito la medicina scientifica come epistemologicamente fondata sulle scienze naturali e biologiche. Ma a ben vedere questo riferimento teoretico è quantomeno incompiuto. Da un lato, come abbiamo visto, la medicina è divenuta sempre più dipendente dallo studio oggettivo degli organi, dei tessuti, delle cellule e dei processi biochimici e quindi dai progressi delle scienze di base. Ma allo stesso tempo essa ha manifestato una certa resistenza ad assimilare alcune fondamentali definizioni ed innovazioni teoriche introdotte nelle scienze della vita, come le nuove possibilità di spiegare il comportamento e i fenomeni psicosociali offerte dalle neuroscienze e soprattutto la teoria dell'evoluzione, paradigma fondante delle discipline biologiche contemporanee.

Questo grave ritardo culturale della moderna medicina scientifica è in larga parte attribuibile a ragioni storiche contingenti. La sua fondazione teorica e metodologica è avvenuta prima che il pensiero evoluzionistico diventasse patrimonio comune e condiviso delle scienze della vita. Di conseguenza, la medicina scientifica si è dovuta forzatamente modellare sui postulati e sui metodi di una biologia attestata sull'approccio meccanicista e riduzionista con cui la biologia della prima metà dell'Ottocento affrontava i problemi dell'indagine sulle strutture e sulle funzioni degli organismi viventi.

Negli ultimi quindici anni, tuttavia, qualcosa sta cambiando. L'approccio darwiniano ed evoluzionistico si è progressivamente innestato nel dibattito sui metodi e sugli strumenti concettuali della ricerca biomedica, interessando altresì la riflessione sugli obiettivi e i criteri che istruiscono la prassi dell'azione sanitaria, l'intervento terapeutico e la formazione dei medici e del persona-

le sanitario. Pur variamente presente nel dibattito medico sin dall'epoca di Darwin (Corbellini 1998; 2004), l'approccio evoluzionistico alla comprensione delle malattie è diventato un tema di specifica riflessione e ricerca all'inizio degli anni Novanta del Novecento per merito soprattutto dei lavori dello psichiatra Randolph Nesse e del biologo George Williams (Nesse e Williams 1991; 1995) e ormai sistematizzato in alcune raccolte di saggi (Donghi 1998; Stearns 1999; Trevathan, Smith e McKenna 1999; Corbellini e Canali 2004), alcuni dedicati in particolare alla psichiatria (Stevens e Price 1996; Canali 2001; Canali e Pani 2003).

Il postulato che fa da sfondo al ragionamento evoluzionistico in medicina è che ogni individuo è espressione di un programma genetico e che tale programma rappresenta un prodotto storico e unico dell'evoluzione, plasmato dai meccanismi della filogenesi: variazione genetica e selezione naturale. Secondo la medicina evoluzionistica dunque i fenomeni epidemiologici, la specifica vulnerabilità individuale alle malattie, i modi e i tempi con cui ogni individuo risponde a un agente patogeno, si ammala o recupera la salute dipenderebbero anche da processi storici, filogenetici. Tale prospettiva suggerisce che una spiegazione adeguata delle malattie non può arrestarsi all'esame delle cause prossime che innescano il processo patogenetico, ma deve considerare l'azione di cause remote, cioè far ricorso a categorie esplicative e argomentazioni di tipo evoluzionistico.

L'applicazione in campo medico dell'approccio darwiniano apre un ventaglio nuovo di spiegazioni delle malattie. Seguendo Nesse e Williams (1995), queste sono raggruppabili in sette categorie:

- 1) *Difesa*. Ciò che noi riteniamo un sintomo patologico, talora è in realtà un meccanismo di adattamento, come nel caso della tosse, del vomito e della febbre.
- 2) *Conflitto con altri elementi in evoluzione*, per esempio i microrganismi patogeni o altri organismi che possono in qualche modo compromettere certe funzioni dell'organismo umano.
- 3) *Il disadattamento del nostro corpo alle veloci modificazioni dell'ambiente e dei modi di vita nell'epoca moderna*. I nostri sistemi funzionali sono sostanzialmente quelli che l'evoluzione umana aveva reso disponibili circa duecentomila anni fa e adatti a un ambiente estremamente diverso da quello in cui viviamo oggi.
- 4) *Compromessi evolutivi a livello genetico*. Un gene può conferire certi vantaggi in specifici contesti ambientali, ma aumentare la suscettibilità a sviluppare alcune patologie. Un caso è quello delle emoglobinopatie, come

l'anemia falciforme e la talassemia, che in condizione eterozigotica possono dare vantaggi selettivi in ambienti malarici ma espongono al pericolo di anemie più o meno gravi.

- 5) *Compromessi evolutivi a livello dei tratti fenotipici complessi.* Ogni tratto, somatico o comportamentale, di un individuo è l'espressione di un complesso equilibrio genetico ed epigenetico fra strutture somatiche e funzioni psicologiche. Certi geni fanno parte allo stesso tempo di sistemi biochimici che controllano processi diversi, per esempio lo sviluppo di un organo, la sintesi di un ormone, una particolare reazione enzimatica. Così, una mutazione genetica che aumenta l'efficienza di una funzione può compromettere l'efficacia di un'altra attività biologica ed esporre l'organismo all'insorgenza di particolari malattie.
- 6) *Vincoli storici e dipendenza da traiettorie evolutive.* L'evoluzione procede usando e coadattando il "vecchio" materiale biologico e psicologico della specie. Il migliore accomodamento fra i materiali e le funzioni biologiche preesistenti difficilmente coincide con la migliore e più efficace soluzione pensabile per una struttura funzionale.
- 7) *Fattori casuali.* Occorre infine ricordare, e siamo nella categoria di spiegazione evolutiva dei fattori casuali, che il processo evolutivo non segue un disegno preordinato di schemi di sviluppo tesi alla massima efficienza. Esso utilizza, adatta e rimodella apparati funzionali ed elementi anatomici già presenti (una sorta di bricolage biologico) e si svolge largamente per l'intervento di fattori casuali, variazioni genetiche, modificazioni ambientali, cambiamenti dei rapporti etologici e così via.

Ancora piuttosto trascurata nel dibattito sulla medicina darwiniana, la prospettiva storica dell'evoluzionismo contempla anche la dimensione culturale. Questo piano è fondamentale per comprendere l'origine delle malattie nell'uomo. Nell'*Homo sapiens*, l'evoluzione è quasi esclusivamente trasformazione culturale. Il mutamento e la produzione culturale sono centrali sull'affermazione della specie umana e sulla trasformazione del suo rapporto con l'ambiente e con le altre specie, anche in quanto potenziali patogeni. L'evoluzione culturale e i suoi prodotti hanno modulato la selezione naturale, finendo per essere in qualche modo rappresentati a livello genetico nelle popolazioni e negli individui, per avere così un ruolo significativo nella vulnerabilità verso talune patologie.

Altra interessante implicazione del pensiero evoluzionistico in medicina è la cosiddetta *life-history theory* (Stearns 1992; Roff 1992). Questa prospettiva concettuale sottolinea il fatto che la gran parte dei tratti fisiologici e comportamentali, siano essi normali o patologici, possono essere meglio compresi anche con-

siderandoli come effetti della selezione naturale sulle tappe e le caratteristiche che definiscono il corso della vita, dalla maturazione allo sviluppo, dalla riproduzione alla senescenza. La *life-history theory* si riferisce in generale ai fenomeni biologici che implicano la presenza e la manifestazione nell'arco della vita, di compromessi fra tratti biologici, come l'età programmata della maturazione sessuale, la durata dello sviluppo fetale e delle cure parentali, il tasso di senescenza e la durata media programmata della vita, le dimensioni e il peso corporeo alla maturità, il tasso di fertilità e così via. Per esempio, un animale che tarda la maturazione può raggiungere dimensioni maggiori e avere una maggiore e più durevole fertilità. In questo caso, la selezione avrebbe realizzato il compromesso tra rinuncia alla fertilità precoce e raggiungimento di dimensioni maggiori e fertilità prolungata. Molti tratti contro cui si rivolge l'azione medica sembrano originarsi dai compromessi sui quali ha gettato luce la *life-history theory*. Per esempio, la vulnerabilità degli adolescenti verso l'iperattività, i comportamenti a rischio, la ricerca del piacere e delle ricompense immediate è tratto che espone questa fascia d'età a un rischio elevato di morte violenta o di sviluppare disturbi del comportamento, come i comportamenti di abuso di sostanze. La *life-history theory* suggerisce, tuttavia, che questo tratto, ancorché potenzialmente patogeno o mortale, possiede in questa fascia d'età un valore riproduttivo tale da compensarne i forti rischi. La *life-history theory* aggiunge così una ulteriore e più fine articolazione temporale alla causalità medica rispetto a quella innestata dalla semplice prospettiva evuzionistica, in cui le cause remote comunque sono inquadrare nella dimensione generazionale. Di conseguenza essa garantisce strategie esplicative e di intervento meglio modellabili sul paziente, la cui malattia non racconta solo la storia della specie, ma anche il modo in cui questa si riflette e si dipana nell'arco della sua vita singola.

3.3. *Pensiero tipologico e pensiero popolazionale: dalla biologia alla medicina*

Uno degli aspetti maggiormente interessanti del darwinismo ai fini della comprensione delle malattie, eppure alquanto trascurato dai nuovi approcci evuzionistici alla medicina, è quello che Ernst Mayr ha chiamato “pensiero popolazionale” (Mayr 1982; 1988). Questa idea si fonda sull'osservazione dell'unicità propria degli organismi viventi, sul fatto che ogni individuo di una specie a riproduzione sessuata è unico e costitutivamente differente da tutti gli altri. In questo senso, le popolazioni – dal cui insieme risulta ogni specie vivente – sono gruppi di individui unici e differenti l'uno dall'altro.

Introdotta da Darwin, il pensiero popolazionale ha permesso il definitivo superamento del pensiero essenzialistico o tipologico, secondo cui le specie erano forme, essenze date, caratterizzate dalla costanza, dalla fissità e dalla discontinuità e per cui, conseguentemente, le variazioni tra gli individui erano viste semplicemente come discostamento dalla norma, deviazione dal progetto dell'essenza.

Il pensiero popolazionale prende le mosse dalla diversità, considera l'individuo come *primum* del processo evolutivo, pone l'accento sulla storicità dei fatti della vita. Secondo questo approccio non esistono quindi parametri biologici o individui tipici. I valori medi dei processi biologici sono soltanto astrazioni. Le differenze, le variazioni, costituiscono invece la realtà biologica e rappresentano l'aspetto veramente significativo delle popolazioni e di ogni fatto della vita. Sulle variazioni e sull'individualità può agire la selezione naturale e così dar luogo all'evoluzione, alla trasformazione filogenetica.

Consolidato da tempo nelle scienze biologiche, il pensiero popolazionale stenta ad affermarsi in medicina. La versione essenzialistica della malattia, anche chiamata ontologica, è quella su cui fondamentalmente si formano tuttora i medici. In essa il paziente, quando non irrilevante, è indifferenziato, ovvero solo l'occasionale mezzo attraverso cui prende corpo lo stato morboso. Quest'ultimo è la vera individualità, natura che tende a coincidere con l'identità della classe nosologica entro cui la condizione patologica è inquadrata. In questo approccio, la malattia ha cause univoche, il suo decorso naturale e la risposta al trattamento nei vari pazienti sono essenzialmente gli stessi. La visione tipologica in medicina è legata a una concezione formale, storica e ingegneristica del corpo e delle sue funzioni. Le fa da sfondo una visione meccanicistica che equipara i processi patologici al guasto in una macchina. Conseguentemente il concetto tipologico di malattia tende a enfatizzare la localizzazione del guasto e a focalizzare l'intervento terapeutico sul pezzo, l'organo danneggiato o difettoso. Poco importa come si è arrivati alla rottura o al malfunzionamento. Tale visione inoltre rimanda a un'idea forte della cura quale processo che deve e può portare al ripristino delle condizioni ottimali di funzionamento nella macchina guasta e quindi quale intervento indipendente dalla storia della macchina, dalle sue condizioni attuali.

L'essenzialismo in medicina infine contempla le malattie in quanto entità distinte e la discontinuità fra la salute e gli stati patologici, rendendo difficile la spiegazione del passaggio dall'una all'altra condizione e scarsamente significativi tutti i processi e gli eventi non immediatamente precedenti le manifestazioni acute.

Trasposto alla medicina, invece, il pensiero popolazionale mette in rilievo il carattere radicalmente particolare dei processi con cui ogni persona intera-

gisce con l'ambiente e con gli stimoli patogeni, ovvero l'individualità della suscettibilità alle malattie e della maniera in cui esse si esprimono nei singoli pazienti. L'approccio popolazionale così richiama l'attenzione sul carattere individuale della malattia: un aspetto duplice, radicato allo stesso tempo sull'unicità del genotipo e della storia di ogni individuo, di come questa si riflette sulle funzioni stesse del genotipo.

Secondo tale visione quindi la malattia rappresenta l'esito singolare di inadeguate risposte adattative dell'individuo, ovvero l'effetto di qualche incongruenza nell'espressione delle dimensioni genetica, epigenetica o dello sviluppo, psicologico e sociale, che caratterizzano l'individualità di una persona.

L'applicazione del pensiero popolazionale quindi permette alla medicina di tematizzare in termini positivi l'unicità e la storicità dell'individuo, la singolarità radicale delle condizioni morbose cui può andare incontro nel corso della vita, frutto dell'incontro tra la storia filogenetica (espressa nel terreno di vulnerabilità) e la storia delle esperienze dell'organismo. In questo senso, il pensiero popolazionale ed evolutivistico indica la necessità di un approccio terapeutico fortemente centrato sul paziente.

4. Dalla medicina evolutivistica alla medicina genomica: verso una nuova logica della malattia

L'evoluzione della medicina evolutivistica sembra potenzialmente in grado di delineare una nuova logica della malattia capace di spiegare in termini genuinamente biologici il ruolo della dimensione socio-culturale nella vulnerabilità alle malattie, nella patogenesi e nel decorso delle condizioni morbose. Questa nuova logica, secondo Charles Sriver (1982; 1984) e Barton Childs (1999), può emergere dalla sintesi della nuova genetica con il pensiero evolutivistico e il classico approccio fisiopatologico alla comprensione della malattia: una struttura concettuale in grado di dar forma a un'idea della malattia fondata sul gene e sulle sue relazioni con l'ontogenesi, la filogenesi, la cultura, la società, allo stesso modo che con la selezione naturale e l'evoluzione biologica. Questa logica presenterebbe una serie di evidenti vantaggi teorici:

- 1) Essa può costituire un unico modello concettuale per la spiegazione e la ricerca delle cause e della fisiopatologia di tutte le malattie, siano esse genetiche o acquisite.
- 2) In questa logica le malattie sono viste come processi che si originano nelle incongruenze, nella dissonanza fra le tappe dei processi fisiologici tesi al-

l'omeostasi dell'organismo, dalle funzioni geniche al comportamento. Ognuna di queste tappe prende origine da un gene, che è a sua volta prodotto dell'evoluzione biologica ed espressione dei vari aspetti dell'ambiente, inclusi quelli concernenti l'organizzazione sociale e la cultura. Sono queste cause remote che determinano chi si ammalerà, quando, come, che preparano eventualmente il terreno all'azione delle cause prossime in direzione dell'esordio della malattia.

- 3) Questa logica sostanzia ancor più radicalmente la necessità della medicina individualizzata già suggerita dall'approccio evoluzionistico (Childs, Wiener e Valle 2005). Poiché l'evoluzione è una conseguenza della riproduzione e delle variazioni a livello genetico, questa logica enfatizza l'individualità del paziente. Ogni paziente rappresenta un'espressione unica della malattia, con bisogni terapeutici specifici. Insistendo sull'importanza della genomica funzionale, questa nuova logica è in grado di concettualizzare l'individualità acquisita della malattia, non semplicemente ereditata e dipendente dalla storia della specie, l'individualità legata alla storia di una vita. Il gene infatti è espressione dell'ambiente e della cultura non soltanto in quanto effetto della selezione naturale, risultato di una storia che precede quella dell'organismo di cui fa parte. I geni regolano la sintesi proteica in funzione dei fattori ambientali. Stimoli interni ed esterni quali gli stadi dello sviluppo, le concentrazioni di ormoni e di mediatori chimici, la disponibilità di certi nutrienti, di certe sostanze, la presenza di certi contaminanti nell'ambiente, lo stress, l'apprendimento, l'interazione sociale, influenzano la formazione e il comportamento dei fattori di trascrizione genica, modulando sensibilmente l'espressione dei geni, ciò che si indica con il termine "regolazione epigenetica". In questo modo si dà forma all'individualità acquisita e quindi al carattere singolare che le condizioni morbose assumeranno in una specifica persona.
- 4) Questa nuova logica rivaluta l'importanza della prevenzione e delle azioni igienico-sanitarie il cui significato è stato ingiustamente ridimensionato dalle idee di efficacia e di efficienza delle terapie e dei trattamenti sviluppati dalla biomedicina. In questo senso tale logica accoglie le sollecitazioni a integrare l'ambiente, i valori, la cultura nella concettualizzazione biomedica della malattia. La vulnerabilità genetica porta infatti alle malattie quando si scontra con aspetti incongruenti dell'ambiente ovvero delle abitudini e degli stili di vita dell'individuo. Allora è chiaro che, attraverso la conoscenza del genoma umano, ampliandosi le possibilità di conoscere le basi della vulnerabilità verso le malattie, l'intervento sull'ambiente e sugli stili di vita potrà diventare sempre più determinante nell'azione medica. La

rivalutazione dell'igiene e della medicina preventiva implicita in questa nuova logica si accompagnerà a una migliore integrazione tra la dimensione della concettualizzazione e dell'azione sanitaria e il piano teorico e pratico della clinica. Ciò potrebbe avvenire sulla base del riconoscimento che entrambe le dimensioni dipendono dalla logica dei geni e dell'evoluzione, sull'idea che la specie umana è un insieme di sistemi aperti geneticamente unici, ognuno essendo il risultato delle pressioni selettive esercitate dalla natura e dalla società.

Scrive Childs:

Il primo principio di questa logica è che nasce dalla biologia dell'*Homo sapiens*, le cui origini si possono rintracciare nella selezione naturale e nell'evoluzione. La malattia è un sottoprodotto della necessità di immagazzinare variazioni per preservare le specie e contrapporsi alla variabilità delle condizioni esterne. La mutazione, espressa attraverso differenze nelle proteine che agiscono come singole unità e nelle tappe dell'omeostasi, compensa questa variazione. È attraverso queste unità e queste tappe che si esprime l'influenza sia delle variazioni genetiche sia ambientali. Esse sono l'interfaccia fra la natura e la cultura e costituiscono la capacità del singolo sistema aperto di adattarsi. Ora, dal momento che le mutazioni sono casuali, alcune delle variazioni nelle unità e nelle tappe saranno sicuramente disadattative, si esprimeranno in maniera incongruente con la realizzazione delle funzioni del sistema di cui fanno parte. Queste unità e queste tappe variabili, insieme alle componenti ambientali con cui sono incongruenti, costituiscono le cause prossime delle malattie le cui cause ultime sono (a) eventi genetici come le mutazioni da cui dipendono le unità e le tappe dell'omeostasi insieme alla ricombinazione, la segregazione, la spinta e le altre forze che giustificano la distribuzione dei loro prodotti; e (b) la storia sociale e culturale delle esperienze con cui la singola unità e la tappa sono incongruenti (Childs 1999, trad. it. pp. 50-51).

Una logica analoga a quella appena esposta è stata recentemente adottata anche nella formulazione di nuovi modelli teorici per la comprensione dei disturbi psichiatrici. In questo caso la sintesi della biologia evuzionistica con le acquisizioni e le teorie esplicative della genomica funzionale sembra poter condurre finalmente al definitivo superamento delle tradizionali divisioni funzionali e della classica divaricazione concettuale fra specialità mediche della mente e discipline fisiopatologiche.

Se infatti è vero che una combinazione di geni dà forma al comportamen-

to, incluso il comportamento sociale, così il comportamento e i fattori sociali – attraverso la loro azione sull'organismo e sul sistema nervoso centrale – modificano l'espressione dei geni e conseguentemente le funzioni delle cellule nervose, modulando di nuovo, circolarmente, il comportamento e la proiezione dell'individuo nella dimensione psicosociale (Andreasen 1997; Kandel 1998; Gabbard 2000). Ma i geni sono prodotti dell'evoluzione biologica e la filogenesi è a sua volta anche il risultato del comportamento, della cultura umana. Esiste, come abbiamo visto, una relazione complessa, circolare, nell'evoluzione, o meglio nella co-evoluzione, fra comportamento, cultura e funzioni dei geni e dunque fra comportamento, cultura, geni e malattia.

La regolazione dell'espressione genica delle cellule nervose incorpora, in senso letterale, i fattori ambientali e psicosociali. Nei processi di trascrizione del gene perciò la cultura può diventare natura. Così, in sostanza, le evidenze accumulate in questi ultimi anni hanno fatto cadere ogni residuo dualismo fra malattie somatiche e disturbi psichiatrici. Non esiste infatti nessuna differenza nella logica della fisiopatologia dei due tipi di malattia in quanto tutti gli stimoli, siano essi ambientali, fisiologici, psicosociali hanno un unico bersaglio finale, la regolazione genica. Allo stesso tempo, la regolazione genica costituisce l'elemento iniziale di una cascata di processi comportamentali e biologici tesi all'adattamento dell'organismo ovvero a un eventuale progressivo discostamento dall'omeostasi e quindi alla malattia.

La sintesi della biologia evoluzionistica con la genomica dunque ci riporta non solo all'integrazione funzionale dei vari apparati del corpo ma alla identità sostanziale dei fenomeni mentali e dei processi somatici, alle interazioni e alle continuità dei meccanismi biologici con gli aspetti comportamentali, con il contesto ambientale e sociale (Canali 2003). Essa ci permette così di spiegare anche la malattia mentale come effetto di un'incongruenza in uno dei piani (dal DNA al comportamento) attraverso cui si dispiegano i processi omeostatici, a livello di uno dei meccanismi concatenati su piani di complessità diversi ma comunque integrati, coerentemente finalizzati dalla storia evolutiva, radicati al livello molecolare dei geni.

RIFERIMENTI BIBLIOGRAFICI

- ANDREASEN, N. C. (1997): "Linking Brain and Mind in the Study of Mental Illnesses: a Project for a Scientific Psychopathology", *Science*, 275, pp. 1586-1593.
- BENCH, R. J. (1989): "Health Science, Natural Science, and Clinical Knowledge", *Journal of Medicine and Philosophy*, 14, pp. 147-164.

- BERNARD, C. (1865): *Introduction à l'étude de la médecine expérimentale*, Paris: J.-B. Baillière (trad. it. *Introduzione allo studio della medicina sperimentale*, Milano: Feltrinelli 1973).
- BERTALANFFY, L. von (1968): *General System Theory: Foundations, Development, Applications*, New York: George Braziller.
- BICHAT, X. (1799): *Traité des membranes en général et de diverses membranes en particulier*, Paris: Richard, Caillé et Ravier.
- (1800), *Recherches physiologiques sur la vie et la mort*, Paris: Brosson, Gabon et Cie.
- CANALI, S. (a cura di) (2001): *Drug-abuse, Evolution, Medicine*, numero monografico di *Medicina delle Tossicodipendenze – Italian Journal of the Addictions*, 31.
- (2003): “Evoluzione, emozioni e malattia”, in Canali e Pani (2003).
- CANALI, S. e PANI, L. (2003): *Emozioni e malattia. Dall'evoluzione biologica al tramonto della psicosomatica*, Milano: Bruno Mondadori.
- CANGUILHEM, G. (1988): “Le statut épistémologique de la médecine”, *History and Philosophy of the Life Sciences*, 10 (Suppl.), pp. 15-29.
- CHANG, Y. P., MAIER-REDELSPERGER, M., SMITH, K. D., CONTU, L., DUCROCO, R., DE MONTALEMBERT, M., BELLOY, M., ELION, J., DOVER, G. J., GIROT, R. (1997): “The Relative Importance of the X-linked FCP Locus and Beta-Globin Haplotypes in Determining Haemoglobin F Levels: a Study of SS Patients Homozygous for Beta S Haplotypes”, *British Journal of Haematology*, 96, pp. 806-814.
- CHILDS, B., WIENER, C. e VALLE, D. (2005): “A Science of the Individual: Implication for a Medical School Curriculum”, *Annual Review of Genomics and Human Genetics*, 6, pp. 313-330.
- CHILDS, B. (1999): *Genetic Medicine. A Logic of Disease*, Baltimore (Md.): The John Hopkins University Press (trad. it. *Medicina genetica. Una logica della malattia*, Roma: Giovanni Fioriti Editore 2004).
- COOPER, S. e PSATY, B. M. (2003): “Genomics and Medicine: Distraction, Incremental Progress, or the Dawn of a New Age?”, *Annals of Internal Medicine*, 138, pp. 576-580.
- CORBELLINI, G. (1998): “Le radici storico-critiche della medicina evoluzionistica”, in P. Donghi (a cura di) (1998, pp. 85-127).
- (2003): “Filosofia della medicina”, in N. Vassallo (a cura di), *Filosofie delle scienze*, Torino: Einaudi.
- (2004): “L'evoluzione della medicina evoluzionistica”, in Corbellini e Canali (a cura di) (2004).
- CORBELLINI, G. e CANALI, S. (a cura di) (2004): *Medicina darwiniana*, Bologna: Apeiron.

- DELKESKAMP-HAYES, C. e GARDELL CUTTER, M. A. (a cura di) (1993): *Science, Technology, and the Art of Medicine. European-American Dialogues*, Dordrecht: Kluwer.
- DIEDERICH, N. J. e GOETZ, C. G. (2008): “The Placebo Treatments in Neurosciences: New Insights from Clinical and Neuroimaging Studies”, *Neurology*, 71, pp. 677-684.
- DIEPPE, P. e BACHMANN, M. (2000): “Clinical Research: Is the Filling Missing from the Medical Research Sandwich?”, *The Journal of the Royal College of Physicians of London*, 34, pp. 282-284.
- DONGHI, P. (a cura di) (1998): *La Medicina di Darwin*, Roma-Bari: Laterza.
- DUPRÉ, J. (2000): “Reductionism”, in W.-H. Newton-Smith (a cura di), *A Companion to the Philosophy of Science*, Oxford: Blackwell, pp. 402-404.
- ENGEL, G. L. (1977): “The Need for a New Medical Model: a Challenge for Biomedicine”, *Science*, 196, pp. 129-136.
- ENGELHARDT, H. T. (1990): “Medical Knowledge and Medical Action: Competing Visions”, in Henk A. M. J. ten Have *et al.* (a cura di), *The Growth of Medical Knowledge*, Dordrecht: Kluwer.
- FOX KELLER, E. (2000): *The Century of the Gene*, Cambridge (Mass.): Harvard University Press.
- FRANK, A. W. (2002): *At the Will of the Body: Reflections on Illness*, Boston: Houghton Mifflin.
- GABBARD, G. O. (2000): “A Neurobiologically Informed Perspective on Psychotherapy”, *British Journal of Psychiatry*, 177, pp. 117-122.
- GIERE, R. N., BICKLE, J. e MAULDIN, R. F. (2006): *Understanding Scientific Reasoning*, v ed., Belmont (Calif.): Thomson.
- GILLETT, C. e LOEWER, B. (a cura di) (2001): *Physicalism and Its Discontents*, Cambridge: Cambridge University Press.
- GLENNAN, S. (1998): “Mechanisms and the Nature of Causation”, *Erkenntnis*, 44, pp. 49-71.
- GOLUB, E. S. (1997): *The Limits of Medicine*, Chicago-London (Ill.): The University of Chicago Press.
- GUTTMACHER, A. E. e COLLINS, F. S. (2002): “Genomic Medicine – a Primer”, *New England Journal of Medicine*, 347, pp. 1512-1520.
- HOLTZMAN, N. A. e MARTEAU, T. M. (2000): “Will Genetics Revolutionize Medicine?”, *New England Journal of Medicine*, 343, pp. 141-144.
- KANDEL, E. R. (1998): “A New Intellectual Framework for Psychiatry”, *American Journal of Psychiatry*, 155, pp. 457-469.
- KITCHER, P. (2001): “Battling the Undead: How (and How Not) to Resist Genetic Determinism”, in R. Singh, K. Krimbas, D. Paul e J. Beatty (a cura di), *Think-*

- ing about Evolution: Historical, Philosophical and Political Perspectives (Festschrift for Richard Lewontin)*, Cambridge: Cambridge University Press, pp. 396-414.
- KLOSTERHALFEN, S. e ENCK, P. (2008): “Neurophysiology and Psychobiology of the Placebo Response”, *Current Opinion in Psychiatry*, 21, pp. 189-195.
- KORNBERG, A. (1987): “The Two Cultures: Chemistry and Biology”, *Biochemistry*, 26, pp. 6888-6891.
- KUHN, Th. S. (1962/1970): *The Structure of Scientific Revolutions*, Chicago (Ill.): The University of Chicago Press (trad. it. *La struttura delle rivoluzioni scientifiche*, Torino: Einaudi, 1979).
- LAPLACE, P. S. de (1878-1912): *Théorie analytique des probabilités*, in *Œuvres complètes de Laplace, publiées sous les auspices de l'Académie des sciences, par MM. les secrétaires perpétuels*, Paris: Gauthier - Villars.
- LAURENT, A. (1994): *Storia dell'individualismo*, Bologna: Il Mulino.
- LEWINSOHN, R. (1998): “Medical Theories, Science, and the Practice of Medicine”, *Social Science and Medicine*, 46, pp. 1261-1270.
- LOUIS, P. C. A. (1828): “Recherche sur les effets de la saignée dans plusieurs maladies inflammatoires”, *Archives générales de médecine*, 18, pp. 321-336.
- (1835): *Recherches sur les effets de la saignée dans quelques maladies inflammatoires et sur l'action de l'émétique et des vésicatoires dans la pneumonie*, Paris: Librairie de l'Académie Royale de Médecine.
- (1836): *Researches on the Effects of Bloodletting in Some Inflammatory Diseases*, Boston: Hilliard, Gray and Company.
- MACHAMER, P., LINDLEY, D. e CRAVER, C. (2000): “Thinking About Mechanisms”, *Philosophy of Science*, 67, pp. 1-25.
- MACKIE, J. L. (1973): *The Cement of the Universe. A Study of Causation*, Oxford: Oxford University Press.
- MARCKMANN, G. (2001): “Teaching Science vs. the Apprentice Model – Do We Really Have the Choice?”, *Medicine, Health Care and Philosophy*, 4, pp. 85-89.
- MARCUM, J. A. (2008): *Humanizing Modern Medicine: an Introductory Philosophy of Medicine*, Berlin: Springer.
- MARCUM, J. A. e VERSCHUUREN, G. M. N. (1986): “Homeostatic Regulation and Whitehead's Philosophy of Organism”, *Acta Biotheoretica*, 35, pp. 123-133.
- MAYR, E. (1982): *The Growth of Biological Thought. Diversity, Evolution and Inheritance*, Cambridge (Mass.): Harvard University Press (trad. it. *Storia del pensiero biologico*, Torino: Einaudi 1990).
- (1988): *Toward a New Philosophy of Biology*, Cambridge (Mass.): Harvard University Press.
- MCKUSICK, V. A. (2001): “The Anatomy of the Human Genome: a Neo-Vesalian

- Basis for Medicine in the 21st Century”, *Journal of the American Medical Association*, 286, pp. 2289-2295.
- MCGUIRE, M. e TROISI, A. (1998): *Darwinian Psychiatry*, Oxford-New York-Tokyo: Oxford University Press (trad. it. *Psichiatria darwiniana*, Roma: Giovanni Fioriti Editore 2003).
- MCWHINNEY, I. R. (1978): “Medical Knowledge and the Rise of Technology”, *Journal of Medicine and Philosophy*, 3, pp. 293-304.
- MIETTINEN, O. S. (2001a): “The Modern Scientific Physician: 1. Can Practice be Science?”, *Canadian Medical Association Journal*, 165, pp. 441-442.
- (2001b): “The Modern Scientific Physician: 2. Medical Science versus Scientific Medicine”, *Canadian Medical Association Journal*, 165, pp. 591-592.
- NAGEL, E. (1961): *The Structure of Science*, New York: Harcourt, Brace, Jovanovic (trad. it. *La struttura della scienza*, Milano: Feltrinelli 1968).
- NEDERBRAGT, H. (2000): “The Biomedical Disciplines and the Structure of Biomedical and Clinical Knowledge”, *Theoretical Medicine and Bioethics*, 21, pp. 553-566.
- NELKIN, D. e LINDEE, S. (1995): *The DNA Mystique: The Gene as a Cultural Icon*, New York: Freeman.
- NESSE, R. M. e WILLIAMS, G. C. (1991): “The Dawn of Darwinian Medicine”, *Quarterly Review of Biology*, 66, pp. 1-22.
- (1995): *Why We Get Sick*, New York: Random House (trad. it. *Perché ci ammaliano*, Torino: Einaudi 1999).
- NORDIN, I. (1999): “The Role of Science in Medicine”, *Theoretical Medicine and Bioethics*, 20, pp. 229-244.
- OYAMA, S. (2000): *The Ontogeny of Information: Developmental Systems and Evolution*, II ed. rivista e accresciuta, Durham (N.C.): Duke University Press.
- PAUL, N. (1998): “Incurable Suffering from the ‘Hiatus Theoreticus’? Some Epistemological Problems in Modern Medicine and the Clinical Relevance of Philosophy of Medicine”, *Theoretical Medicine and Bioethics*, 19, pp. 229-251.
- PAULING, L., ITANO, H. A., SINGER, S. J. e WELLS, I. C. (1949): “Sickle Cell Anemia, a Molecular Disease”, *Science*, 110, pp. 543-548.
- PELLEGRINO, E. D. e THOMASMA, D. C. (1981): *A Philosophical Basis of Medical Practice*: Oxford-New York: Oxford University Press.
- PETERS, R. (1936): *Biochemical Lesion and Lethal Synthesis*, Oxford: Pergamon Press.
- (1969): “The Biochemical Lesion and Its Historical Development”, *British Medical Bulletin*, 25, pp. 223-226.
- PETTIT, Ph. (1993), “A Definition of Physicalism”, *Analysis*, 53, pp. 213-223.

- PLOTZ, P. H. (1997): “Deconstructing Disease: an Anatomy of Illness in the Age of Molecular Biology”, *Perspectives in Biology and Medicine*, 40, pp. 160-164.
- POLAND, J. (1994): *Physicalism: the Philosophical Foundations*, Oxford: Oxford University Press.
- QUINE, W. V. O. (1951): “Two Dogmas of Empiricism”, *The Philosophical Review*, 60, pp. 20-43.
- REES, J. (2004): “The Fundamentals of Clinical Discovery”, *Perspectives in Biology and Medicine*, 47, pp. 597-607.
- ROFF, D. (1992): *The Evolution of Life Histories: Theory and Analysis*, New York: Chapman & Hall.
- SARKAR, S. (1992): “Models of Reduction and Categories of Reductionism”, *Synthese*, 91, pp. 167-194.
- SCHADE, H. (1935): *Molekular Pathologie der Entzündung*, Dresden-Leipzig: Theodor Steinl.
- SCHWARTZ, M. A., WIGGINS, O. (1985): “Science, Humanism, and the Nature of Medical Practice: a Phenomenological View”, *Perspectives in Biology and Medicine*, 28, pp. 331-366.
- SCRIVER, C. R. (1982): “Treatment of Inborn Errors of Metabolism. The Nature/Nurture Argument Specified”, in M. A. Crawford, D. A. Gibbs e R. W. Watts (a cura di), *Advances in the Treatment of Inborn Errors of Metabolism. Proceedings of the 2nd Clinical Research Centre Symposium Held in September 1981*, New York: John Wiley & Sons Inc.
- (1984): “An Evolutionary View of Disease in Man”, *Proceedings of the Royal Society*, 220, pp. 273-298.
- SELDIN, D. (1981): “Presidential Address: The Boundaries of Medicine”, *Transactions of the Association of American Physicians*, 94, pp. 75-86.
- STEARNS, C. (a cura di) (1999): *Evolution in Health & Disease*, Oxford-New York-Tokyo: Oxford University Press.
- STEARNS, S. (1992): *The Evolution of Life Histories*, Oxford: Oxford University Press.
- STEVENS, A. e PRICE, J. (1996): *Evolutionary Psychiatry*, London-New York: Routledge, 1996.
- SUMMERS, K. M. (1996): “Relationship Between Genotype and Phenotype in Monogenic Diseases: Relevance to Polygenic Diseases”, *Human Mutation*, 7, pp. 283-293.
- TAUBER, A. I. e SARKAR, S. (1992): “The Human Genome Project: Has Blind Reductionism Gone too Far?”, *Perspectives in Biology and Medicine*, 35, pp. 220-235.

- TEMKIN, O. (1963): "The Scientific Approach to Disease: Specific Entity and Individual Sickness", in A. C. Crombie (a cura di), *Scientific Change*, New York: Basic Books.
- THAGARD, P. (1999): *How Scientists Explain Disease*, Princeton (N.J.): Princeton University Press (trad. it. *La spiegazione scientifica della malattia*, Milano: McGraw-Hill, 2001).
- TREVATHAN, W. R., SMITH, E. O. e MCKENNA, J. J. (a cura di) (1999): *Evolutionary Medicine*, Oxford-New York-Tokyo: Oxford University Press.
- UEXKÜLL, T. e WESIACK, W. (1991): *Theorie der Humanmedizin*, II ed., München: Urban & Schwarzenberg.
- VALLANCE, A. K. (2006): "Something out of Nothing: the Placebo Effect", *Advances in Psychiatric Treatment*, 12, pp. 287-296.
- VAN DIJCK, J. (2005): *The Transparent Body: A Cultural Analysis of Medical Imaging*, Seattle (Wash.): University of Washington Press.
- VENTER, J. C. et al. (2001): "The Sequence of the Human Genome", *Science*, 291, pp. 1304-1351.
- VIRCHOW, R. (1855): "Cellular-Pathologie", *Virchows Archiv*, 8, pp. 3-39.
- (1858): *Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre*, Berlin: A. Hirschwald.
- WEATHERALL, D. J. (1999): "From Genotype to Phenotype: Genetics and Medical Practice in the New Millennium", *Philosophical Transactions of the Royal Society of London*, series B, Biological Sciences, 354, pp. 1995-2010.
- (2000): "Single Gene Disorders or Complex Traits: Lessons from the Thalassemias and other Monogenic Diseases", *British Medical Journal*, 321, pp. 1117-1120.
- WEINER, H. (1977): "The Illusion of Simplicity: the Medical Model Revisited", *American Journal of Psychiatry*, 135, pp. 27-33.
- WHITBECK, C. (1977): "Causation in Medicine. The Disease Entity Model", *Philosophy of Science*, 44, pp. 619-637.
- WOLF, U. (1997): "Identical Mutations and Phenotypic Variation", *Human genetics*, 100, pp. 305-321.
- WULFF, H. R. (1990): "Function and Value of Medical Knowledge in Modern Disease", in Henk A. M. J. ten Have et al. (a cura di), *The Growth of Medical Knowledge*, Dordrecht: Kluwer.
- ZIMAN, J. (2002): *La vera scienza. Natura e modelli operativi della prassi scientifica*, Bari: Dedalo.

Deduzione, induzione e abduzione nelle scienze mediche

Roberto Festa

Dipartimento di Filosofia, Università di Trieste
e-mail: festa@units.it

Vincenzo Crupi

Dipartimento di Area Critica Medico-chirurgica, Università di Firenze
e-mail: vincenzo.crupi@unifi.it

Pierdaniele Giaretta

Dipartimento di Filosofia, Università di Padova
e-mail: pierdaniele.giaretta@unipd.it

1. Introduzione
2. Forme fondamentali di inferenza
3. Inferenze abduttive nelle scienze mediche
4. Il metodo ipotetico-deduttivo nelle scienze mediche
5. Le ipotesi causali e i metodi di Mill nelle scienze mediche
6. Osservazioni conclusive

SOMMARIO. Nel presente contributo ci si propone di trattare alcune forme di inferenza coinvolte nella valutazione di ipotesi nelle scienze mediche. A partire dalla tradizionale distinzione fra inferenze deduttive e induttive, si introdurranno e discuteranno diversi tipi di inferenza ben noti in filosofia della scienza (l'abduzione, il metodo ipotetico-deduttivo e i metodi di Mill), illustrandone attraverso esempi storici la rilevanza per la formazione e l'evoluzione delle conoscenze mediche.

PAROLE CHIAVE: deduzione, induzione, abduzione, scienze mediche.

1. Introduzione

Nelle scienze mediche, come in molte altre aree della ricerca scientifica, la valutazione di ipotesi sulla base di dati empirici riveste un ruolo centrale. Tale

valutazione coinvolge diverse forme di ragionamento o, in altri termini, diversi tipi di *inferenze*. Tradizionalmente si distinguono due forme fondamentali di ragionamento, vale a dire la deduzione e l'induzione. Nel seguito, cominceremo quindi la nostra trattazione illustrando gli aspetti generali delle inferenze deduttive e induttive (paragrafo 2). Ci concentreremo poi su alcuni specifici e importanti tipi di inferenza e mostreremo, con l'aiuto di esempi tratti dalla ricerca medica, quale sia il loro ruolo nella valutazione di ipotesi. In particolare, considereremo il modo in cui le inferenze abduttive (paragrafo 3), il metodo ipotetico-deduttivo (paragrafo 4) e i metodi di Mill (paragrafo 5) vengono applicati nella valutazione delle ipotesi in medicina.

2. Forme fondamentali di inferenza

2.1. Deduzione e induzione

In ogni tipo di *inferenza* si mettono in relazione un insieme di enunciati, detti *premesse*, e un singolo enunciato, detto *conclusione*. Un caso ben noto di inferenza *deduttiva* è rappresentato dal sillogismo categorico. Eccone un esempio:¹

- (1) *Premesse:*
 Coloro che presentano livelli elevati di ormoni tiroidei circolanti sono tachicardici.
 I malati del morbo di Basedow presentano livelli elevati di ormoni tiroidei circolanti.

Conclusione:
 I malati del morbo di Basedow sono tachicardici.

L'inferenza (1) si può ricondurre al seguente schema generale:

- (2) SILLOGISMO CATEGORICO
1. Tutti i Q sono R
 2. Tutti i P sono Q
-
- Tutti i P sono R .

¹ Preso a prestito da Federspil (1980, p. 34).

In questa notazione le premesse sono indicate dai numeri 1 e 2 sulla sinistra, mentre la linea orizzontale che separa l'ultima premessa dalla conclusione indica il carattere deduttivo di un'inferenza.

La cogenza del sillogismo categorico appare alquanto intuitiva. Anche per questo, probabilmente, si tratta di uno dei primi tipi di inferenza valida identificati nello studio della logica, a partire dagli scritti di Aristotele che hanno inaugurato la storia della disciplina. Facendo riferimento agli sviluppi contemporanei della logica, è possibile indicare che cosa stia a fondamento della validità del sillogismo categorico, così come di qualsiasi altro tipo di inferenza deduttiva.

Innanzitutto, la validità deduttiva riguarda primariamente una forma inferenziale, per esempio lo schema (2), e può essere intuitivamente presentata nel modo seguente: non è possibile "riempire di contenuti" la forma inferenziale in modo tale che le premesse risultino vere e la conclusione falsa. Nel caso dello schema (2) i "contenuti" possono essere inseriti mediante la sostituzione di "P", "Q" e "R" con opportune espressioni linguistiche, quali, per esempio, "neuroni", "cellule", "tessuti". È facile vedere che nessuna di queste sostituzioni può dare come esito delle premesse vere e una conclusione falsa e perciò lo schema è deduttivamente valido. In secondo luogo, si applica la nozione di validità deduttiva anche alle specifiche inferenze formulate nel linguaggio naturale, definendo un'inferenza deduttivamente valida quando la forma dell'inferenza è deduttivamente valida nel senso sopra indicato. Per esempio, l'inferenza (1) è deduttivamente valida poiché la forma (2) lo è. In questo modo la validità deduttiva di una specifica inferenza viene ricondotta alla validità deduttiva della sua forma. In riferimento a questa seconda accezione della nozione di validità deduttiva è bene tenere presente alcune cautele e precisazioni, che possono essere introdotte considerando l'esempio seguente:

(1*) *Premesse:*
Fido è un cane.

Conclusione:
Fido è un animale.

La forma dell'inferenza (1*) è:

(2*) *Premesse:*
L'individuo x è P .

Conclusione:
L'individuo x è Q

che non è deduttivamente valida, poiché è possibile esemplificare lo schema (2*) in modo tale da ottenere una premessa vera e una conclusione falsa. L'inferenza (1*) può essere considerata deduttivamente valida solo se si considera anche l'enunciato "Tutti i cani sono animali" come una sua premessa *implicita*. Esempi di questo genere suggeriscono di prendere in considerazione la forma dell'inferenza solo dopo avere esplicitato con opportune premesse tutto ciò che è implicitamente presupposto in una specifica inferenza. (Come si può facilmente immaginare, l'individuazione di ciò che è implicitamente presupposto può talora rivelarsi un compito alquanto impegnativo.)

Per indicare che la conclusione di un'inferenza è conseguenza deduttivamente valida delle premesse si impiegano talora anche le seguenti espressioni, che nel seguito considereremo interscambiabili: la conclusione *segue logicamente* dalle premesse; la conclusione è *deducibile* dalle premesse; la conclusione *segue necessariamente* dalle premesse; le premesse *implicano (logicamente)* la conclusione. L'equivalenza di questi modi alternativi di esprimersi richiederebbe a rigore una specifica discussione, ma tutti hanno in comune l'idea che le inferenze deduttive sono in grado di *trasmettere la verità*: se si ammette la verità delle premesse, allora si deve anche ammettere la verità della conclusione. Come ora vedremo, lo stesso non accade nelle inferenze *induttive*. Un ben noto esempio è rappresentato dalle cosiddette *generalizzazioni induttive*. Si consideri in proposito la seguente inferenza:

(3) *Premessa:*

Mille malati del morbo di Basedow finora osservati presentano livelli elevati di ormoni tiroidei circolanti.

Conclusione:

I malati del morbo di Basedow presentano livelli elevati di ormoni tiroidei circolanti.

L'inferenza (3) si può ricondurre al seguente schema generale:

(4) GENERALIZZAZIONE INDUTTIVA

1. Tutti gli n elementi Q finora osservati sono R

Tutti i Q sono R .

Come si può vedere, la linea orizzontale, che nello schema delle inferenze deduttive separa le premesse dalla conclusione, viene qui sostituita da una dop-

pia linea orizzontale, abitualmente impiegata per indicare il carattere induttivo dell'inferenza. È facile rendersi conto che la generalizzazione induttiva, diversamente dalle inferenze deduttive, *non* garantisce la verità della conclusione una volta ammessa la verità delle premesse. È infatti perfettamente possibile immaginare che i mille malati del morbo di Basedow finora osservati presentino livelli elevati di ormoni tiroidei circolanti e che nondimeno ciò non valga per *tutti* i malati del morbo di Basedow – in particolare, che non valga per almeno qualcuno di quelli non ancora osservati. Tuttavia, è naturale ritenere che la premessa, se vera, fornisca alla conclusione un certo sostegno – in un senso su cui ora dovremo soffermarci.²

2.2. *Plausibilità e conferma induttiva*

Nelle scienze empiriche, la valutazione di un'ipotesi H avviene sulla base di dati osservativi o sperimentali – o elementi di evidenza – descrivibili mediante un enunciato E . Normalmente, H non è deducibile da E . Per questo motivo la valutazione di H riflette solitamente l'impiego di diverse forme di inferenza induttiva. Le relazioni induttive che intercorrono fra E e H possono essere di almeno due tipi distinti, corrispondenti alle due domande seguenti:

- (i) H è *plausibile* alla luce di E ?
- (ii) La *plausibilità* iniziale di H è *accresciuta* da E ?

Come si vede, abbiamo formulato le domande (i) e (ii) facendo uso della nozione di *plausibilità*. Nella maggior parte delle inferenze induttive coinvolte nella vita quotidiana e nella ricerca scientifica, infatti, si impiega un concetto informale di *plausibilità*, al quale nel seguito ci atterremo. Di norma, diciamo semplicemente che “ H è *plausibile*” (o, in modo spesso intercambiabile, che è “*probabile*”) per indicare che abbiamo una certa fiducia nella verità di H .

Le domande (i) e (ii) suggeriscono due tipi di inferenze induttive. Operando il primo tipo di inferenza, asseriamo che sulla base della premessa E dobbiamo ritenere *plausibile* la conclusione H . Operando il secondo tipo di inferenza, d'altra parte, asseriamo che la premessa E *accresce* la *plausibilità* ini-

² Un'analisi più approfondita dei vari tipi di inferenze deduttive e induttive è al di là dei nostri presenti scopi. Esistono in proposito ottime trattazioni manualistiche, a cui rinviamo il lettore interessato. Si vedano, per esempio, Copi e Cohen (2001), Hacking (2001), Hurley (1999), Salmon (1983), Skyrms (1999) e Varzi, Rolt e Rohatyn (1998).

ziale di H , cioè accresce la fiducia che si riponeva nella verità di H prima di venire a conoscenza delle informazioni contenute in E . Nel seguito ci concentreremo sulle inferenze induttive del secondo tipo, poiché esse sembrano svolgere un ruolo di primo piano – raramente contestato – nella pratica scientifica. Seguendo una terminologia consolidata, le indicheremo come inferenze di *conferma (induttiva)*.³ Nella nostra notazione, la più semplice illustrazione dell’inferenza di conferma induttiva può essere così rappresentata:

(5) CONFERMA INDUTTIVA

$$\begin{array}{l} 1. \quad E \\ \quad \quad = \\ \quad \quad H. \end{array}$$

Di solito, la plausibilità iniziale dell’ipotesi H è fondata su una *conoscenza di sfondo*, spesso condivisa da una certa comunità di ricerca, e non sempre esplicitamente verbalizzata. In molti casi, comunque, è possibile e utile formulare esplicitamente, mediante opportuni enunciati, almeno una parte rilevante di tale conoscenza di sfondo, che nel seguito indicheremo congiuntamente con S . È bene precisare fin d’ora che S comprende solitamente un insieme alquanto composto di enunciati. Alcuni hanno carattere *teorico* e sono presupposti dai ricercatori in quanto considerati altamente plausibili. Altri hanno carattere *empirico* e specificano le particolari condizioni in cui determinate osservazioni o misurazioni sperimentali si verificano. Infine, possono far parte di S enunciati di carattere *logico*, che stabiliscono relazioni fra altri enunciati (teorici ed empirici) rilevanti. Facendo riferimento alla conoscenza di sfondo S è possibile rappresentare lo schema di un’inferenza di conferma “relativizzata”, che riprenderemo ripetutamente nel seguito:

(6) CONFERMA INDUTTIVA RELATIVIZZATA

$$\begin{array}{l} 0. \quad S \\ 1. \quad E \\ \quad \quad = \\ \quad \quad H. \end{array}$$

³ È bene chiarire che, a differenza di quanto fatto per le inferenze deduttive, nel seguito considereremo alcune forme tipiche di conferma induttiva ma non mireremo a indicare le condizioni generali nelle quali la plausibilità di un’ipotesi H alla luce di E sia maggiore della plausibilità iniziale di H .

Effettuando l'inferenza (6) affermiamo che la premessa 1, vale a dire l'enunciato E , conferma H relativamente a S ; cioè che le informazioni contenute in E , se aggiunte alla conoscenza di sfondo S , accrescono la plausibilità iniziale di H . Ciò a sua volta significa che la plausibilità di H alla luce di ($S \& E$) è maggiore della plausibilità iniziale di H , valutata alla luce di S soltanto. (Qui e nel seguito indicheremo con “&” la congiunzione logica di due enunciati.)

3. Inferenze abduttive nelle scienze mediche

3.1. L'abduzione

Già a partire dall'opera del filosofo statunitense Charles Sanders Peirce (1839-1914), che introdusse il concetto di abduzione e svolse le prime ricerche sistematiche sulla natura del ragionamento abduttivo e sul suo ruolo nell'indagine scientifica, si parla di abduzione in due accezioni: (i) con riferimento a un particolare genere di *inferenze* dotate di caratteristiche peculiari che le distinguerebbero sia dalle inferenze deduttive sia da quelle induttive – o almeno dalle forme tradizionalmente più discusse di inferenze induttive; (ii) con riferimento a *procedure euristiche* impiegate nella generazione, o scoperta, delle ipotesi scientifiche. Poiché in questa sede siamo interessati all'analisi delle forme di ragionamento coinvolte nella valutazione di ipotesi, ci occuperemo solo della prima accezione di abduzione, vale a dire delle *inferenze abduttive*.⁴

Tra gli studiosi che negli ultimi decenni si sono occupati delle inferenze abduttive vi è ampio consenso su due punti. Il primo è che le premesse di un'inferenza abduttiva comprendono la descrizione di qualche interessante fatto osservato, mentre la conclusione dell'inferenza è data da un'ipotesi esplicativa che fornisce una buona spiegazione di tale fatto. Il secondo punto è che un'inferenza abduttiva determina un aumento della plausibilità iniziale della conclusione. Nella discussione che segue ci proponiamo di esporre in che senso l'inferenza abduttiva è appunto una “inferenza a una buona spiegazione”. Mostriamo inoltre che essa può venire intesa come una particolare forma di inferenza induttiva, e più precisamente come una particolare forma di conferma relativizzata.

⁴ Un'utile trattazione storica e teorica dei diversi aspetti del ragionamento abduttivo, con numerosi ulteriori riferimenti, è fornita da Niiniluoto (1999).

Un importante aspetto della ricerca scientifica riguarda le relazioni esplicative tra ipotesi ed eventi osservati. Per i nostri scopi, faremo uso di una nozione di *spiegazione (deduttiva)* che soddisfa i requisiti indicati qui di seguito:

- (7) Data la conoscenza di sfondo S , l'ipotesi H fornisce una *spiegazione (deduttiva)* di un evento (descritto da un enunciato) E nel caso in cui:
- (i) siamo certi della verità di E ;
 - (ii) S non implica logicamente E ;
 - (iii) $(S \ \& \ H)$ implica logicamente E .

Il riferimento alla spiegazione svolge un ruolo di grande rilievo nella ricerca scientifica e, in particolare, nelle inferenze abduttive. Non di rado, infatti, accade che gli scienziati si imbattano in un evento E in qualche misura sorprendente, cioè non pienamente spiegabile (deducibile) in base alla conoscenza di sfondo S . Di fronte a un evento sorprendente, una mossa piuttosto naturale è quella di cercarne una spiegazione. Nei casi coronati da successo, si può così giungere alla scoperta che una determinata ipotesi H offre una possibile spiegazione (deduttiva) di E .⁵ Molti studiosi sostengono che la scoperta delle potenzialità esplicative di un'ipotesi H rispetto a qualche sorprendente evento E osservato in precedenza conduce a una conferma di H , accrescendone la plausibilità iniziale. Chi condivide questo punto di vista sarà pronto ad adottare il seguente schema di inferenza abduttiva:

- (8) INFERENZA ABDUTTIVA
- | | | |
|----|---|-------|
| 0. | S | |
| 1. | H fornisce una spiegazione deduttiva di E | |
| | | H . |

Si può notare che l'inferenza abduttiva (8) rappresenta una particolare forma di conferma relativizzata (si veda lo schema (6)). Effettuando tale inferenza affermiamo che l'ipotesi H è confermata (relativamente a S) dal fatto di fornire una spiegazione dell'evento noto E . Si osservi che in questo caso l'informazione confermantente espressa nella premessa 1 non riflette alcuna nuova scoperta

⁵ Si noti che H può essere un'ipotesi già formulata in precedenza oppure essere stata concepita nel tentativo di spiegare E .

ta di carattere empirico, comprendendo piuttosto la scoperta di relazioni *logiche* intercorrenti fra H ed E (alla luce di S).

Nella storia delle scienze empiriche, è molto comune che si argomenti in favore di un'ipotesi in base alla sua capacità di rendere conto di fenomeni già noti ma altrimenti privi di spiegazione, conformemente allo schema di inferenza abduttiva appena descritto. Per esempio, due fra i maggiori successi inizialmente riconosciuti della meccanica newtoniana furono di carattere esplicativo, consistendo nella capacità di fornire una spiegazione deduttiva del moto dei pianeti (approssimativamente conforme alle leggi precedentemente formulate da Keplero) e del moto dei gravi (approssimativamente conforme alla legge galileiana).

3.2. Inferenze abduttive nelle scienze mediche: due esempi

Esempi di inferenze abduttive nel senso appena descritto ricorrono anche nelle scienze mediche. Qui ne illustreremo nei dettagli due: il primo è tratto dagli albori della medicina scientifica e riguarda l'ipotesi della contaminazione nell'eziologia della febbre puerperale, formulata da Semmelweis verso la metà dell'Ottocento; il secondo appartiene invece alla ricerca medica più recente e si riferisce a un'influente ipotesi sulla patogenesi dell'AIDS.

ESEMPIO 1. *L'ipotesi della contaminazione nell'eziologia della febbre puerperale (prima parte)*. Il medico ungherese Ignaz Semmelweis (1818-65) condusse le sue ricerche sull'eziologia della febbre puerperale fra il 1844 e 1848, mentre lavorava al primo reparto di maternità dell'Ospedale Generale di Vienna. Il punto di partenza delle sue indagini era costituito dall'osservazione (E) che una percentuale preoccupante delle donne che partorivano nel suo reparto contraeva una malattia grave e spesso fatale, nota appunto come febbre puerperale, e che tale percentuale era ampiamente maggiore di quella riscontrata nel secondo reparto di maternità dello stesso ospedale. Semmelweis considerò diverse ipotesi sull'origine della febbre puerperale, fra le quali almeno tre sembravano in grado di fornire una possibile spiegazione dello strano fenomeno E . (i) L'ipotesi iatrogena (H_1) suggeriva di spiegare E come una conseguenza di visite maldestre condotte dagli studenti di medicina, alla luce della circostanza che questi ultimi esercitavano il loro tirocinio esclusivamente nel primo reparto. (ii) Secondo l'ipotesi psicosomatica (H_2), invece, le pazienti del primo reparto erano maggiormente vulnerabili alla malattia per il fatto che nel primo reparto, a differenza del secondo, il prete che somministrava l'estrema unzio-

ne, preceduto da un inserviente che suonava una campanella, doveva attraversare le corsie per raggiungere la morente, terrificando così le pazienti al punto da debilitarle. (iii) Secondo l'ipotesi della contaminazione (H_3), infine, a provocare la morte delle partorienti era in molti casi la contaminazione da particelle di materia cadaverica trasportate da professori, assistenti e studenti che – nel primo reparto soltanto – passavano direttamente dalla dissezione di cadaveri all'esame fisico delle pazienti. Si osservi che H_3 postulava una circostanza all'epoca tutt'altro che pacifica, vale a dire che la pratica comunemente adottata – e seguita anche nel primo reparto – di lavarsi le mani con acqua e sapone fosse insufficiente a rimuovere completamente le particelle cadaveriche.

Poiché ciascuna delle ipotesi H_1 - H_3 forniva una spiegazione del fenomeno E , tutte e tre le ipotesi risultavano confermate – sulla base di inferenze abduttive – da questo specifico, e comune, successo esplicativo. Tuttavia, nel 1847, Semmelweis venne a conoscenza del tragico evento (E^*) della morte di un suo collega, il professor Kolletschka. Kolletschka si era ferito accidentalmente nel corso di un'autopsia. Poco dopo era deceduto manifestando sintomi molto simili a quelli della febbre puerperale. L'ipotesi H_3 offriva una spiegazione anche di E^* , mentre H_1 e H_2 non erano ovviamente in grado di farlo. In considerazione di questo e altri successi esplicativi di H_3 , non condivisi da H_1 e H_2 , Semmelweis concentrò sempre più la sua attenzione sull'ipotesi della contaminazione. Come vedremo più avanti, inoltre, le inferenze abduttive appena descritte furono accompagnate e seguite da diversi controlli empirici appositamente escogitati, che condussero infine Semmelweis alla risoluzione del suo problema proprio in favore dell'ipotesi H_3 (si veda l'Esempio 3).⁶

ESEMPIO 2. *L'“ipotesi del lavandino” nella patogenesi dell'AIDS (prima parte)*. Negli anni novanta del secolo scorso si considerava ormai stabilita l'esi-

⁶ La scoperta di Semmelweis, di cui egli fornì un vivido e dettagliato resoconto (Semmelweis 1861), è uno degli episodi più noti e romanzeschi della storia della ricerca medica. Per un verso, il pionieristico lavoro di Semmelweis è stato considerato un luminoso esempio dell'inventiva e dell'audacia di un ricercatore osteggiato dal suo ambiente, al punto da attrarre l'attenzione di Celine (1952), medico egli stesso oltre che scrittore geniale e controverso. Per altro verso, le indagini di Semmelweis sono state viste come un caso paradigmatico di ricerca scientifica in medicina. Data la varietà di posizioni metodologiche e teoriche esistenti, non è allora sorprendente che i filosofi della scienza abbiano proposto una ricca rassegna di interpretazioni, talora divergenti, circa la natura dei metodi di Semmelweis e il significato della sua parabola scientifica. Così, studiosi diversi hanno rilevato nelle ricerche di Semmelweis soprattutto la centralità del ragionamento abduttivo (Lipton 2004, pp. 75-98) o di quello analogico (Thagard

stenza di un legame causale fra l'infezione da Hiv e le manifestazioni cliniche dell'Aids. Si supposeva che l'Hiv causasse l'Aids deprimendo progressivamente un'importante componente del sistema immunitario, cioè i cosiddetti linfociti CD4+. Si riteneva infatti, sulla base di precedenti osservazioni di laboratorio, che queste cellule rappresentassero un bersaglio biologico privilegiato del virus, che sarebbe stato in grado di distruggerle, per esempio attraverso il comune meccanismo della citolisi. Tuttavia, l'elaborazione di un resoconto pienamente soddisfacente della storia naturale della malattia sembrava imbattersi in una notevole anomalia (*E*), talora indicata come "il paradosso centrale della patogenesi virale": la proporzione di linfociti CD4+ infetti in pazienti malati appariva decisamente troppo bassa (anche secondo le stime più generose, non più di 1 su 100) per determinare l'irrimediabile compromissione del sistema immunitario (Sheppard, Ascher e Krokwa 1993). Secondo un'efficace metafora impiegata da alcuni studiosi, sembravano esserci "più cadaveri che pallottole" (Ascher *et al.* 1995).

Nel 1995, David Ho e altri studiosi presentarono la cosiddetta "ipotesi del lavandino", che per qualche tempo ha rappresentato il modello patogenetico dell'Aids più popolare e discusso (Ho *et al.* 1995; Wei *et al.* 1995). In breve, Ho e i suoi collaboratori elaborarono un semplice modello matematico secondo il quale il declino della quantità misurata di linfociti CD4+ (l'abbassamento del livello dell'acqua nel lavandino) emergeva come l'effetto relativamente lento di un elevatissimo turnover (circa due miliardi di cellule al giorno), prodotto dalla continua distruzione da parte dell'Hiv (lo scarico del lavandino) e da un continuo sforzo di compensazione del sistema immunitario (il rubinetto), che da ultimo, nel volgere di anni, avrebbe condotto al collasso del sistema (il lavandino si svuota). A sostegno dell'ipotesi di Ho (*H*), alcuni ricercatori argomentarono che essa forniva una possibile soluzione del "paradosso delle pallottole", offrendo una spiegazione del fatto (*E*) che la gran parte dei linfociti CD4+ non apparissero infetti. In estrema sintesi, tale spiegazione può

1999, pp. 137-138), l'adozione di una logica induttiva della scoperta (Pera 1983) o un'implicita applicazione dell'approccio bayesiano al "problema di Duhem" (Giorello e Moriggi 2004). Altri ancora hanno ricostruito il lavoro di Semmelweis nei termini del modello popperiano per congetture e confutazioni (Antiseri 1977) o secondo il modello ipotetico-deduttivo (Hempel 1966, pp. 15-19; si veda anche il nostro Esempio 3). Non sono mancati neppure gli studiosi che vi hanno ravvisato un'applicazione dei metodi induttivi di Mill (Copi e Cohen 2001, p. 504) o un'anticipazione dei metodi statistici oggi impiegati dall'epidemiologia clinica (Salmon 1983; Vineis 1999, pp. 17 ss.). Infine, c'è chi ha visto nelle difficoltà di Semmelweis a convincere i colleghi della validità delle sue teorie un classico esempio della tenacia dei paradigmi accettati, in linea con le tesi di Kuhn (Gillies 2005).

essere delineata come segue: i numerosi linfociti colpiti dal virus, una volta infetti, vengono rapidamente distrutti e continuamente sostituiti da nuove cellule; di conseguenza, in ogni istante la maggior parte dei linfociti presenti sono di recente creazione, e non sono ancora stati infettati (Maddox 1995).

Anche in questo caso, come in quello di Semmelweis, l'argomentazione abduttiva in favore dell'ipotesi di Ho contribuì ad attirare su di essa l'attenzione dei ricercatori, motivando i controlli sperimentali cui fu in seguito sottoposta (si veda l'Esempio 4).

4. Il metodo ipotetico-deduttivo nelle scienze mediche

4.1. Il metodo ipotetico-deduttivo

Alcune procedure deduttive e induttive ampiamente utilizzate nella valutazione delle ipotesi scientifiche sono quelle coinvolte nel cosiddetto *metodo ipotetico-deduttivo*. Tale metodo si basa sull'idea che possiamo mettere alla prova un'ipotesi deducendone alcune previsioni di carattere osservativo o sperimentale, e controllando poi se si realizzano oppure no. Per i nostri scopi, faremo uso di una nozione di *predizione (deduttiva)* che soddisfa i requisiti indicati qui di seguito:

- (9) Data la conoscenza di sfondo S , l'ipotesi H fornisce una *predizione (deduttiva)* di un evento (descritto da un enunciato) E nel caso in cui:
- (i) non siamo certi della verità di E ;
 - (ii) S non implica logicamente E ;
 - (iii) $(S \& H)$ implica logicamente E .

Evidentemente, i controlli osservativi o sperimentali volti ad accertare se un evento predetto E si realizza possono avere due esiti: (i) si stabilisce che E non si realizza o, equivalentemente, che si verifica l'evento non- E ; (ii) si stabilisce che l'evento E effettivamente si realizza. Come vedremo tra poco in modo più dettagliato, nel primo caso possiamo concludere (deduttivamente) che l'ipotesi H è falsificata da E (alla luce di S), mentre nel secondo caso possiamo concludere che H è confermata (induttivamente) da E (relativamente a S).

Molti episodi della storia della scienza possono essere interpretati come applicazioni più o meno consapevoli ed esplicite del metodo ipotetico-deduttivo. Nel 1695, per esempio, l'astronomo Edmund Halley applicò la meccanica

newtoniana a una cometa che aveva osservato nel 1682 e ne dedusse la previsione che la cometa avrebbe impiegato circa settantacinque anni per compiere un'orbita completa e tornare visibile dalla Terra. Halley morì nel 1743, quindici anni prima del ritorno della cometa, che ricomparì puntualmente il giorno di Natale del 1758. In accordo con il metodo ipotetico-deduttivo, il successo della previsione di Halley apparve ai ricercatori come una straordinaria conferma della meccanica newtoniana. Questo tipo di procedura di ricerca e di valutazione delle ipotesi è stato ampiamente usato anche nelle scienze mediche fin dai loro esordi, almeno a partire dalle ricerche di William Harvey (*De motu cordis*, 1628) sul movimento del cuore e la circolazione del sangue negli animali.⁷ La stessa insistenza di Claude Bernard (*Introduzione allo studio della medicina sperimentale*, 1865) sulla necessità del “ragionamento sperimentale” come base della medicina moderna riflette l'adozione dei principi di fondo di ciò che oggi è appunto comunemente indicato come metodo ipotetico-deduttivo.⁸

Come anticipato, il metodo ipotetico-deduttivo si fonda su due tipi di inferenze, corrispondenti ai due possibili esiti del controllo empirico di un evento predetto E . Cominciamo dal caso in cui l'osservazione attesti che E non si verifica, cioè in cui si osservi non- E . Il corrispondente tipo di inferenza, che chiameremo *falsificazione ipotetico-deduttiva*, può essere rappresentato come segue:

(10) FALSIFICAZIONE IPOTETICO-DEDUTTIVA

- | | |
|----|--|
| 1. | S |
| 2. | H fornisce una predizione deduttiva di E |
| 3. | non- E |
| | |
| | non- H . |

È possibile mostrare che lo schema di inferenza (10) soddisfa i criteri della validità deduttiva, vale a dire che dalla verità delle premesse consegue con ne-

⁷ Secondo Johansson e Lynøe (2009, p. 142), che ne forniscono un'esposizione dettagliata, le ricerche di Harvey sono “giustamente lodate come un eccellente esempio del modo in cui la ricerca medica dovrebbe essere condotta”. Oltre all'impiego del metodo ipotetico-deduttivo, Johansson e Lynøe rinvencono nel lavoro di Harvey vari altri tipi di inferenze induttive quali generalizzazioni, inferenze abduttive, analogiche e di altro genere.

⁸ Secondo Federspil (1980, p. 8), “non è difficile rendersi conto che tutta la medicina scientifica, nel suo complesso, è stata elaborata con questa metodologia. [...] L'analisi delle conoscenze biomediche, considerate globalmente, rivela una costante aderenza ai principi del metodo ipotetico-deduttivo”.

cessità la verità della conclusione (nella quale si asserisce che H è falsa). Esso rappresenta quindi una particolare forma di inferenza deduttiva.⁹

Il secondo tipo di inferenza che ci interessa, che chiameremo *conferma ipotetico-deduttiva*, rappresenta invece una forma di conferma induttiva relativizzata, illustrata come segue:

(11) CONFERMA IPOTETICO-DEDUTTIVA

- | | | |
|----|--|-------|
| 0. | S | |
| 1. | H fornisce una predizione deduttiva di E | |
| 2. | E | |
| | | H . |

Vale la pena di integrare la presentazione dello schema di inferenza (11) con un principio di carattere quantitativo, formulato in base alla nozione di evidenza *sorprendente*. Tale nozione deve essere definita in relazione alla conoscenza di sfondo S : diciamo che l'evento E è tanto più sorprendente quanto meno tale evento appare plausibile in base a S . Nelle analisi filosofiche e nella pratica della scienza è ampiamente diffusa la convinzione che un'ipotesi H che fornisce una predizione dell'evento E è tanto maggiormente confermata dall'osservazione di E quanto più tale osservazione è sorprendente nel senso appena definito.

4.2. Il metodo ipotetico-deduttivo nelle scienze mediche: due esempi

Riprendendo la discussione dei due casi storici introdotti in precedenza (si veda il paragrafo 3.2), illustriamo qui di seguito due esempi di applicazione del metodo ipotetico-deduttivo nelle scienze mediche. Il primo dei due esempi, che sviluppa la precedente esposizione delle ricerche di Semmelweis, illustra sia la falsificazione sia la conferma ipotetico-deduttiva. Il secondo, ancora tratto dalla recente ricerca sulla patogenesi dell'AIDS, illustra invece soltanto la falsificazione di un'ipotesi.

⁹ Alcuni studiosi, a partire da Karl Popper (1902-94), hanno considerato la falsificazione ipotetico-deduttiva come la pietra miliare del corretto metodo scientifico. Per un'applicazione dell'approccio falsificazionista di Popper alla ricerca medica, si veda Antiseri (1977).

ESEMPIO 3. *L'ipotesi della contaminazione nell'eziologia della febbre puerperale (seconda parte)*. Avendo concentrato la sua attenzione sulle ipotesi H_1 - H_3 , Semmelweis le sottopose a svariati controlli sperimentali che condussero alla falsificazione di H_1 e H_2 e alla conferma dell'ipotesi della contaminazione H_3 . Consideriamo, per esempio, la falsificazione dell'ipotesi psicosomatica H_2 . Semmelweis si chiese se vi fossero effetti osservabili deducibili dall'ipotesi che si potessero mettere alla prova dell'osservazione. Supponendo che l'arrivo del sacerdote non venisse rilevato dalle pazienti, l'ipotesi H_2 avrebbe fornito una ben precisa predizione (E): una pronta e significativa diminuzione della mortalità da febbre puerperale nel primo reparto, in cui Semmelweis operava, a livelli simili a quelli del secondo. Egli decise quindi di persuadere il sacerdote a raggiungere le pazienti morenti silenziosamente e senza farsi notare. Poiché ciononostante non si registrò alcuna diminuzione della mortalità (non- E), l'ipotesi psicosomatica H_2 risultò falsificata. Successivamente, Semmelweis sottopose a controllo empirico anche H_3 . Anche in questo caso, si chiese quali predizioni osservabili conseguissero dall'ipotesi. Pensò così che, se la contaminazione da particelle cadaveriche fosse stata all'origine della febbre puerperale (H_3), allora, in presenza di opportune misure antisettiche in grado di eliminare chimicamente il materiale infettivo dalle mani di medici e studenti, l'incidenza della malattia nel primo reparto sarebbe dovuta scendere ai livelli del secondo (E). Emise perciò una direttiva che, a partire dal maggio del 1847, faceva obbligo di lavarsi le mani con una soluzione di ipoclorito di calcio prima di procedere alla visita delle pazienti. La mortalità per febbre puerperale nel suo reparto cominciò ben presto a diminuire, e per il 1848 calò fino all'1,27%, a fronte dell'1,33% registrato nel secondo reparto. In tal modo, H_3 risultava confermata dagli esiti dell'esperimento. Tale conferma appariva inoltre molto forte, per il fatto che l'effetto osservato a seguito del lavaggio delle mani sarebbe apparso altamente sorprendente alla luce della sola conoscenza di sfondo.

ESEMPIO 4. *L'“ipotesi del lavandino” nella patogenesi dell'AIDS (seconda parte)*. La medicina sperimentale contemporanea ha elaborato diversi metodi per indagare la “dinamica” di diverse famiglie di cellule del corpo umano *in vivo*. Una delle procedure più ingegnose e affidabili per questo tipo di osservazioni può essere descritta sinteticamente come segue. In primo luogo, si somministra a un individuo (per via orale, o per iniezione) una sostanza (come glucosio o acqua) che è stata arricchita con deuterio, un isotopo dell'idrogeno – la cui assunzione non presenta rischi – che viene incorporato nella sintesi del DNA durante la riproduzione cellulare. In seguito, a scadenze di tempo appo-

sitamente programmate, vengono prelevati opportuni campioni di tessuto o di sangue. Le cellule della popolazione di interesse vengono quindi purificate in laboratorio e sottoposte alla misurazione dell'arricchimento isotopico del DNA. In tal modo si rileva la proporzione di DNA isotopico in funzione del tempo, dalla quale attraverso metodi di calcolo standard si possono stimare dati come il tasso di riproduzione o la sopravvivenza media nella popolazione di cellule interessate.

Le tecniche sperimentali appena descritte si sono rivelate decisive per sottoporre a controllo l'ipotesi di Ho e colleghi (*H*) sulla patogenesi dell'AIDS. Come si ricorderà (si veda l'Esempio 2), tale ipotesi postulava – e rappresentava matematicamente – un frenetico turnover dei linfociti CD4+ durante l'infezione da HIV. Dall'ipotesi *H* si poteva perciò derivare deduttivamente una ben precisa predizione (*E*): il calcolo sperimentale del tasso di riproduzione dei linfociti CD4+ in pazienti HIV-positivi (ma non sottoposti a trattamento antiretrovirale) avrebbe dovuto rivelare, in media, livelli significativamente più alti di quelli fisiologici ottenuti in pazienti comparabili ma HIV-negativi. In un importante studio sulla dinamica dei linfociti CD4+, Hellerstein *et al.* (1999) non osservarono però alcuna differenza fra i due gruppi, mostrando che la previsione *E* risultava smentita e falsificando così l'"ipotesi del lavandino". Riferendosi ai lavori di Ho e colleghi, Hellerstein e collaboratori esplicitarono le conseguenze negative dello studio, rilevando che in base ai loro risultati "alcuni modelli possono essere esclusi" (ivi, p. 86). Nel commentare gli stessi risultati, un altro affermato studioso del settore concluse che essi "pongono fine a quattro anni di avvincente (sebbene spesso aspro) dibattito riguardo all'ipotesi della produzione/distruzione dei linfociti CD4+", vale a dire appunto riguardo all'"ipotesi del lavandino" (si veda Pantaleo 1999).¹⁰

5. Le ipotesi causali e i metodi di Mill nelle scienze mediche

5.1. Le ipotesi causali

Le ipotesi formulate nella ricerca scientifica spesso riguardano relazioni *causali*. Un'ipotesi di questo genere, vale a dire un'*ipotesi causale*, asserisce che eventi di un determinato tipo causano eventi di un altro tipo o, in altri termini, che una determinata *condizione* o *circostanza* è causa di un certo *fenomeno*.

¹⁰ Per una ricostruzione più dettagliata di questo episodio della ricerca sulla patogenesi dell'AIDS, del suo contesto storico e dei suoi sviluppi, rimandiamo a Crupi (2007).

Come ogni altro genere di ipotesi, le ipotesi causali possono figurare come conclusioni di inferenze induttive. L'induzione di ipotesi causali può essere utilmente guidata da *metodi di induzione eliminativa*, così chiamati perché mirano all'eliminazione di tutte le ipotesi causali considerate all'inizio di un processo di indagine ad eccezione di una, che viene così inferita induttivamente in base alle informazioni disponibili. Questi metodi – già definiti da Francesco Bacone (1561-1626) e poi ripresi da John F. W. Herschel (1792-1871) e soprattutto da John Stuart Mill (1806-73) – sono generalmente noti come “metodi di Mill”. Come vedremo a breve, i metodi di Mill forniscono una buona rappresentazione di forme di ragionamento induttivo spontaneamente adottate da coloro che – nella vita quotidiana e nella scienza – si propongono di valutare ipotesi che specificano i rapporti causali fra determinati aspetti, proprietà o fattori qualitativi. A questo proposito occorre segnalare che almeno fino alla metà del Novecento la maggior parte delle ipotesi causali nella ricerca medica avevano carattere qualitativo, e che ancora oggi questo tipo di ipotesi causali svolge un ruolo centrale nelle scienze mediche. Di conseguenza, una forma assai diffusa di ragionamento induttivo tra i ricercatori medici era ed è ancora oggi costituita dall'impiego tacito e spontaneo di schemi di inferenza essenzialmente riconducibili ai metodi di Mill.¹¹

Come si è detto, ipotesi causali della forma $H = \text{“La condizione } C \text{ è causa del fenomeno } F\text{”}$ sono comuni nelle scienze empiriche, ma il loro significato non è affatto privo di ambiguità. Vi sono infatti almeno due modi fondamentalmente diversi in cui H può essere interpretata. H è spesso intesa come l'affermazione che C è condizione *necessaria* per il verificarsi di F , cioè come l'asserzione che, in assenza di C , F non può accadere. Per esempio, affermando che il bacillo di Koch – cioè il *Mycobacterium tuberculosis* – causa la tubercolosi si intende che la tubercolosi non può insorgere senza il bacillo. D'altra parte, H può anche essere intesa come l'affermazione che C è condizione *sufficiente* del verificarsi di F , cioè come l'asserzione che, in presenza di C , F non può non accadere. Per esempio, si può dire che la decapitazione è causa sufficiente della morte.

Come notato da Fletcher e Fletcher (2005, p. 3), la conoscenza delle cause delle malattie – che nei manuali di medicina sono solitamente discusse sotto i titoli di “eziologia”, “patogenesi” e simili – costituisce un essenziale contributo delle scienze mediche alla pratica clinica. Tale conoscenza guida infatti i medici nel loro approccio a fondamentali compiti clinici quali la prevenzione,

¹¹ Non è quindi un caso che Copi e Cohen (2001), uno dei più completi e diffusi manuali di logica informale, illustri i metodi di Mill attraverso circa tre dozzine di esempi tratti dalla storia, anche recente, della medicina.

la diagnosi e il trattamento. Le notevoli implicazioni cliniche della conoscenza delle cause delle malattie stanno alla base del forte interesse per le ipotesi causali che è da sempre un tratto distintivo delle scienze mediche. Si pensi, per esempio, alla scoperta che C è condizione necessaria della malattia F , cioè alla scoperta che F non può presentarsi in assenza di C . Tale scoperta dà fondamento, tra l'altro, alla ricerca terapeutica: infatti, l'eliminazione della malattia F sarebbe garantita dall'identificazione di interventi opportuni in grado di rimuovere la sua condizione necessaria C . Molti successi nel contrasto delle malattie infettive si sono basati precisamente su ricerche di questo genere: si è dapprima scoperto che un certo tipo di germe era condizione necessaria di una malattia, per poi individuare sostanze o strumenti in grado di debellare quel germe così da evitare anche l'insorgenza della malattia.¹²

La distinzione fra condizioni necessarie e condizioni sufficienti rimuove solo in parte l'ambiguità associata all'impiego di ipotesi causali. Essa ci permette comunque di stabilire una stretta relazione fra i due tipi di ipotesi causali e le corrispondenti *ipotesi di regolarità*. A dispetto delle diverse interpretazioni della causalità discusse nella letteratura filosofica (sulle quali ci è impossibile soffermarci in questa sede), vi è infatti ampio consenso sulle due considerazioni che seguono:

- (i) dall'ipotesi causale " C è condizione necessaria di F " consegue l'ipotesi di regolarità "se non si presenta C , allora non si presenta neppure F ";
- (ii) dall'ipotesi causale " C è condizione sufficiente di F " consegue l'ipotesi di regolarità "se si presenta C , allora si presenta anche F ".

Occorre notare che in virtù di (i) e (ii) è possibile eliminare un'ipotesi causale se, sulla base di osservazioni, l'ipotesi di regolarità che ne consegue risulta falsificata. Come emergerà nel seguito, questa semplice considerazione apre la strada all'applicazione dei metodi di Mill nella valutazione di ipotesi causali. Per questo motivo, e per semplicità di esposizione, parleremo d'ora in poi liberamente di "ipotesi causali" anche in riferimento alle corrispondenti ipotesi di regolarità.

¹² Storicamente, il ruolo delle ipotesi causali nella ricerca medica è emerso in modo esplicito e sistematico a partire dalla seconda metà dell'Ottocento nell'ambito dell'approccio metodologico biosperimentale, o fisiopatologico, caratterizzato appunto dal tentativo di elaborare spiegazioni causali delle malattie. A questo riguardo, Corbellini (2007, pp. 26 ss.) nota che i due ricercatori ai quali si deve la sistematizzazione di questo approccio – Claude Bernard (1813-78) e Carl Ludwig (1816-95) – sostengono una concezione del metodo sperimentale in medicina largamente coincidente con quella elaborata da Mill.

5.2. *Il metodo della concordanza*

Nel suo *Sistema di logica deduttiva e induttiva* (1843), Mill delinea cinque metodi induttivi. Qui ci occuperemo solo dei primi due (per molti versi i più fondamentali), noti come metodo della concordanza e della differenza, che si applicano rispettivamente alle condizioni necessarie e alle condizioni sufficienti di un determinato fenomeno. Come vedremo, ciascuno dei due metodi consente di selezionare – come conclusione di un’appropriata inferenza induttiva – una specifica ipotesi causale a dispetto di altre inizialmente considerate.

Supponiamo di avere formulato un’ipotesi iniziale secondo la quale la condizione necessaria del fenomeno F è una fra le condizioni C_1, \dots, C_5 . Possiamo denotare tale ipotesi iniziale come la disgiunzione $D = “H_1, \text{ oppure } \dots, \text{ oppure } H_5”$, dove H_i è l’ipotesi che C_i è condizione necessaria di F . Il metodo della concordanza viene impiegato per eliminare dalla lista C_1, \dots, C_5 delle possibilità inizialmente ammesse da D almeno alcune tra le condizioni non necessarie di F . Eliminare una condizione C_i dal novero delle possibili condizioni necessarie di F significa mostrare che in qualche caso F si presenta anche in assenza di C_i . Di conseguenza, il metodo della concordanza richiede di esaminare casi in cui si presenta F e di controllare, per ciascuna condizione C_i , se essa è presente oppure no. L’applicazione del metodo della concordanza è illustrata dall’acquisizione dell’evidenza E relativa dall’osservazione dei tre casi rappresentati nella Tavola 1, in cui viene registrata, per ciascun caso, la presenza (+) o assenza (–) delle condizioni C_1, \dots, C_5 .

Tavola 1

Caso	Possibili condizioni necessarie					Fenomeno F
	C_1	C_2	C_3	C_4	C_5	
1	+	+	+	–	–	+
2	+	+	–	–	+	+
3	+	–	–	+	+	+

Vediamo che le condizioni C_2, \dots, C_5 si presentano in alcuni casi, ma non in tutti: ciò significa che nessuna di esse è condizione necessaria di F . Possiamo quindi concludere deduttivamente, sulla base dell’evidenza E , che le corrispondenti ipotesi causali H_2, \dots, H_5 sono false. L’unica condizione che non viene eliminata da E è C_1 , poiché essa si presenta in tutti e tre i casi considerati. La circostanza per cui i casi considerati concordano nella presenza di una sola specifica condizione, eliminando in tal modo tutte le altre, è ciò che suggerisce di denominare il primo metodo milliano come “metodo della concordanza”.

Si noti che, anche se E esclude tutte le ipotesi considerate in D tranne la sola H_1 , non è possibile concludere che E verifica H_1 , nel senso di implicarla logicamente. Infatti, resta logicamente possibile che H_1 sia falsa anche se E è vera, dal momento che i tre casi descritti da E non ci permettono in alcun modo di escludere che prima o poi ci imbatteremo in un ulteriore caso in cui la concordanza viene meno, essendo C_1 assente pur in presenza di F . Se ciò accadesse, significherebbe evidentemente che le vere condizioni necessarie di F non erano incluse nella lista contemplata dall'ipotesi iniziale D , per quanto essa potesse apparire inizialmente plausibile. Ciò significa che non è possibile *dedurre* H_1 dall'evidenza empirica E , e quindi che – diversamente da quanto talora suggerito in alcune esposizioni divulgative dei metodi di Mill – il metodo della concordanza *non* consente di raggiungere, sulla base dell'evidenza empirica, la certezza che una determinata ipotesi causale è vera. Ciò che invece il metodo consente di elaborare è un'inferenza induttiva, e più precisamente – come ora mostreremo – un particolare tipo di conferma relativizzata.

In molti casi analoghi a quello che abbiamo appena descritto, sembra naturale sottoscrivere le seguenti valutazioni comparative di plausibilità:

(i) Alla luce della conoscenza di sfondo S , D è più plausibile di H_1 .

In riferimento a (i) è sufficiente osservare che D contempla una lista di possibili condizioni necessarie di F più estesa della sola C_1 indicata da H_1 .

(ii) D è almeno altrettanto plausibile alla luce di $(S \& E)$ di quanto lo era alla luce di S .

In riferimento a (ii), si noti che E non falsifica D né, più in generale, fornisce ragioni per ritenerla meno plausibile.

(iii) Alla luce di $(S \& E)$, D è tanto plausibile quanto lo è H_1 .

In riferimento a (iii), si noti che l'evidenza E “restringe” le ipotesi ammesse da D alla sola H_1 , escludendo tutte le altre.

È facile vedere che (i)-(iii) hanno come conseguenza:

(iv) H_1 è più plausibile alla luce di $(S \& E)$ che alla luce di S .

Notiamo a questo punto che, come si è chiarito più sopra (si veda il paragrafo

2.2), sottoscrivere l'asserzione (iv) equivale a effettuare la seguente inferenza di conferma relativizzata:

$$\begin{array}{l} 0. \quad S \\ 1. \quad E \\ \quad = \\ \quad H_1. \end{array}$$

Prima di passare all'esposizione del metodo milliano della differenza, presenteremo sinteticamente un esempio di applicazione del metodo della concordanza nella ricerca medica.

ESEMPIO 5. *Le cause genetiche del morbo di Alzheimer.* Dopo avere esaminato centinaia di famiglie colpite dal morbo di Alzheimer, un gruppo di ricerca dell'Università di Washington è riuscito a eliminare una dopo l'altra tutte le possibili cause genetiche del morbo inizialmente ipotizzate – in base alla considerazione che ciascuna di esse era assente in alcuni pazienti – con l'eccezione di una, costituita dalla presenza di una piccola zona con caratteristiche peculiari sul cromosoma 14 (Schellenberg *et al.* 1992). Risultò in tal modo confermato il ruolo causale di questa anomalia, presente solo negli ammalati, come condizione necessaria del morbo.¹³

5.3. Il metodo della differenza

Supponiamo ora di avere formulato un'ipotesi iniziale secondo la quale la condizione sufficiente del fenomeno F è una fra le condizioni C_1, \dots, C_5 . Anche in questo caso, denotiamo tale ipotesi iniziale come la disgiunzione $D = "H_1, \text{ oppure } \dots, \text{ oppure } H_5"$, dove H_i è l'ipotesi che C_i è condizione sufficiente di F . Il metodo della differenza viene impiegato per eliminare dalla lista C_1, \dots, C_5 delle possibilità inizialmente ammesse da D almeno alcune tra le condizioni non sufficienti di F . Eliminare una condizione C_i dal novero delle possibili condizioni sufficienti di F significa mostrare che in qualche caso, pur presenza di C_i , F non si presenta. Di conseguenza, il metodo della differenza richiede di esaminare casi in cui *non* si presenta F e di controllare, per ciascuna condizione C_i , se essa è presente oppure no. L'applicazione del metodo della differenza è

¹³ L'esempio è tratto da Copi e Cohen (2001, p. 498), che lo impiegano con scopi simili.

illustrata dall'acquisizione dell'evidenza E relativa dall'osservazione dei tre casi rappresentati nella Tavola 2.

Caso	Possibili condizioni sufficienti					Fenomeno
	C_1	C_2	C_3	C_4	C_5	F
1	–	–	–	+	+	–
2	–	–	+	+	–	–
3	–	+	+	–	–	–

Si può vedere che la Tavola 2 si ottiene dalla Tavola 1 sostituendo, in ogni occorrenza, “+” con “–”, e viceversa. Avremmo, naturalmente, potuto considerare una tavola diversa. Il riferimento a due tavole speculari, tuttavia, sottolinea la somiglianza strutturale fra il metodo della concordanza e il metodo della differenza. Esso permette altresì di riprodurre immediatamente, *mutatis mutandis*, le osservazioni fatte a proposito della Tavola 1, con le seguenti conclusioni: (i) l'evidenza E rappresentata nella Tavola 2 conduce all'eliminazione di tutte le possibili condizioni sufficienti considerate dall'ipotesi disgiuntiva D tranne C_1 ; (ii) l'ipotesi causale H_1 – secondo la quale C_1 è condizione sufficiente di F – è induttivamente confermata da E (relativamente alla conoscenza di sfondo S).

Nell'esempio della Tavola 2, C_1 è assente in tutti e tre i casi considerati. Ciò significa che i tre casi concordano tra di loro per l'assenza di C_1 . Non risulta quindi per nulla evidente il significato dell'espressione “metodo della differenza”, coniata da Mill. Tale espressione è invece suggerita dal particolare tipo di esempi – rappresentato nella Tavola 3 – che Mill impiega per illustrare il suo metodo per la ricerca di cause sufficienti.

Caso	Possibili condizioni sufficienti					Fenomeno
	C_1	C_2	C_3	C_4	C_5	F
1	+	+	+	+	+	+
2	–	+	+	+	+	–

Vediamo qui che la sola *differenza* tra il caso in cui il fenomeno F si verifica (caso 1) e quello in cui non si verifica (caso 2) è costituita dalla presenza di C_1 nel primo e dalla sua assenza nel secondo. Con riferimento a esempi del tipo

illustrato nella Tavola 3, Mill formula il metodo della differenza in questi termini: se abbiamo osservato un caso in cui si verifica il fenomeno F e un altro in cui non si verifica, e se la sola differenza è la presenza di una determinata condizione C_i nel primo soltanto dei due casi considerati, allora possiamo indurre la corrispondente ipotesi causale H_i , secondo la quale C_i è condizione sufficiente di F . Mill delinea quindi il metodo della differenza con riferimento a un particolare genere di evidenza, caratterizzata dall'osservazione di due soli casi. Occorre comunque notare che tale formulazione è pienamente compatibile con quella da noi presentata in precedenza. L'osservazione del caso 2, infatti, conduce all'eliminazione delle ipotesi H_2, \dots, H_5 , confermando così H_1 . A nostro avviso, il caso 1 – che pure risulta logicamente ridondante, nel senso che non conduce all'eliminazione di alcuna delle possibili condizioni sufficienti considerate – viene comunque presentato da Mill per sottolineare che l'ipotesi iniziale è appunto D e che essa è compatibile con le osservazioni svolte.

Il seguente esempio illustra un'importante classe di applicazioni del metodo della differenza nelle scienze mediche.

ESEMPIO 6. *Il metodo degli interventi simulati.* Nelle scienze biomediche un procedimento tradizionalmente impiegato per comprendere le funzioni fisiologiche di un organo consiste nella sua rimozione chirurgica effettuata in animali da laboratorio e nell'attenta osservazione degli effetti che ne conseguono. La rimozione chirurgica, tuttavia, è normalmente accompagnata da alcune circostanze collaterali, quali la concomitante lesione o asportazione di altri organi durante l'intervento e lo stress postoperatorio. Di conseguenza, è difficile stabilire se gli effetti osservati dipendano dall'una o dall'altra fra le possibili condizioni sufficienti che si sono determinate con l'intervento, vale a dire dalla rimozione dell'organo in questione o da qualcuna delle circostanze concomitanti. Per far fronte a questo problema, Claude Bernard propose il suo metodo degli *interventi simulati*. Per stabilire se la scomparsa (o l'attenuazione) di un certo processo fisiologico riscontrato in una animale dovesse essere attribuita alla rimozione di un determinato organo o a qualcuna delle altre circostanze collaterali associate all'intervento chirurgico, Bernard suggerì di prendere un animale del tutto simile sul quale effettuare un intervento diverso dal primo per la mancata rimozione dell'organo, ma identico a esso per tutti gli altri aspetti rilevanti (modalità di incisione, durata ecc.). Indichiamo quindi con F la scomparsa (o attenuazione) di un certo processo fisiologico, con R la rimozione dell'organo e con C_1 - C_3 le circostanze collaterali associate all'intervento chirurgico. Supponiamo poi che i risultati ottenuti nell'intervento di rimozione e in quello simulato siano compendati nella Tavola 4.

Tavola 4

Caso	Possibili condizioni sufficienti			Effetti fisiologici	
	R	C_1	C_2	C_3	F
1. Intervento di rimozione	+	+	+	+	+
2. Intervento simulato	-	+	+	+	-

Secondo il metodo milliano della differenza, i dati della Tavola 4 conducono all'eliminazione di C_1 - C_3 come possibili condizioni sufficienti di F , confermando induttivamente l'ipotesi causale che tale effetto vada attribuito a R , cioè alla rimozione dell'organo.¹⁴

6. Osservazioni conclusive

L'idea, oggi largamente accettata, che la filosofia della scienza debba basarsi su una conoscenza approfondita della storia delle scienze ha spinto gli studiosi ad analizzare le procedure, i problemi e i fondamenti delle singole discipline scientifiche. In tal modo, a partire dall'ambito della filosofia *generale* della scienza – rivolta a problemi metodologici di ampia portata relativi a qualunque indagine scientifica – si sono ramificate e sviluppate diverse filosofie *speciali* delle scienze, quali la filosofia della fisica, la filosofia della biologia o la filosofia dell'economia.¹⁵ A queste pare del tutto naturale affiancare anche la *filosofia delle scienze mediche*.¹⁶ Di fatto, però, quasi nessuno dei più importanti filosofi della scienza del Novecento ha discusso dettagliatamente esempi illustrativi delle forme di ragionamento usate dai ricercatori medici. Si tratta di una circostanza alquanto sorprendente, considerato che la filosofia generale della scienza, per sua natura, può utilmente avvalersi di esempi tratti dai più disparati settori della ricerca scientifica. Per di più, a quanto ci risulta,

¹⁴ Sulle relazioni concettuali fra il metodo della differenza di Mill e le considerazioni metodologiche di Bernard si veda Schaffner (1993, pp. 145-152). Schaffner nota che la raccomandazione, su cui Bernard insisteva con particolare vigore, di effettuare sempre interventi simulati nelle indagini sulle funzioni degli organi è stata ampiamente accolta. L'intervento simulato costituisce infatti una procedura di controllo ormai diffusa in questo tipo di ricerche. Negli anni sessanta del Novecento è stata utilizzata, per esempio, nelle indagini sperimentali sulle funzioni del timo.

¹⁵ Per un'utile panoramica, si veda Vassallo (2003).

¹⁶ Si vedano Corbellini (2003), Federspil *et al.* (2008), Giaretta *et al.* (2008), Wulff, Stig e Rosenberg (1990).

un'indagine sistematica delle forme di ragionamento coinvolte nella ricerca medica non è ancora stata tentata neanche dagli stessi studiosi di filosofia delle scienze mediche.¹⁷ Per questi motivi, la trattazione qui presentata deve essere considerata ampiamente preliminare, e tale da lasciare inevitabilmente ai margini molti temi che altrimenti meriterebbero un'analisi più approfondita. In conclusione, vorremmo segnalarne soprattutto due.

In primo luogo, concentrandoci sulla *valutazione* di ipotesi, abbiamo tralasciato di discutere le forme di ragionamento eventualmente coinvolte nella *scoperta* – cioè nella generazione ed elaborazione – delle ipotesi. La possibilità di una “logica della scoperta” è tema di dibattito ricorrente in filosofia della scienza, che si ripropone nelle scienze mediche come nelle altre discipline. Il lettore interessato può fare riferimento ad alcuni classici testi sull'argomento, come Nickles (1980) e Simon (1977), e a Schaffner (1993, cap. 2) e Thagard (1999) per trattazioni più direttamente legate alla medicina.

In secondo luogo, abbiamo limitato la nostra discussione a ipotesi che potessero essere interpretate ed espone in termini qualitativi. Siamo ben consapevoli, d'altra parte, che le ipotesi di carattere quantitativo – e in particolare le ipotesi statistiche – rivestono un interesse significativo per la scienza medica contemporanea, soprattutto a seguito dell'avvento e degli sviluppi dell'epidemiologia clinica. Basti pensare alle ipotesi statistiche relative alla prevalenza di una determinata patologia, alla sensibilità e specificità di un test diagnostico o alle correlazioni fra una certa condizione – come un fattore di rischio, o un trattamento terapeutico – e le sue possibili conseguenze. Riteniamo comunque che la valutazione di ipotesi quantitative di questo tipo possa essere ricostruita attraverso opportune estensioni e raffinamenti delle forme di ragionamento qui espone.¹⁸ Resta inteso, tuttavia, che tale suggerimento dovrà trovare attuazione in altra sede.

RIFERIMENTI BIBLIOGRAFICI

- ANTISERI, D. (1977): “Introduzione”, nell'edizione italiana di Semmelweis (1861).
 ASCHER, M.S. *et al.* (1995): “Paradox Remains”, *Nature*, 375, p. 196.

¹⁷ Il punto di riferimento più ricco di spunti resta a tutt'oggi Schaffner (1993) (si vedano in particolare i capp. 4 e 5).

¹⁸ Pensiamo, in particolare, all'arricchimento degli strumenti logici qui impiegati attraverso l'esplicito riferimento alla teoria della probabilità in una prospettiva bayesiana. Per un'introduzione ormai classica alla filosofia della scienza bayesiana, si veda Howson e Urbach (1993), a cui ci permettiamo di aggiungere Festa (1996) e (1999).

- CÉLINE L. F. (1952): *Semmelweis*, Paris: Gallimard (trad. it. *Il dottor Semmelweis*, Milano: Adelphi 2006).
- COPI, I. M. e COHEN, C. (2001): *Introduction to Logic*, New York: MacMillan (trad. it. *Introduzione alla logica*, Bologna: Il Mulino 2002).
- CORBELLINI, G. (2003): “Filosofia della medicina”, in Vassallo (2003).
 ——— (2007): *EBM. Medicina basata sull’evoluzione*, Roma-Bari: Laterza.
- CRUPI, V. (2007): “The Sink and the Murder Scene: Rise and Fall of a Causal Model in AIDS Pathogenesis”, *L&PS – Logic & Philosophy of Science*, 5, pp. 9-32.
- FEDERSPIL, G. (1980): *I fondamenti del metodo in medicina clinica e sperimentale*, Padova: Piccin.
- FEDERSPIL, G. et al. (a cura di) (2008): *Filosofia della medicina*, Milano: Raffaello Cortina.
- FESTA, R. (1996): *Cambiare opinione. Temi e problemi di epistemologia bayesiana*, Bologna: CLUEB.
 ——— (1999): “Bayesian Confirmation”, in M. C. Galavotti e A. Pagnini (a cura di), *Experience, Reality, and Scientific Explanation*, Dordrecht: Kluwer.
- FLETCHER, R. H. e FLETCHER, S. W. (2005): *Clinical Epidemiology: The Essentials*, Philadelphia: Lippincott Williams & Wilkins.
- GIARETTA, P. et al. (a cura di) (2009): *Filosofia della medicina*, Bologna: Il Mulino.
- GILLIES, D. (2005): “Hempelian and Kuhnian Approaches in the Philosophy of Medicine: The Semmelweis Case”, *Studies in the History and Philosophy of Biological and Biomedical Sciences*, 36, pp. 195-181.
- GIORELO, G. e MORIGGI, S. (2004): “Tra diagnosi e scoperta. Una rilettura del caso Semmelweis”, in G. Federspil e P. Giaretta (a cura di), *Forme della razionalità medica*, Soveria Mannelli: Rubbettino.
- HACKING, I. (2001): *An Introduction to Probability and Inductive Logic*, Cambridge: Cambridge University Press (trad. it. *Introduzione alla probabilità e alla logica induttiva*, Milano: Il Saggiatore 2005).
- HELLERSTEIN, M. et al. (1999): “Directly Measured Kinetics of Circulating T-lymphocytes in Normal and HIV-1 Infected Humans”, *Nature Medicine*, 5, pp. 83-89.
- HEMPEL, C. G. (1966): *Philosophy of Natural Science*, Englewood Cliffs (N.J.): Prentice Hall (trad. it. *Filosofia delle scienze naturali*, Bologna: Il Mulino 1980).
- HO, D.D. et al. (1995): “Rapid Turnover of Plasma Virions and CD4 Lymphocytes in HIV-1 Infection”, *Nature*, 373, pp. 123-126.
- HOWSON, C. e URBACH, P. (1993): *Scientific Reasoning: The Bayesian Approach*, La Salle (Ill.): Open Court.

- HURLEY, P. J. (1999): *A Concise Introduction to Logic*, Belmont (Calif.): Wadsworth.
- JOHANSSON, I. e LYNØE N. (2009): *Medicine and Philosophy: A Twenty-First Century Introduction*, Frankfurt a.M.: Ontos Verlag.
- LIPTON, P. (2004), *Inference to the Best Explanation*, 2nd edition, London: Routledge.
- MADDOX, J. (1995): “More Conviction on HIV and AIDS”, *Nature*, 377, p. 1.
- NICKLES, T. (a cura di) (1980): *Scientific Discovery: Case Studies*, Dordrecht: Reidel.
- NIINILUOTO, I. (1999): “Defending abduction”, *Philosophy of Science*, 66 Supplement, pp. S436-S451.
- PANTALEO, G. (1999): “Unraveling the strands of HIV’s Web”, *Nature Medicine*, 5, pp. 27-28.
- PERA, M. (1983): “Per una teoria induttivista della diagnosi clinica e della scoperta medica”, in C. Scandellari e G. Federspil (a cura di), *Scoperta e diagnosi in medicina. Discussioni sul metodo*, Padova: Piccin.
- SALMON, W. C. (1983): *Logic*, Englewood Cliffs (N.J.): Prentice Hall.
- SCHAFFNER, K. F. (1993): *Discovery and Explanation in Biology and Medicine*, Chicago (Ill.): The University of Chicago Press.
- SCELLENBERG, G. D. *et al.* (1992): “Genetic Linkage Evidence for a Familial Alzheimer’s Disease Locus on Chromosome 14”, *Science*, 258, pp. 668-671.
- SEMMELWEIS, I. F. (1861): *Die Ätiologie, der Begriff und die Prophylaxis des Kindbettfiebers*, Wien-Leipzig: Hartlebens Verlag (trad. it. *Come lavora uno scienziato. Eziologia, concetto e profilassi della febbre puerperale*, Roma: Armando 1977).
- SIMON, H. (1977): *Models of Discovery and Other Topics in the Methods of Science*, Dordrecht: Reidel.
- SHEPPARD, H. W., ASCHER, M. S. e KROWKA, J. F. (1993): “Viral Burden and HIV Disease”, *Nature*, 364, p. 291.
- SKYRMS, B. (1999): *Choice and Chance: An Introduction to Inductive Logic*, 4th edition, Belmont (Calif.): Wadsworth (trad. it. della prima edizione *Introduzione alla logica induttiva*, Bologna: Il Mulino 1974).
- THAGARD, P. (1999): *How Scientists Explain Disease*, Princeton (N.J.): Princeton University Press (trad. it. *La spiegazione scientifica della malattia*, Milano: McGraw-Hill 2000).
- VARZI, A., NOLT, J. e ROHATYN, D. (1998): *Logic*, New York: McGraw-Hill (trad. it. *Logica*, Milano: McGraw Hill 2007).
- VASSALLO, N. (a cura di) (2003): *Filosofie delle scienze*, Torino: Einaudi.
- VINEIS, P. (1999): *Nel crepuscolo della probabilità. La medicina tra scienza ed etica*, Torino: Einaudi.

WEI, X. *et al.* (1995): “Viral Dynamics in Human Immunodeficiency Virus Type 1 Infection”, *Nature*, 373, pp. 117-122.

WULFF, H.R., STIG, A.P. e ROSENBERG, R. (1990): *Philosophy of Medicine: An Introduction*, Oxford: Blackwell (trad. it. *Filosofia della medicina*, Milano: Raffaello Cortina 1995).

Dai giochi d'azzardo ai test diagnostici: la teoria della decisione nella medicina clinica*

Vincenzo Crupi

Dipartimento di Area Critica Medico-chirurgica, Università di Firenze
e-mail: vincenzo.crupi@unifi.it

Roberto Festa

Dipartimento di Filosofia, Università di Trieste
e-mail: festa@units.it

1. La teoria della scelta razionale
2. Decisioni terapeutiche
3. Decisioni diagnostiche
4. Osservazioni conclusive

SOMMARIO. L'obiettivo del presente contributo è introdurre gli aspetti centrali dello studio delle decisioni, applicandoli a problemi illustrativi della medicina clinica. Vedremo quindi quali siano gli elementi che definiscono un problema decisionale e ne forniscono una rappresentazione adeguata. Mostreremo inoltre come sia possibile sottoporre ad analisi tali elementi, identificando in ciascuna situazione il corso d'azione ottimale, vale a dire quello che permette al decisore di perseguire con maggiore efficacia gli obiettivi che si propone. Nel corso della trattazione, discuteremo altresì quali siano le origini, il significato e i limiti dell'analisi decisionale applicata alla medicina clinica.

PAROLE CHIAVE: razionalità, decisione, diagnosi, terapia.

* Desideriamo ringraziare Giovanni Federspil, Pierdaniele Giaretta e Katya Tentori per fruttuose discussioni sui temi trattati in questo contributo.

*In zweifelhaften Fällen,
sollte man sich stets für das richtige entscheiden.¹*

K. KRAUS

1. La teoria della scelta razionale

Per gli esseri umani, decidere è quasi come nutrirsi: bene o male, non possono farne a meno. L'esigenza di prendere decisioni ci è imposta dalla limitatezza delle risorse a nostra disposizione, a cominciare dal tempo e dal denaro. Dalla più piccola delle scelte quotidiane a quella di seguire un certo corso di studi, dall'acquisto di un determinato modello di automobile fino al voto politico, il problema ricorrente è la necessità di individuare uno specifico corso d'azione da perseguire a dispetto di altri possibili.

Certo, l'esperienza comune mostra quanto le decisioni possano essere fra loro diverse sotto molti punti di vista. Possono avere conseguenze più o meno serie. Possono richiedere la considerazione di pochi dati semplici e chiari, o di una gran quantità di informazioni complesse e di difficile interpretazione. Eppure, nonostante questa innegabile varietà, è spesso possibile analizzare le decisioni in base a un quadro concettuale unitario e ormai consolidato, noto appunto come *teoria delle decisioni* o *teoria della scelta razionale*.

I primi elementi della teoria della scelta razionale si possono rintracciare nell'opera del filosofo e matematico Blaise Pascal (1623-62), che è anche uno dei padri riconosciuti della *teoria della probabilità*. (Come si vedrà, esistono importanti legami fra le due teorie.) Le indagini di Pascal su questo tema furono ispirate da preoccupazioni che all'epoca impegnavano i giocatori d'azzardo, particolarmente sensibili – per ovvie ragioni – all'elaborazione di metodi affidabili per determinare quanto fossero vantaggiose le scommesse.² Pascal, che avendone posto le fondamenta aveva intuito le enormi potenzialità della teoria, non esitò a presentarne un'applicazione divenuta celebre: analizzò appunto in termini di scommessa una scelta che ogni uomo è chiamato a fare, mettendo letteralmente in gioco la sua vita. Nel paragrafo 233 dei suoi *Pensieri*, sostenne infatti che adottare una condotta di vita cristiana è la scelta più razionale anche per chi si senta attratto dai piaceri del libertinaggio. Il suo argomento ha una forma matematica: tutto considerato, sostiene Pascal, “scommettere su Dio” presenta un bilancio di benefici e costi attesi enormemente più

¹ “In caso di dubbio, bisognerebbe sempre scegliere la cosa giusta.”

² Per un'avvincente ricostruzione storica, si veda Devlin (2008).

vantaggioso della scelta opposta, perché la possibilità di ottenere l'eterna beatitudine – per quanto incerta – ha un valore incommensurabile.³ Come avremo modo di vedere più nei dettagli, la teoria della scelta razionale ha continuato a servirsi comunemente di questo parallelismo fra scommesse e problemi decisionali della vita reale.

Successivi contributi di figure centrali della storia della filosofia e della matematica, come Jeremy Bentham (1748-1832) e Jakob Bernoulli (1654-1705), testimoniano il persistente intreccio fra le due discipline che caratterizza l'elaborazione dei principi razionali della decisione. Ancora negli anni venti del Novecento, sono due giovani geni della filosofia e della matematica a mettere a punto in una forma compiuta gli elementi fondamentali della teoria. Nel 1926, il ventitreenne Frank Plumpton Ramsey (1903-30), originale seguace di Russell e Wittgenstein, elabora la prima trattazione assiomatica della scelta in condizioni di incertezza in un importante articolo rimasto inedito fino alla sua prematura scomparsa (Ramsey 1931). Due anni più tardi, il suo coetaneo ungherese John von Neumann (1903-57) pubblica una teoria matematica dei giochi in cui il comportamento strategico ottimale emerge da un calcolo basato sulle conoscenze e gli obiettivi di agenti perfettamente razionali (von Neumann 1928). Per veder maturare i frutti di questi lavori inaugurali sarà però necessario attendere la metà del secolo, con la pubblicazione dei due epocali volumi *Theory of Games and Economic Behavior*, ancora di von Neumann con la collaborazione dell'economista austriaco Oskar Morgenstern (1902-77) (von Neumann e Morgenstern 1944/47), e *The Foundations of Statistics* dello statistico Leonard Savage (1917-71) (Savage 1954). È da notare, peraltro, che alcuni dei principi della teoria avevano nel frattempo trovato una originale trattazione nell'opera del grande matematico italiano Bruno De Finetti (1906-85) (De Finetti 1937), del quale Savage non esitò a considerarsi seguace.

Nella scienza contemporanea, la teoria della scelta razionale è divenuta il cuore della cosiddetta “economia neoclassica” (si vedano in proposito Motterlini e Piattelli Palmarini 2005; Motterlini 2006). Essa si è altresì rivelata un importante punto di riferimento per una varietà di discipline interessate al comportamento, dalla psicologia cognitiva alle scienze politiche, dalle neuroscienze fino all'etologia, arrivando infine a toccare anche lo studio delle decisioni

³ Al di là del suo perdurante interesse filosofico (si veda Jordan 1994), la “scommessa di Pascal” è rimasta un vivido caso illustrativo della teoria delle decisioni, ripreso perciò dalla manualistica, come in Baron (2000, p. 230). Del resto, il filosofo e storico della scienza Ian Hacking ha identificato proprio nell'argomento di Pascal “il primo contributo convincente alla teoria delle decisioni” (Hacking 1975, p. viii).

cliniche. Il nesso fra comportamento economico e medicina clinica può apparire sorprendente, ma lo è solo a prima vista. È tipico infatti degli agenti economici (consumatori, imprenditori, investitori) dover selezionare il più vantaggioso fra corsi di azione alternativi, ponderando i loro possibili costi e benefici. Ebbene, numerose situazioni cliniche rivelano caratteristiche del tutto analoghe, indicando così come strumenti rigorosi e generali per l'analisi dei problemi decisionali risultino altamente rilevanti per riflettere sulla pratica della medicina. Ciò che può sorprendere, piuttosto, è il relativo ritardo con cui lo studio delle decisioni è stato recepito nella riflessione sulla medicina clinica. Se si fa eccezione per alcuni contributi pionieristici ancorché fondamentali (Ledley e Lusted 1959; Weinstein e Fineberg 1980) si può dire che una più diffusa consapevolezza della rilevanza di questi temi si sia affermata solo in tempi recenti (si vedano Chapman e Sonnenberg 2000; Hunink *et al.* 2001; Motterlini e Crupi 2005; Scandellari 2005), in parte assecondata dall'avvento della cosiddetta *evidence-based medicine* (Sackett *et al.* 1997; Shulkin 2000; Festa 2004; Crupi 2005).

Gli sviluppi appena descritti, seppur tardivi, appaiono pienamente motivati. Secondo una classificazione diffusa, infatti, l'erogazione di cure mediche coinvolge nel complesso almeno *tre* grandi classi di problemi decisionali, collocate a livelli differenti (Sutherland e Till 1993). Il livello più generale riguarda l'adozione di linee guida o di interi programmi di intervento (per esempio, un programma di screening) da parte di grandi organizzazioni sanitarie che si rivolgono alla società nel suo insieme. A un livello intermedio, si può collocare l'identificazione di procedure stabilite localmente (per esempio, da una singola struttura ospedaliera) per specifici gruppi di pazienti. Infine, al cuore stesso della medicina clinica si trovano le innumerevoli scelte che determinano l'azione dei medici nell'interazione con i singoli pazienti. La tripartizione è indicativa e non esclude, evidentemente, che si diano legami o sovrapposizioni fra i diversi livelli. Essa ci permette comunque di precisare che la trattazione presentata di seguito verterà principalmente sulla terza classe di problemi decisionali, quelli relativi al percorso diagnostico-terapeutico di singoli pazienti.

Per introdurre il piano della nostra esposizione, è utile aggiungere che, a partire dalla prima presentazione di un problema clinico, medico e paziente si trovano di fronte a tre alternative fondamentali. La prima è, semplicemente, *non fare nulla*. Si tratta di una prospettiva solitamente poco attraente, che può persino apparire in contrasto con la vocazione stessa della medicina clinica; ciò non toglie, tuttavia, che resti una possibilità in linea di principio sempre aperta, e anche meritevole di considerazione almeno in situazioni specifiche. La seconda possibilità è quella di elaborare una scelta *terapeutica*, vale a dire orientata alla somministrazione di un particolare trattamento fra quelli dispo-

nibili. Solitamente, si ritiene indicata una scelta di questo genere a condizione che il quadro diagnostico appaia sufficientemente chiaro. Se così non è, risulta spesso opportuno considerare una terza strategia, vale a dire rinviare l'eventuale trattamento a dopo la raccolta di informazioni cliniche aggiuntive (per esempio, attraverso l'esecuzione di un esame di laboratorio), che qui indicheremo come l'elaborazione di una scelta *diagnostica*. Lo schema appena delineato, e in particolare la distinzione fra scelte terapeutiche e scelte diagnostiche, determina una naturale suddivisione dei problemi decisionali della medicina clinica che guiderà la nostra discussione. Dedicheremo quindi il prossimo paragrafo alle decisioni terapeutiche, rivolgendoci successivamente a quelle diagnostiche.

2. Decisioni terapeutiche

2.1. Scelte in condizioni di certezza: ponderare costi e benefici

Si è detto di come la teoria della scelta razionale sia stata considerata centrale nello studio dei fenomeni economici. Non è un caso, quindi, che il contesto delle transazioni commerciali rappresenti un ambito di applicazione privilegiato della teoria e fornisca efficaci illustrazioni dei suoi principi. Supponiamo, per esempio, che siate interessati all'acquisto di una casa e che abbiate a disposizione due *opzioni* alternative, vale a dire l'acquisto di uno fra gli appartamenti *A* e *B*. Supponiamo anche che le uniche differenze rilevanti fra *A* e *B* riguardino la metratura e il prezzo di vendita (si tratta, evidentemente, di una semplificazione alquanto drastica, introdotta a scopi puramente illustrativi). L'appartamento *A* è, poniamo, di 100 mq e costa 300 mila euro; *B* è di 75 mq e costa 200 mila euro. Le opzioni *A* e *B* si distinguono quindi per i loro *esiti*: perseguendo l'una o l'altra, vi troverete a vivere in un ambiente più o meno spazioso affrontando una spesa più o meno impegnativa.

È naturale supporre che la vostra scelta dipenderà da quanto le due opzioni vi appaiono attraenti in base alla desiderabilità dei loro esiti. La teoria della scelta razionale rappresenta questo tipo di valutazione attribuendo al decisore una *funzione di utilità* (solitamente indicata con *U*) definita sugli esiti delle opzioni fra cui scegliere. L'utilità complessiva di un'opzione sarà poi data dalla somma delle utilità dei suoi esiti (Keeney e Raiffa 1976), cosicché:

$$(1) \quad \begin{aligned} U(A) &= U(100 \text{ mq}) + U(\text{spesa di } 300 \text{ mila euro}) \\ U(B) &= U(75 \text{ mq}) + U(\text{spesa di } 200 \text{ mila euro}) \end{aligned}$$

È bene chiarire che i valori di utilità sono interpretati in termini del tutto soggettivi, tali quindi da riflettere i desideri e gli obiettivi di un particolare decisore, che possono differire notevolmente da quelli di un altro. La teoria non impone alcun vincolo a questo livello. Ciò che essa prevede, piuttosto, è che un decisore, se agisce razionalmente, sceglierà invariabilmente l'opzione che *massimizza* la sua utilità. Sceglierà di acquistare l'appartamento *A*, per esempio, se la somma dell'utilità dei suoi esiti è maggiore di quella degli esiti di *B*. In condizioni ordinarie, naturalmente, l'utilità soggettiva crescerà a fronte di una metratura maggiore e di un prezzo più conveniente (cioè più basso). Nel problema degli appartamenti *A* e *B*, tuttavia, non è possibile ottenere il meglio su entrambi i fronti. È invece necessario stabilire quale dei due aspetti far prevalere. In termini tecnici, si dice che il problema decisionale presenta un *trade-off* fra i due attributi rilevanti (l'ampiezza degli appartamenti e la loro convenienza).⁴

Per i nostri presenti scopi è di particolare interesse sottolineare come i *trade-off* siano moneta corrente nei problemi clinici. Nella scelta fra trattamenti alternativi per una condizione patologica nota, in particolare, è comune dover confrontare opzioni alternative che si differenziano per attributi quali, per esempio, la facilità di assunzione e la presenza di effetti collaterali, come accade nello scenario seguente (adattato da Schwartz e Chapman 1999, p. 150):

Una paziente di ventisette anni lamenta per la prima volta prurito, bruciore e perdite vaginali. L'esame fisico determina una diagnosi di candidiasi vaginale. Si considerino le due seguenti opzioni di trattamento, entrambe efficaci dal punto di vista terapeutico.

Tioconalozo (es. Vagistat 1): crema intravaginale da applicare in una singola dose; comporta un temporaneo aggravamento dei sintomi (es. prurito).

Clotrimazolo (es. Mycelex 7): crema intravaginale da applicare per sette giorni consecutivi; non comporta effetti collaterali.

⁴ È tanto semplice spiegare che cosa sia un *trade-off* quanto è difficile tradurre il termine in italiano, al punto che non sembra esservi alcuna traduzione generalmente accettata. In certi contesti, può essere utile riferirsi alla nozione di "conflitto" decisionale; il conflitto, cioè, fra due possibili ragioni per la scelta che spingono in direzioni opposte. Un altro termine talvolta impiegato è quello di "compromesso", come nell'espressione "compromesso qualità-prezzo". Un'ulteriore possibilità è quella di fare riferimento al termine "equilibrio", considerando che il decisore deve valutare quale fra le opzioni di scelta realizzi l'"equilibrio" più soddisfacente fra gli attributi rilevanti.

Proprio come nell'esempio non medico da cui abbiamo preso le mosse, si tratta qui di ponderare i vantaggi e gli svantaggi associati ai due corsi di azione possibili, vale a dire se il guadagno di utilità implicato da una più semplice modalità di somministrazione sia maggiore o minore di quello associato all'assenza di effetti collaterali.

2.2. Scelte in condizioni di incertezza: l'utilità attesa

Uno degli aspetti più interessanti della teoria della scelta razionale è che può essere elegantemente estesa a situazioni più complesse di quelle considerate finora, perché tali da comprendere elementi di *incertezza* o di *rischio*. Per illustrare questo tipo di casi, è comune rappresentare i problemi di scelta in relazione a diverse possibili scommesse. Facciamo ancora un esempio di scelta fra due opzioni.

- (A) Vinci 100 euro se ottieni un sei con un dado bilanciato; altrimenti non vinci nulla.
- (B) Vinci 20 euro se ottieni testa con una moneta bilanciata; altrimenti non vinci nulla.

Quale delle due scommesse sceglieresti di giocare?

In questo scenario, gli esiti possibili delle due opzioni sono in tutto tre: non vincere nulla, vincere 20 euro oppure vincerne 100. A ciascuno di essi un particolare decisore potrà attribuire un certo valore di utilità. Le utilità in questione, tuttavia, non sono di per sé sufficienti per una valutazione complessiva delle due opzioni, perché gli esiti di queste ultime non sono *certi*, bensì dipendenti dall'avverarsi di stati di cose (il risultato del lancio del dado, oppure della moneta) che non sono sotto il controllo del decisore. In questi casi, la quantità che orienta la scelta razionale è detta utilità *attesa* (solitamente indicata con *EU*, per *expected utility*). L'utilità attesa di un'opzione è definita come la somma delle utilità dei suoi possibili esiti, ciascuna moltiplicata per la corrispondente probabilità. Secondo questa definizione:

$$\begin{aligned} (2) \quad EU(A) &= [1/6 \times U(100 \text{ euro})] + [5/6 \times U(0 \text{ euro})] \\ EU(B) &= [1/2 \times U(20 \text{ euro})] + [1/2 \times U(0 \text{ euro})]. \end{aligned}$$

Se i valori di utilità fossero identici alle corrispondenti vincite monetarie, la scommessa *A* risulterebbe maggiormente vantaggiosa, cioè associata a una più

elevata utilità attesa (per rendersene conto, è sufficiente notare che $1/6 \times 100$ è maggiore di $1/2 \times 20$). Occorre precisare, tuttavia, che dato il carattere soggettivo della funzione di utilità non c'è motivo per cui tale rapporto di identità debba necessariamente valere (e anzi è in contrasto con le effettive valutazioni di molte persone). Se per esempio vi interessasse soltanto guadagnare abbastanza per pagarvi una pizza stasera, 20 euro sarebbero più che sufficienti. Di conseguenza, la differenza fra la vincita di 20 e quella di 100 euro sarebbe molto ridotta (o nulla) in termini di utilità soggettiva. La scommessa *B* risulterebbe così la più attraente semplicemente per il fatto che presenta una più elevata probabilità di vincita. Ancora una volta, quindi, la teoria non definisce la razionalità dei decisori vincolando i loro valori di utilità, ma piuttosto assumendo che essi scelgano in maniera conseguente, vale a dire massimizzando la loro utilità attesa, quale che sia.

Occorre a questo punto rendere esplicito che nella presente trattazione intendiamo fare riferimento alla teoria della scelta razionale in senso *normativo*, cioè come uno strumento che indica come gli individui *dovrebbero* risolvere i problemi decisionali che si trovano ad affrontare, date le utilità soggettive pertinenti che riflettono i loro desideri e obiettivi. In questa prospettiva, è naturale chiedersi *perché* gli individui dovrebbero conformarsi ai principi della teoria nel fronteggiare le decisioni del mondo reale, proponendosi quindi di massimizzare la loro utilità attesa. Una possibile risposta rimanda ai fondamenti assiomatici della teoria. Si prenda, a titolo di esempio, il cosiddetto “assioma di transitività”, secondo il quale se un agente preferisce *A* a *B* e *B* a *C*, allora deve preferire *A* a *C*. Non solo l'assioma appare perfettamente naturale, ma si può mostrare che dalla sua violazione conseguono esiti altamente indesiderabili.⁵ Ebbene, un fondamentale risultato matematico della teoria mostra che se un agente *non* massimizza la sua utilità attesa, allora è destinato a violare, nelle sue scelte, questo o altri principi assiomatici ben precisi e razionalmente vincolanti.

A questo punto è facile vedere come il calcolo dell'utilità attesa fornisca un

⁵ Un agente con preferenze non transitive preferirebbe quindi *A* a *B*, *B* a *C*, e al contempo *C* ad *A*. Ecco come si può trasformare un agente siffatto in un “bancomat” (*money-pump*). Gli si mette a disposizione l'opzione *C*. Poiché l'agente preferisce *B* a *C*, sarà disposto a ottenere *B* in cambio di *C* più una qualche somma di denaro. A questo punto, poiché preferisce *A* a *B*, sarà disposto a ottenere *A* in cambio di *B* più una qualche somma di denaro. Infine, poiché preferisce *C* ad *A*, sarà ora disposto a ottenere *C* in cambio di *A* più una qualche somma di denaro. L'agente è allora incorso in un esborso monetario pur ritrovandosi esattamente al punto di partenza, vale a dire in possesso di *C*. E la serie delle transazioni può essere ripetuta a piacimento. Riguardo alla fondazione assiomatica della teoria della scelta razionale, si veda la dettagliata esposizione fornita da Hargreaves Heap *et al.* (1992, cap. 1).

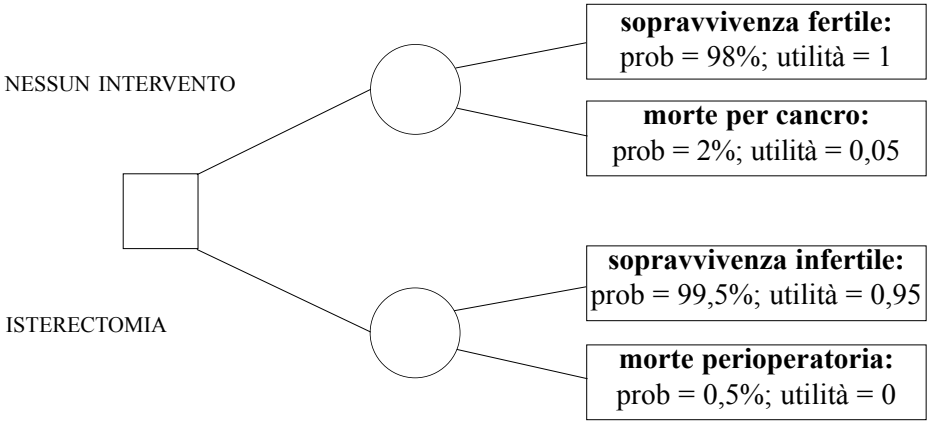
quadro di analisi in grado di chiarire la struttura delle decisioni cliniche e di orientare l'azione. Medici e pazienti, infatti, operano continuamente scelte in condizioni di incertezza e di rischio, in cui la stima della probabilità di una condizione patologica o degli effetti di un trattamento svolge una funzione determinante. Se una donna incinta accusa dolore e gonfiore alle gambe, il suo medico deve valutare la possibilità che sia affetta da tromboflebite (una complicazione relativamente comune in gravidanza), la probabilità che si verifichi un'embolia nel caso che non vengano somministrati anticoagulanti e i possibili rischi per il nascituro nel caso che essi vengano somministrati. In un paziente anziano con emicrania e febbre ricorrenti, è necessario considerare la diagnosi di arterite temporale e la possibilità di una terapia steroidea, con i suoi benefici (una riduzione del rischio di cecità) ed eventuali effetti collaterali (come la necrosi ossea). Di fronte a un paziente giovane con un nevo sul torace la cui immagine dermoscopica appaia lievemente sospetta, è necessario stabilire se il rischio dello sviluppo di una patologia grave (come un melanoma) giustifichi l'escissione chirurgica.

Per un'esemplificazione più dettagliata, si consideri il caso di una donna di ventinove anni che ha sviluppato un cancro invasivo alla cervice dell'utero. Il tumore primario è stato asportato con successo, ma si tratta di valutare l'opportunità di un ulteriore intervento chirurgico di isterectomia radicale per evitare che il cancro ritorni. Poiché il nostro esempio ha una funzione illustrativa, lo semplificheremo assumendo che l'isterectomia radicale riduca a *zero* la probabilità della successiva ricorrenza del tumore. In queste condizioni, gli esiti possibili si possono ridurre a quattro: 1. sopravvivenza fertile, 2. morte per successiva ricorrenza del tumore, 3. sopravvivenza infertile, e 4. morte perioperatoria. Lo schema ad albero della figura alla pagina successiva illustra graficamente il problema. Secondo un uso comune, il riquadro rappresenta il nodo decisionale e i cerchi rappresentano le opzioni, alle quali sono associati i possibili esiti.

Le probabilità provengono dalla letteratura scientifica e rappresentano quindi le stime plausibili di un medico informato (si veda Johnson *et al.* 1992) mentre i valori di utilità riflettono le valutazioni di una paziente reale ottenuti da Thorton, Lilford e Johnson (1992) e codificati su una scala che va da 0 (per l'esito meno desiderabile, la morte imminente) a 1 (per l'esito maggiormente desiderabile, la sopravvivenza fertile).

In queste condizioni, si può vedere che l'utilità attesa dell'opzione di *non* intervenire è superiore a quella dell'isterectomia. Infatti:

$$(3) \quad \begin{aligned} EU(\text{nessun intervento}) &= (0,98 \times 1) + (0,02 \times 0,05) = 0,981 \\ EU(\text{isterectomia}) &= (0,995 \times 0,95) + (0,005 \times 0) = 0,945. \end{aligned}$$



In base all'analisi appena svolta, è anche possibile mettere in luce i fattori che ne determinano il risultato: benché l'isterectomia offra maggiori probabilità di sopravvivenza (99,5% contro 98%), la paziente ha mostrato di dare un peso relativamente alto alla conservazione della fertilità (1 contro 0,95).

2.3. Le determinanti dell'utilità attesa e la loro quantificazione

La struttura formale della teoria della scelta razionale impiega stime di probabilità e valori di utilità come elementi di calcolo, ma non indica in che modo essi debbano essere ricavati. Nel caso del comportamento di un consumatore, per esempio, è naturale supporre che il decisore, se razionale, persegua consapevolmente il proprio interesse individuale e che egli sia quindi chiamato a valutare le probabilità e le utilità rilevanti sulla base del suo giudizio personale (facendo uso, naturalmente, di tutte le informazioni pertinenti di cui dispone). Nella medicina clinica la situazione è più complessa. La stima affidabile delle probabilità degli esiti di interesse, infatti, non può che basarsi sulle conoscenze del medico e sulla sua capacità di giudizio. D'altra parte, però, le utilità rilevanti devono riflettere adeguatamente il perseguimento del maggior beneficio *del paziente* in funzione dei *suoi* desideri e obiettivi, i quali evidentemente non sono immediatamente accessibili al medico.

A partire da questa fondamentale osservazione, la decisione clinica può essere interpretata e praticata in maniere differenti (per un'utile classificazione, si veda Emanuel e Emanuel 1992). In un approccio puramente paternalistico,

per esempio, il medico è di fatto l'unico protagonista del processo decisionale, dovendosi incaricare direttamente anche dell'identificazione degli obiettivi che meglio servono l'interesse del malato. In contrasto con l'approccio paternalistico si pone il modello di un processo decisionale condiviso e partecipativo, opportunamente basato sull'"alleanza terapeutica" fra curante e curato. In questa prospettiva, che ha guadagnato in tempi recenti favore crescente, resta inteso che è compito del medico identificare e fornire informazioni sugli esiti attesi delle diverse alternative di intervento. Tuttavia, il paziente è visto come unico e autonomo depositario delle preferenze pertinenti. La scelta più appropriata dovrebbe quindi emergere da un'adeguata combinazione di questi due elementi, che si traducono appunto, rispettivamente, in stime di probabilità e valori di utilità. In questo senso, secondo l'efficace sintesi di McNutt (2004, p. 2516), "il medico dovrebbe essere un navigatore, ma non un pilota".

Una discussione estesa delle questioni appena menzionate ricade al di fuori dei limiti del presente contributo. Vorremmo tuttavia osservare, con l'aiuto di alcuni esempi, come anche in questo contesto l'analisi delle decisioni fornisca un utile quadro concettuale per rappresentare e comprendere alcuni degli aspetti più critici.

Si consideri uno scenario clinico per molti versi emblematico, vale a dire la diagnosi di un tumore alla prostata clinicamente localizzato. Si tratta di una condizione per la quale la stessa comunità medica non ha individuato un trattamento di elezione fra la rimozione chirurgica e le terapie radiologiche (Chodak 1998). Opzioni di intervento alternative risultano infatti paragonabili in termini di efficacia (probabilità di sopravvivenza), ciascuna presentando specifici vantaggi e svantaggi da altri punti di vista (come la conservazione della funzionalità sessuale, o il rischio di complicazioni dovute al trattamento). Situazioni di questo genere mettono in evidenza un fattore cruciale per l'appropriatezza delle scelte di trattamento, vale a dire l'accuratezza con la quale si riesce a determinare il peso che il singolo paziente attribuisce ai diversi aspetti rilevanti per la decisione in esame.

Per i nostri scopi, conta sottolineare che gli elementi che determinano l'utilità attesa di una scelta clinica sono spesso tutt'altro che manifesti, ma piuttosto tali da richiedere un apposito sforzo di ricerca ed elaborazione. Nel caso delle probabilità, ciò comporta la formulazione di giudizi ponderati basati sulla migliore evidenza scientifica disponibile e sulle specifiche caratteristiche della situazione clinica. A ciò si aggiunge appunto l'esigenza di identificare i valori di utilità più appropriati. L'opportuna interrogazione del destinatario dei trattamenti rappresenta in questo senso una procedura del tutto naturale e ra-

gionevole, quando è perseguibile.⁶ Quando non lo è, è comunque possibile affrontare il problema ricorrendo a valutazioni di altra natura, come si evince dall'esempio su cui ora ci soffermeremo.

Immaginate che vi sia una certa patologia (per esempio, una forma di influenza) che comporta un rischio di morte per i bambini piccoli, corrispondente, diciamo, a 10 casi su 10000 (0,1%). Esiste un vaccino che rimuove tale rischio, ma che può determinare esso stesso degli effetti collaterali mortali in 9 casi su 10000 (0,09%). Se assegniamo alla morte utilità 0 e alla sopravvivenza utilità 1, è immediato verificare che:

$$(4) \quad EU(\text{vaccino}) = 0,9991 \times 1 > 0,999 \times 1 = EU(\text{nessun intervento}).$$

In altri termini, eseguire la vaccinazione rappresenta la scelta ottimale. Eppure gli psicologi cognitivi Ritov e Baron (1990) hanno osservato sperimentalmente che, in queste condizioni, la maggioranza delle persone sceglierebbe altrimenti, opponendosi alla vaccinazione di proprio figlio.⁷ Che cosa si nasconde dietro questo comportamento, indicato dagli studiosi come “fallacia dell'omissione”?

Una plausibile ipotesi di spiegazione è la seguente. I genitori percepiscono una differenza rilevante fra gli esiti avversi delle due opzioni. In un caso, alla perdita del figlio si aggiunge il tragico rimpianto di avere *attivamente* contribuito all'evento, esponendolo alla vaccinazione. Nello scenario alternativo, tale aggravante è assente, o comunque meno pressante. Se rappresentiamo questa situazione assegnando valore 0 solo allo scenario che è percepito come maggiormente negativo (la “morte per intervento”) e alla “morte per omissione” un valore sufficientemente differente (per esempio, 0,2), otteniamo un'inversione dei valori delle utilità attese che riflette le preferenze rilevate da Ritov e Baron:

$$(5) \quad \begin{aligned} EU(\text{vaccino}) &= (0,9991 \times 1) + (0,0009 \times 0) = 0,9991 \\ EU(\text{nessun intervento}) &= (0,999 \times 1) + (0,001 \times 0,2) = 0,9992. \end{aligned}$$

Dal nostro punto di vista, il caso della vaccinazione è interessante perché ci impone ancora una volta di riflettere su *quali* siano le utilità rilevanti per la de-

⁶ Il repertorio dei metodi suggeriti per guidare tale interrogazione è ormai piuttosto esteso. Per una rassegna e discussione, si possono vedere Karel (2000) e Stiggelbout (2000).

⁷ Si veda Asch *et al.* (1994) per uno studio analogo relativo a un vaccino reale contro la pertosse.

cisione. La seconda analisi del problema decisionale (le equazioni al punto (5)), sebbene formalmente accettabile, risente in modo cruciale dell'introduzione di un fattore (la diversa percezione di responsabilità dei genitori in relazione all'azione in confronto all'omissione) che risulta però del tutto irrilevante dal punto di vista delle utilità relative al benessere del bambino, nel cui esclusivo interesse, presumibilmente, la decisione dovrebbe essere presa. Per di più, la percezione di una significativa differenza fra azione e omissione potrebbe rivelarsi illusoria alla prova dei fatti.⁸ Benjamin Franklin, toccato dal problema in prima persona agli albori delle tecniche di vaccinazione, ebbe a scrivere in proposito (si veda Schmidt 1976):

Nel 1736 persi uno dei miei figli, un bel bambino di quattro anni, per vaiolo contratto naturalmente. Ho amaramente rimpianto di non averglielo inoculato. Ricordo ciò per il bene dei genitori che non prendono in considerazione questo intervento, presumendo che non si perderebbero mai l'eventuale morte di un bambino per questo motivo. Il mio esempio mostra che il rimpianto può essere lo stesso in entrambi i casi e che perciò occorre scegliere il mezzo più sicuro.

È da segnalare come la rilevanza cruciale di un diverso atteggiamento riguardo all'azione e all'inazione, così come della valutazione delle utilità di un individuo che non è in condizione di esprimerle, può estendersi a questioni particolarmente delicate ai confini fra medicina ed etica, come la morte medicalmente assistita.⁹

3. Decisioni diagnostiche

3.1. *L'utilità delle ricerche di informazioni cliniche*

Come abbiamo anticipato, accanto all'identificazione degli interventi terapeutici più appropriati – su cui ci siamo soffermati nei paragrafi precedenti – esiste una seconda classe di problemi decisionali che riveste grande importanza per la pratica clinica. Si tratta delle scelte che riguardano l'opportunità di rac-

⁸ Sull'affascinante problema della previsione delle proprie utilità future – che, per quanto rilevante, ci è qui impossibile trattare – rimandiamo all'importante discussione di Kahneman (2000) e a quella, più direttamente legata all'ambito clinico, di Ubel *et al.* (2005).

⁹ Riguardo ai rapporti fra teoria della decisione e bioetica, si veda l'ampia e stimolante discussione offerta da Baron (2006).

cogliere nuove informazioni e, in caso affermativo, il modo più efficace per farlo.

Che la raccolta di nuove informazioni possa giovare alla soluzione di un problema decisionale appare intuitivamente ovvio, e può essere motivato analiticamente come segue. Negli esempi fin qui discussi, si è visto che le decisioni in condizioni di incertezza o di rischio condividono un fondamentale tratto distintivo: gli esiti di interesse (per esempio, la guarigione) non dipendono soltanto dalla scelta dell'una o dell'altra opzione (per esempio, la somministrazione di un certo farmaco), ma anche dal verificarsi di stati di cose sui quali il decisore non ha controllo (per esempio, gli effettivi sviluppi patologici in corso, dei quali non si ha piena conoscenza). Come si è visto, per identificare il corso d'azione ottimale il decisore deve affidarsi alla valutazione della probabilità di questi ultimi stati di cose. Con l'acquisizione di nuove informazioni rilevanti, tali probabilità potranno essere opportunamente aggiornate, modificando così l'utilità attesa delle diverse opzioni.¹⁰

Tuttavia, assegnare uno specifico valore di utilità allo svolgimento di una ricerca di informazioni può risultare tutt'altro che banale alla luce di qualche ulteriore considerazione. Un test diagnostico, evidentemente, può avere diversi esiti (nel caso più semplice, positivo oppure negativo). È chiaro d'altra parte che, per orientare la decisione di eseguire o meno il test in questione, il calcolo dell'utilità attesa deve aver luogo *prima* che il suo esito sia stato ottenuto. Come è possibile eseguire un'analisi di questo tipo? Per illustrarlo, ci serviremo ancora una volta di un esempio.

Si consideri un paziente affetto da insufficienza epatica cronica progressiva, per il quale il medico ha isolato due possibili ipotesi diagnostiche alternative che richiedono trattamenti differenti: epatite cronica attiva o cirrosi. In base alle informazioni di cui dispone, il medico è incerto sulle due possibilità, al punto che assegna ad entrambe la stessa probabilità (cioè il 50%). Seguendo Weinstein e Fineberg (1980), da cui l'esempio è tratto, assumiamo che l'esito clinico di interesse sia la sopravvivenza a due anni (con utilità uguale a 1) a fronte della mancata sopravvivenza a due anni (con utilità uguale a 0). Se il paziente è affetto da epatite, la sua probabilità di sopravvivenza a due anni è del 67%, ma potrebbe essere portata all'85% con un trattamento a base di steroidi. Se viceversa il paziente soffre di cirrosi, la sua probabilità di so-

¹⁰ Tale processo di aggiornamento probabilistico trova la sua più diffusa formalizzazione nel cosiddetto approccio bayesiano al ragionamento clinico. Per una trattazione recente ed estesa, che comprende ulteriori riferimenti bibliografici, ci permettiamo di rimandare a Festa, Butasi e Crupi (2009).

pravvivenza a due anni è del 50% e scenderebbe al 48% in presenza dello stesso trattamento, a causa del rischio di possibili effetti collaterali (come emorragie gastrointestinali, o complicazioni tromboemboliche). Date queste condizioni, il trattamento a base di steroidi garantirebbe una maggiore utilità attesa. Infatti:

$$\begin{aligned}
 (6) \quad EU(\text{steroidi}) &= [P(\text{epatite}) \times 0,85] + [P(\text{cirrosi}) \times 0,48] \\
 &= [0,5 \times 0,85] + [0,5 \times 0,48] = 0,665 \\
 EU(\text{nessun intervento}) &= [P(\text{epatite}) \times 0,67] + [P(\text{cirrosi}) \times 0,50] \\
 &= [0,5 \times 0,67] + [0,5 \times 0,50] = 0,585.
 \end{aligned}$$

Esiste però un'ulteriore possibilità, vale a dire rimandare la scelta di trattamento a dopo l'esecuzione di una biopsia epatica, una procedura diagnostica moderatamente impegnativa che richiede il prelievo chirurgico di un campione di tessuto dal fegato. Il problema che ci interessa è quindi determinare in che misura l'esecuzione della biopsia possa comportare un prevedibile beneficio rispetto alle prospettive cliniche del paziente.

A titolo illustrativo, supporremo per il momento che il rischio di morte associato alla procedura biopica sia trascurabile e che l'esito di tale procedura determini *con certezza* la diagnosi corretta. Indicando con B^+ e B^- i due possibili risultati della biopsia (rispettivamente, positivo e negativo per l'epatite), quest'ultimo presupposto si esprime con le seguenti probabilità *condizionate* (il simbolo “|” si legge “dato che” o “supponendo che”):

$$\begin{aligned}
 (7) \quad P(\text{epatite}|B^+) &= 1, \text{ e quindi } P(\text{cirrosi}|B^+) = 0 && \text{[massimo valore} \\
 &&& \text{predittivo positivo]} \\
 P(\text{cirrosi}|B^-) &= 1, \text{ e quindi } P(\text{epatite}|B^-) = 0 && \text{[massimo valore} \\
 &&& \text{predittivo negativo]}.
 \end{aligned}$$

Si noti anche che, ipotizzando appunto che l'esame rilevi con certezza la presenza di ciascuna patologia, la probabilità semplice (non condizionata) di ciascun risultato del test è identica a quella inizialmente stimata per ciascuna patologia. Vale a dire:

$$\begin{aligned}
 (8) \quad P(B^+) &= P(\text{epatite}) = 0,5 \\
 P(B^-) &= P(\text{cirrosi}) = 0,5.
 \end{aligned}$$

Occorre ora considerare *entrambi* i possibili esiti dell'esame. Se la biopsia fosse positiva, si opterebbe per il trattamento steroideo, con utilità attesa pari a:

$$(9) \quad EU_{B^+}(\text{steroidi}) = [P(\text{epatite}|B^+) \times 0,85] + [P(\text{cirrosi}|B^+) \times 0,48] \\ = (1 \times 0,85) + (0 \times 0,48) = 0,85.$$

Se d'altra parte la biopsia fosse negativa, ci si asterebbe dal trattamento steroideo, con utilità attesa pari a:

$$(10) \quad EU_{B^-}(\text{nessun intervento}) = [P(\text{epatite}|B^-) \times 0,67] + [P(\text{cirrosi}|B^-) \times 0,50] \\ = (0 \times 0,67) + (1 \times 0,50) = 0,50.$$

Ma poiché la biopsia non è ancora stata eseguita, è necessario integrare questi ultimi due valori di utilità attesa, moltiplicandoli per la probabilità di occorrenza dei corrispondenti esiti dell'esame. Dal momento che i due esiti B^+ e B^- sono equiprobabili (si vedano le equazioni al punto (8)), il calcolo in questione corrisponde a una semplice media:

$$(11) \quad [P(B^+) \times 0,85] + [P(B^-) \times 0,50] = (0,5 \times 0,85) + (0,5 \times 0,50) = 0,675.$$

Soffermiamoci sul significato della quantità appena calcolata, vale a dire 0,675. Si tratta dell'utilità attesa della scelta che verrà eseguita a seguito dell'ottenimento del risultato della biopsia, calcolata quando tale risultato *non* è ancora stato osservato. Si vede quindi che, nelle condizioni date, l'esecuzione della biopsia determina un incremento dell'utilità attesa da 0,665 a 0,675 in confronto all'opzione di scegliere immediatamente un trattamento (procedendo così alla somministrazione di steroidi senza raccogliere alcuna informazione aggiuntiva). Ciò corrisponde ad un *guadagno* di utilità attesa equivalente a $0,675 - 0,665 = 0,01$, che rappresenta lo specifico valore associato alla ricerca di informazioni attraverso l'esame bioptico. Può apparire un beneficio modesto, ma è utile considerarne le implicazioni: su cento pazienti come quello descritto che vengono sottoposti a biopsia epatica ce ne sarà *uno in più* in vita a due anni di distanza in confronto a quanto avverrebbe omettendo l'esecuzione dell'esame.

Per i nostri scopi illustrativi, abbiamo inizialmente supposto che la biopsia epatica rimuova completamente l'incertezza diagnostica nel caso appena discusso – vale a dire che i suoi valori predittivi positivo e negativo siano massimi, equivalendo entrambi a 1. Occorre tuttavia notare che lo schema di analisi qui presentato può essere esteso al caso maggiormente realistico in cui si tratti di valutare l'esecuzione di un test diagnostico informativo ma imperfetto. Si tratterà di ripercorrere la stessa procedura di calcolo servendosi delle più opportune stime del valore predittivo positivo e negativo del test in questione.

Abbiamo altresì supposto che il rischio di morte associato alla procedura biptica fosse trascurabile, ma anche da questo punto di vista è possibile modificare il presupposto e raffinare opportunamente l'analisi. Senza scendere nei dettagli (per i quali rimandiamo alla più estesa trattazione di Weinstein e Fineberg 1980, pp. 93 ss.), facciamo l'ipotesi che la biopsia epatica comporti un significativo rischio specifico di mortalità, per esempio del 5% (magari a causa di qualche complicazione aggiuntiva del quadro clinico complessivo del paziente). In tal caso, è facile intuire (ed è possibile dimostrare formalmente) che i potenziali costi dell'indagine biptica superano i suoi benefici attesi, rendendone svantaggiosa l'esecuzione.

Alla luce di queste ultime precisazioni, si può vedere come l'analisi del valore di una ricerca di informazioni diagnostiche riguardi una classe di problemi della massima rilevanza per la medicina clinica. Esiste infatti una naturale tendenza degli esami maggiormente informativi a implicare maggiori rischi e un più elevato dispendio di risorse. Per questo motivo, è spesso necessario chiedersi se valga la pena di eseguire la ricerca; se, cioè, i vantaggi attesi dall'acquisizione di informazioni aggiuntive siano ancora tali, al netto dei costi che essa può comportare.

3.2. *Sequenze di esami e trabocchetti decisionali*

Come abbiamo appena visto, la teoria della scelta razionale permette di analizzare i problemi decisionali relativi alla ricerca di informazioni cliniche aggiuntive, e dei modi più opportuni per svolgerla. Fra i molti possibili sviluppi di questo tipo di analisi, va segnalato il complesso problema della pianificazione di una sequenza di indagini diagnostiche. In uno dei primi studi dedicati a questo tema, Dennis Fryback si è concentrato sul problema specifico di una massa renale sospetta individuata radiograficamente, rispetto alla quale si debba stabilire una fra tre ipotesi diagnostiche alternative, vale a dire se si tratti di un tumore, di una cisti o di una "variante normale" non patologica (si veda Fryback e Thornbury 1976; nel seguito faremo principalmente riferimento al resoconto di Hastie e Dawes 2001, pp. 281-283, con alcuni adattamenti).

All'epoca dello studio, non era possibile distinguere le tre possibili condizioni cliniche attraverso una singola procedura (per esempio, un metodo di diagnostica per immagini più potente della radiografia). Per rilevare la presenza di una cisti, si tentava il prelievo di liquido dalla massa attraverso un intervento poco invasivo, detto "aspirazione", che comportava rischi minimi.

Per determinare la presenza di un tumore, invece, si doveva ricorrere a un trattamento notevolmente più invasivo, disagiata e rischioso: una arteriografia con l'inserimento di una sonda in grado di prelevare un campione di tessuto da sottoporre a biopsia. La questione al centro della ricerca di Fryback era: quale dei due esami deve essere eseguito *per primo*? Per apprezzare la rilevanza del problema è utile considerare quanto segue. Se il primo esame eseguito fosse risultato negativo, il secondo sarebbe stato comunque necessario, per scartare (o confermare) la possibilità di una variante normale. Se il primo esame eseguito fosse risultato positivo, tuttavia, si sarebbe potuto risparmiare al paziente l'esecuzione del secondo, con i suoi costi clinici e personali.

La procedura decisionale comunemente adottata rifletteva una regola assai semplice: esegui per primo l'esame relativo alla condizione patologica che appare la più probabile fra le due. In altri termini, se l'immagine radiografica era maggiormente indicativa di una cisti piuttosto che di un tumore, si prescriveva per prima l'aspirazione; altrimenti, l'arteriografia. La regola può apparire a prima vista ragionevole, eppure Fryback mostrò che in molte circostanze produce scelte subottimali. Per rendercene conto, assegniamo valore di utilità 1 all'esito meno sfavorevole (l'esecuzione della sola aspirazione) e valore 0 a quello più negativo (l'esecuzione di entrambi gli esami). Resta da stabilire quale sia l'utilità dell'esecuzione della sola arteriografia nell'intervallo fra 0 e 1. In base al giudizio di medici e pazienti, Fryback ricavò una stima equivalente a 0,1. Ciò significa che l'aspirazione era generalmente percepita come dieci volte più tollerabile (o meno avversa) dell'arteriografia. A questo punto, l'utilità attesa della scelta di *eseguire per prima l'aspirazione* è data da:¹¹

$$\begin{aligned} (12) \quad & [P(\text{cisti}) \times U(\text{asp})] + [P(\text{tumore}) \times U(\text{asp} + \text{art})] + [P(\text{normale}) \times U(\text{asp} + \text{art})] \\ & = [P(\text{cisti}) \times 1] + [P(\text{tumore}) \times 0] + [P(\text{normale}) \times 0] \\ & = P(\text{cisti}). \end{aligned}$$

Per contro, l'utilità attesa di *eseguire per prima l'arteriografia* equivale a:

$$\begin{aligned} (13) \quad & [P(\text{tumore}) \times U(\text{art})] + [P(\text{cisti}) \times U(\text{asp} + \text{art})] + [P(\text{normale}) \times U(\text{asp} + \text{art})] \\ & = [P(\text{tumore}) \times 0,1] + [P(\text{cisti}) \times 0] + [P(\text{normale}) \times 0] \\ & = P(\text{tumore}) \times 0,1. \end{aligned}$$

¹¹ Il lettore più attento noterà forse che le espressioni presentate di seguito impiegano direttamente le probabilità delle diagnosi in sostituzione di quelle dei corrispondenti risultati dei test. Con ciò si suppone implicitamente che tali risultati si possano considerare perfettamente affidabili.

Ora chiediamoci in quali condizioni il secondo valore sia maggiore del primo. Poniamo pertanto:

$$(14) \quad P(\text{tumore}) \times 0,1 > P(\text{cisti})$$

Ottenendo:

$$(15) \quad P(\text{tumore})/P(\text{cisti}) > 10$$

Per come è stato definito il problema, quindi, non è sufficiente – come supponeva la prassi clinica usuale – che un tumore sia più probabile di una cisti perché iniziare dall'arteriografia rappresenti la scelta ottimale; è necessario, piuttosto, che sia *almeno dieci volte* più probabile!

Per concludere la nostra discussione delle decisioni diagnostiche, vorremmo rilevare come la teoria della scelta razionale chiarisca il problema della pianificazione diagnostica anche al di là di specifici risultati quantitativi come quello appena presentato. Essa mette in luce, infatti, una condizione generale che deve essere soddisfatta affinché un'indagine sia meritevole di considerazione. Ci riferiamo a ciò che Leonard Savage ha denominato “principio della cosa sicura” (*sure thing principle*) e illustrato attraverso l'esempio di un imprenditore che debba decidere a proposito di un investimento in tempo di elezioni. Se l'investimento gli appare conveniente quale che sia lo schieramento politico che andrà al governo, osserva Savage, allora non c'è motivo per cui il decisore ricerchi informazioni per prevedere l'esito della consultazione elettorale, o ne attenda lo svolgimento (si veda Savage 1954, p. 21). Non tutte le informazioni possibili, quindi, meritano di essere ricercate. Perché una tale ricerca possa risultare utile ai fini di una decisione, essa deve riguardare un'informazione che, una volta ottenuta, sia in grado di influenzare in un senso o nell'altro le scelte future. Nel contesto clinico, ciò si applica immediatamente ai test diagnostici, che per definizione possono avere esiti diversi. Ebbene, devono esserci almeno due diversi esiti possibili (nel caso più semplice, il test è positivo oppure negativo) a seguito dei quali risulta opportuno perseguire corsi d'azioni fra loro differenti (per esempio, eseguire o non eseguire un intervento chirurgico). È in effetti possibile dimostrare formalmente che se una certa ricerca di informazioni non soddisfa questa condizione minima, allora la sua esecuzione non può comportare un contributo positivo in termini di utilità attesa.

L'apparente ovvietà del “principio della cosa sicura” non ne elimina la rilevanza come guida del comportamento, soprattutto perché, come ora vedremo, in certe condizioni i decisori in carne e ossa si rivelano inclini a violarlo,

con spiacevoli conseguenze. Immaginate, per esempio, di essere uno studente e di aver appena sostenuto un esame molto impegnativo. Se foste stati promossi, comprendereste un biglietto in offerta per una vacanza in un posto rinomato? E se invece foste stati respinti? Tversky e Shafir (1992) hanno osservato che i due terzi di un gruppo di studenti universitari avrebbero scelto nello stesso modo in entrambi i casi, cioè che l'esito dell'esame non influenzava la loro decisione in un senso o nell'altro. Eppure una consistente maggioranza (61%) di un secondo gruppo di studenti, ai quali *non* veniva detto se erano stati promossi o respinti, avrebbe preferito subire una piccola penale pur di rinviare la scelta dell'acquisto del biglietto a dopo la comunicazione dei risultati. Pagando così di tasca propria pur di ottenere un'informazione che fornisse loro una "buona ragione" per decidere. Un'identica scelta come quella di partire, infatti, avrebbe potuto essere motivata dal bisogno di riposo in vista della necessaria ripetizione dell'esame, nel caso fossero stati respinti; oppure dal legittimo desiderio di "premiarsi", nel caso fossero stati promossi.

L'esigenza psicologica di trovare "buone ragioni" per decidere può quindi promuovere una ricerca di informazioni cui è associato un dispendio di risorse a conti fatti inutile (Shafir, Simonson e Tversky 1993). E non solo nel pianificare le proprie vacanze, ma anche in contesti clinici le cui conseguenze hanno ben altra portata. David Eddy ha per esempio analizzato nei dettagli l'impiego della mammografia in casi di sospetto cancro al seno, osservando nel comportamento di medici esperti la tendenza a cadere in un trabocchetto del tutto analogo (Eddy 1982). Per determinare la natura (benigna o maligna) di un nodulo, molti di loro ordinano infatti una mammografia anche se poi, *qualunque sia il risultato*, procederanno comunque prescrivendo una biopsia: per escludere definitivamente la diagnosi di cancro (nel caso la mammografia sia risultata negativa); oppure per confermare definitivamente quella stessa diagnosi (nel caso la mammografia sia risultata positiva). In questo modo, evidentemente, si stabilisce di eseguire un esame clinico i cui risultati di fatto non influenzeranno il successivo trattamento. In una situazione del genere, quindi, il test mammografico può essere senz'altro omesso, a favore di un'esecuzione più tempestiva dell'intervento biopsico (che in seguito, lo ripetiamo, verrebbe svolto *in ogni caso*).¹² A questo proposito, occorre osservare che i costi asso-

¹² Si potrebbe obiettare che l'immagine mammografia risulterebbe utile per guidare l'eventuale rimozione chirurgica del nodulo. Questo osservazione, tuttavia, non indebolisce l'analisi di Eddy. Se infatti è *questo* il contributo che si attribuisce alla mammografia, allora l'indagine diagnostica dovrebbe comunque iniziare dalla biopsia, prevedendo l'esecuzione della mammografia *solo* nel caso di una biopsia positiva, che indica la necessità di una nodulectomia.

ciati a una ricerca di informazioni comprendono anche il *tempo* necessario per il suo completamento, durante il quale c'è sempre un rischio (più o meno alto, ma ineliminabile) che le condizioni del paziente, se malato, peggiorino o si complicino. In questo senso, gli esiti indesiderati di una ricerca di informazioni priva di valore non riguardano soltanto lo sperpero di risorse economiche, ma possono incidere sulla stessa efficacia delle cure.

4. Osservazioni conclusive

Nei paragrafi precedenti abbiamo presentato la teoria della scelta razionale e ne abbiamo discusso alcune applicazioni di potenziale interesse per chi sia chiamato ad esercitare la medicina clinica o a riflettere su di essa. Chiaramente, il carattere di questa trattazione è introduttivo. Non abbiamo alcuna difficoltà a riconoscere che gli scenari clinici qui discussi rappresentano per molti versi delle idealizzazioni. Essi andrebbero quindi considerati come esempi puramente illustrativi delle due grandi classi di problemi decisionali che abbiamo identificato, vale a dire le scelte terapeutiche e quelle diagnostiche. Inoltre, il ricorso ad approfondimenti bibliografici, ai quali abbiamo regolarmente rimandato il lettore interessato, riflette la quantità di temi rilevanti che sono stati soltanto sfiorati – non meno importanti di quelli che è stato possibile esporre più estesamente. Vorremmo ora concludere la nostra discussione con alcune brevi osservazioni generali sulla teoria delle decisioni e i suoi rapporti con la medicina clinica.

Secondo il premio Nobel per l'economia Herbert Simon, la teoria della scelta razionale è un vero e proprio “gioiello intellettuale” dei nostri tempi (Simon 1983, p. 3). A fronte di questo riconoscimento, peraltro, Simon non ha esitato a discutere in modo molto critico il significato della teoria. A partire dai suoi primi fondamentali contributi (si veda Simon 1957), egli ha sottolineato con forza che i decisori in carne e ossa *non* compiono le loro scelte come efficienti calcolatori di probabilità e utilità attese, a causa della limitatezza delle risorse computazionali disponibili alla mente umana. In effetti, la successiva ricerca sulla *psicologia* della decisione ha ampiamente documentato il ricorso prevalente all'intuizione e a procedure di ragionamento semplificate che producono violazioni sistematiche dei principi della scelta razionale (Kahneman 2002). Si tratta di risultati di notevole interesse per qualsiasi disciplina interessata al tema della razionalità, a cominciare dalla filosofia (Stich 1990; Marconi 2001). Occorre tuttavia sottolineare che nella decisione – così come, poniamo, nell'aritmetica o nella grammatica – l'identificazione degli errori e la

loro correzione presuppongono una teoria che indichi le soluzioni corrette in base a principi il più possibile rigorosi. In breve, i limiti *descrittivi* della teoria non ne intaccano il valore *normativo*; semmai, lo enfatizzano. In questo senso, la teoria della scelta razionale resta un punto di riferimento importante per la medicina, specie nel momento in cui si va affermando un interesse costruttivo per il tema dell'errore clinico (si vedano Delvecchio 2005; Crupi, Gensini e Motterlini 2006; Crupi 2008).

I medici, per parte loro, risultano particolarmente sensibili a due principali rilievi critici riguardo alle applicazioni cliniche della teoria delle decisioni. In primo luogo, è ricorrente l'osservazione che l'analisi decisionale è incompatibile con i ritmi serrati cui le attività di cura sono spesso sottoposte. Proprio perché nella medicina clinica il tempo rappresenta una risorsa importante e talvolta cruciale, di cui disporre opportunamente nella scelta e nell'azione, la preoccupazione appare ragionevole. A un esame più attento, tuttavia, essa è destinata a ridimensionarsi almeno in parte. Innanzi tutto, la teoria delle decisioni identifica alcuni principi generali e puramente qualitativi (è il caso del "principio della cosa sicura" nella ricerca di informazioni) la cui applicazione può migliorare le scelte pur richiedendo uno sforzo di riflessione minimo e molto circoscritto. In secondo luogo, la semplice identificazione degli elementi costitutivi di un problema di scelta – cioè la corretta definizione di un albero decisionale – può in molti casi determinare un approccio più chiaro e coerente per il medico e il paziente anche senza che si proceda a una elaborazione quantitativa. Il ricorso a quest'ultima risulta poi potenzialmente utile (e spesso effettivamente vantaggioso) solo in un limitato sottoinsieme di casi particolarmente problematici.

Secondo un'altra obiezione alquanto diffusa, l'analisi delle decisioni svilirebbe e deformerebbe la pratica della medicina clinica trascurando che la cura è un'"arte" più che una scienza, basata quindi sull'esercizio del "giudizio clinico" piuttosto che sull'utilizzo di diagrammi e formule. A questo proposito, è utile ricordare un'efficace analogia di Weinstein e Fineberg (1980, p. 390): il rapporto fra un albero decisionale e la realtà clinica richiama quello fra una mappa e il relativo territorio. In entrambi i casi, una rappresentazione adeguata richiede una notevole capacità di giudizio e l'imprescindibile contributo di competenze esperte. Nel caso medico, ciò si riflette nella selezione e nella valutazione dei corsi d'azione da includere fra le opzioni e delle conseguenze a breve e lungo termine che meritano considerazione. In questo modo, l'analisi decisionale non discredita il giudizio clinico; piuttosto ne sollecita l'impiego rispetto a problemi i cui singoli elementi vengono opportunamente esplicitati e organizzati.

Discutendo di questi temi ormai trent'anni fa, William Schwartz ha riflettuto sul modo in cui la conoscenza esperta si consolida ed evolve nel tempo (Schwartz 1979). Si dà per scontato che tale evoluzione implichi l'acquisizione di nuove informazioni, in particolare i più aggiornati e affidabili risultati accumulati dalla ricerca medica. Curiosamente, osserva Schwartz, non è altrettanto comune il riconoscimento dell'esigenza di arricchire e raffinare "le nostre capacità di *elaborare* le informazioni" (*ibid.*, p. 559). Servendosi di un'altra utile analogia, è come se all'impegno nel tenere in ordine il database non si accompagnasse la preoccupazione di migliorare il software. A distanza di tre decenni, si può dire che l'invito a compensare questa asimmetria non abbia perso la sua rilevanza.

RIFERIMENTI BIBLIOGRAFICI

- ASCH, D. A. *et al.* (1994): "Omission Bias and Pertussis Vaccination", *Medical Decision Making*, 14, pp. 118-123.
- BARON, J. (2000): *Thinking and Deciding*, New York: Cambridge University Press.
- (2006): *Against Bioethics*, Cambridge (Mass.): MIT Press (trad. it. *Contro la bioetica*, Milano: Raffaello Cortina 2008).
- CHAPMAN, G.B. e SONNENBERG, F.A. (a cura di) (2000): *Decision Making in Health Care*, Cambridge: Cambridge University Press.
- CHODAK, G.W. (1998): "Comparing Treatments for Localized Prostate Cancer: Persisting Uncertainty", *Journal of the American Medical Association*, 280, pp. 1008-1010.
- CRUPI, V. (2005): "Aspetti cognitivi della razionalità medica: dall'evidenza alla decisione", *Nuova civiltà delle macchine*, 23, pp. 59-68.
- (2008): "Errore, decisione e razionalità in medicina", *L'arco di Giano*, 55, pp. 29-41.
- CRUPI, V., GENSINI, G. F. e MOTTERLINI, M. (a cura di) (2006): *La dimensione cognitiva dell'errore in medicina*, Milano: Franco Angeli.
- DE FINETTI, B. (1937): "La prévision: ses lois logiques, ses sources subjectives", *Annales de l'Institut Henri Poincaré*, 7, pp. 1-68.
- DELVECCHIO, G. (2005): *Decisione ed errore in medicina*, Torino: Centro Scientifico Editore.
- DEVLIN, K. (2008): *The Unfinished Game. Pascal, Fermat and the Seventeenth-Century Letter that Made the World Modern*, London: Basic Books.
- EDDY, D. M. (1982): "Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities", in D. Kahneman, P. Slovic e A. Tversky, A. (a cura di), *Judg-*

- ment under Uncertainty: Heuristics and Biases*, New York: Cambridge University Press, pp. 249-267 (trad. it. “Il ragionamento probabilistico nella medicina clinica: problemi e opportunità”, in Crupi, Gensini e Motterlini, a cura di, 2006, pp. 45-67).
- EMANUEL, E. J. e EMANUEL, L. L. (1992): “Four Models of the Physician-Patient Relationship”, *Journal of the American Medical Association*, 267, pp. 2221-2226.
- FESTA, R. (2004): “Principio di evidenza totale, decisioni cliniche ed Evidence Based Medicine”, in G. Federspil e P. Giaretta (a cura di), *Forme della razionalità medica*, Soveria Mannelli: Rubbettino, pp. 47-82.
- FESTA, R., BUTTASI, C. e CRUPI, V. (2009): “Evidenza incerta e probabilità delle diagnosi: estensioni dell’approccio bayesiano alla pratica clinica”, in P. Giaretta, A. Moretto, G. F. Gensini e M. Trabucchi (a cura di), *Filosofia della medicina*, Bologna: Il Mulino, pp. 565-609.
- FRYBACK, D. G. e THORNBURY, J. R. (1976), “Evaluation of a Computerized Bayesian Model for Diagnosis of Renal Cysts versus Tumor versus Normal Variant from Exploratory Urogram Information”, *Investigative Radiology*, 11, pp. 102-111.
- HACKING, I., (1975): *The Emergence of Probability*, Cambridge: Cambridge University Press (trad. it. *L’emergenza della probabilità*, Milano: Il Saggiatore, 1987).
- HARGREAVES HEAP, S. *et al.* (1992): *The Theory of Choice: A Critical Guide*, Cambridge: Blackwell (trad. it. *La teoria della scelta: una guida critica*, Roma-Bari: Laterza, 1996).
- HASTIE, R. e DAWES, R.M. (2001): *Rational Choice in an Uncertain World*, Thousand Oaks (Calif.): Sage.
- HUNINK, M. G. M. *et al.* (2001): *Decision Making in Health and Medicine*, Cambridge: Cambridge University Press.
- JOHNSON, N. *et al.* (1992): “Using Decision Analysis to Calculate the Optimum Treatment for Microinvasive Cervical Cancer”, *British Journal of Cancer*, 65, pp. 717-722.
- JORDAN, J. (a cura di) (1994): *Gambling on God: Essays on Pascal’s Wager*, Lanham (Md.): Rowman & Littlefield.
- KAHNEMAN, D. (2000): “New Challenges to the Rationality Assumption”, in D. Kahneman e A. Tversky (a cura di), *Choices, Values and Frames*, New York: Cambridge University Press, pp. 758-774.
- (2002): “Maps of Bounded Rationality: Psychology for Behavioral Economics”, *American Economic Review*, 93, pp. 1449-1475 (trad. it. “Mappe di razionalità limitata”, in D. Kahneman, D. McFadden e V. L. Smith, *Critica della ragione economica*, Milano: Il Saggiatore, 2005, pp. 77-140).

- KAREL, M. J. (2000): "The Assessment of Values in Medical Decision Making", *Journal of Aging Studies*, 14, pp. 403-422.
- KEENEY, R. L. e RAIFFA, H. (1976): *Decisione with Multiple Objectives: Preference and Value Trade-offs*, New York: Cambridge University Press.
- LEDLEY, R. S. e LUSTED, L. B. (1959): "Reasoning Foundations of Medical Diagnosis", *Science*, 130, pp. 9-21.
- MARCONI, D. (2001): *Filosofia e scienza cognitiva*, Roma-Bari: Laterza.
- MCNUTT, R. A. (2004): "Shared Medical Decision Making: Problems, Process, Progress", *Journal of the American Medical Association*, 292, pp. 2516-2518.
- MOTTERLINI, M. (2007): *Economia emotiva*, Milano: Rizzoli.
- MOTTERLINI, M. e CRUPI, V. (2005): *Decisioni mediche: un punto di vista cognitivo*, Milano: Raffaello Cortina.
- MOTTERLINI, M. e PIATTELLI PALMARINI M. (2005): "Introduzione", in D. Kahneman, D. McFadden e V. L. Smith, *Critica della ragione economica*, Milano: Il Saggiatore, 2005, pp. 9-24.
- RAMSEY, F. P. (1931): "Truth and probability", in Id., *Foundations of Mathematics and Other Essays*, London: Routledge & Kegan, pp. 156-198; ristampato in Id., *Philosophical Papers*, Cambridge: Cambridge University Press, 1990, pp. 52-94.
- RITOV, I. e BARON, J. (1990): "Reluctance to Vaccinate: Omission Bias and Ambiguity", *Journal of Behavioral Decision Making*, 3, pp. 263-277.
- SACKETT, D.L. et al. (1997): *Evidence-based Medicine. How to Practice and Teach EBM*, London: Churchill Livingstone (trad. it. *La medicina basata sull'evidenza. Come insegnare e praticare l'EBM*, Torino: Centro Scientifico, 2000).
- SAVAGE, L.J. (1954): *The Foundations of Statistics*, New York: Wiley.
- SCANDELLARI, C. (2005): *La diagnosi clinica. Principi metodologici del processo decisionale*, Milano: Masson.
- SCHMIDT, W. M. (1976): "Health and Welfare of Colonial American Children", *American Journal of Diseases of Children*, 130, pp. 694-701.
- SCHWARTZ, W. B. (1979): "Decision Analysis: A Look at Chief Complaints", *New England Journal of Medicine*, 300, pp. 556-559.
- SCHWARTZ, J. A. e CHAPMAN, G. B. (1999): "Are More Options Always Better? The Attraction Effect in Physicians' Decisions about Medications", *Medical Decision Making*, 19, pp. 315-323 (trad. it. "È sempre preferibile avere più opzioni? L'effetto attrazione nelle scelte di trattamento", in Crupi, Gensini e Motterlini, a cura di, 2006, pp. 141-157).
- SHAFIR, E., SIMONSON, I. e TVERSKY, A. (1993): "Reason-Based Choice", *Cognition*, 49, pp. 11-36 (trad. it. "Scelta e ragioni", in M. Motterlini e F. Guala, a

- cura di, *Introduzione all'economia cognitiva e sperimentale*, Milano: UBE, 2005, pp. 167-194).
- SHULKIN, J. (2000): "Decision Sciences and Evidence-Based Medicine – Two Intellectual Movements to Support Clinical Decision Making", *Academic Medicine*, 75, pp. 816-818.
- SIMON, H.A. (1957): *Models of Man: Social and Rational*, New York: Wiley.
- (1983): *Reason in Human Affairs*, Stanford (Calif.): Stanford University Press.
- STICH, S. (1990): *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*, Cambridge (Mass.): MIT Press (trad. it. *La frammentazione della ragione*, Bologna: Il Mulino, 1997).
- STIGGELBOUT, A.M. (2000): "Assessing Patients' Preferences", in Chapman e Sonnenberg (a cura di) (2000, pp. 289-312).
- SUTHERLAND, H. J. e TILL, J. E. (1993): "Quality of Life Assessments and Levels of Decision Making: Differentiating Objectives", *Quality of Life Research*, 2, pp. 297-303.
- THORTON, J. G., LILFORD, R. J. e JOHNSON, N. (1992): "Decision Analysis in Medicine", *British Medical Journal*, 304, pp. 1099-1103.
- TVERSKY, A. e SHAFIR, E. (1992): "The Disjunction Effect in Choice under Uncertainty", *Psychological Science*, 3, pp. 305-309.
- UBEL, P.A. *et al.* (2005): "Misimagining the Unimaginable: The Disability Paradox and Health Care Decision Making", *Health Psychology*, 24, pp. S57-S62.
- VON NEUMANN, J. (1928): "Zur Theorie der Gesellschaftspiele", *Mathematische Annalen*, 100, pp. 295-320 (trad. ingl. "On the Theory of Games of Strategy", in A. W. Tucker e R. D. Luce, a cura di, *Contributions to the Theory of Games*, Princeton (N.J.): Princeton University Press, 1959, vol. IV, pp. 13-42).
- VON NEUMANN, J. e MORGENSTERN, O. (1944/1947): *Theory of Games and Economic Behavior*, Princeton (N.J.): Princeton University Press.
- WEINSTEIN, M. C. e FINEBERG, H. V. (1980): *Clinical Decision Analysis*, Philadelphia (Pa.): Saunders (trad. it. *L'analisi della decisione in medicina clinica*, Milano: Franco Angeli, 2008).

A World in Time

Daniele Chiffi
University of Salento
e-mail: chiffidaniele@gmail.com

1. Rotating worlds
2. Gödel, Kant and the concept of time
3. Ideality of time
4. Possible worlds. Ontology and cognition
5. Concluding remarks

ABSTRACT. This paper is intended to analyse the concept of time in the Gödelian models of Einstein's field equations, by showing the intimate connections between Gödel's views on (the philosophy of) logic and mathematics and his cosmological work. He came up with these models, when he was working on a paper for a book of the *Library of Living Philosophers* devoted to Einstein about the concept of time in theory of relativity and in the Kantian philosophy. Gödel provided the mathematical structure of some Rotating worlds which are coherent with the theory of relativity, but within which an objective dimension of time cannot be constructed (see Section 1). In Section 2, I underline the role of the Kantian philosophy on time and the particular interpretation that Gödel has offered, which has led him to become a holder of the idealistic view on time. Section 3 will be devoted to Gödel's argument for the ideality of time, while in Section 4 I will make the role of the concept of "possible world" in Gödel's works clear as well as his importance for his views on time. Finally, in Section 5 I will present my concluding remarks which show the link between the Kantian dimension of time in modern physics and the structure of our cognition.

KEYWORDS: time, Gödel, Rotating worlds, possible worlds.

Ogni cosa mortal Tempo interrompe.

F. PETRARCA, *Trionfi*

1. Rotating worlds

The concept of time is a perennial issue in philosophy and in the sciences. It is quite standard to follow McTaggart's distinction between the A-series of time, in which events can be ordered by way of the non-relational predicates "is past", "is present" and "is future", and the B-series in which events can be ordered according to the relational predicates "comes before" and "comes after". Moreover, it is also important to make a distinction between an idealistic view on time and a realistic one, in order to have a better framework on the problem of time. Idealists maintain that time and change are not real, because time is Plato's "moving image of eternity", while realists maintain that time is an essential relation of the events which occur in our world.

After the formulation of the theory of relativity, many philosophers have argued that the philosophical problem of time still remains. An exception to this claim has been provided by Putnam (1967): "I do not believe that there are any longer any philosophical problems about Time; there is only the physical problem of determining the exact physical geometry of the four dimensional continuum we inhabit". On the contrary, Gödel argues: "Time remains the philosophical problem even after Einstein" (Wang 1996); and: "Time is that mysterious and seemingly self-contradictory being which, on the other hand, seems to form the basis of the world's and our own existence" (Gödel 1949b).¹ Gödel acknowledges that time is a fundamental issue in philosophy and that the question of time cannot only be placed upon the framework of the theory of relativity. The concept of time in the theory of relativity can be well understood if we read the following quote from Heisenberg:

In the theory of relativity future and past are separated by a finite time interval the length of which depends on the distance from the observer. Any action can only be propagated by a velocity smaller than or equal to the velocity of light. Therefore, an observer can at a given instant neither know of nor influence any event at a distant point which takes place between two characteristic times. The one time is the instant at which a light signal has to be given from the point of the event in order to reach the ob-

¹ Yourgrau (2005) is an important monograph dedicated to Gödel and Einstein and their works in the theory of relativity. The previous quotes can be found in Yourgrau's book too.

server at the instant of observation. The other time is the instant at which a light signal, given by the observer at the instant of the observation, reaches the point of the event. The whole finite time interval between these two instants may be said to belong to the “present time” for the observer at the instant of observation. Any event taking place between the two characteristic times may be called “simultaneous” with the act of observation (1958, p. 71).

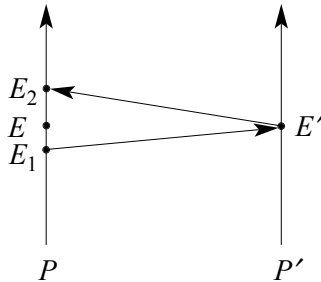
Thus, the “concept of time” needs the “concept of simultaneity” if we want to understand the difference between past and future and to grasp the relativistic present, otherwise we cannot apply any time-metrics, e.g., we need simultaneity between an event and a clock. Furthermore, our conception of time requires inertia, since we need to place the events in time according to the Kantian “enduring magnitude”, which can be considered as a sort of physical inertia. The concept of inertia plays a very important role in the development of the theory of relativity. In effect, Gödel has proposed a solution of Einstein’s field equations in the theory of relativity, which allows for Rotating worlds, i.e., worlds where the “compass of inertia” rotates everywhere relative to matter, namely it rotates relative to the totality of the galactic systems with angular velocity $\sqrt{2(\pi\kappa\rho)}$, where ρ is the density of matter and κ is Newton’s gravitational constant.

A Rotating world (R-world) presents an inertial field, which can be considered as a possible “substitute” for the “absolute space”, which determines the motion of bodies upon which no force acts. The inertial field determines the behaviour of the axis of the pendulum at the pole, where no force acts on it. Matter rotates with respect to this inertial field of the pendulum. Hence, the R-world does not rotate itself as a whole with respect to a particular axis, but this is a different kind of inertia which allows the rotation of the world. Nevertheless, the mutual distances of any two material particles remain constant all of the time. R-worlds present Timelike Closed Lines, which are not geodesics (namely, the shortest trajectories between two points in a particular space). “Timelike” means that the space-time lines do not exceed the speed of the light, while “close” means that the lines create a loop, which allows for time travels. Notice that it is possible to get only a partial ordering between events, because if one tries to define an absolute “before”, then the events cannot be compared or they lay in a cyclic ordering. Moreover, there is no objective lapse of time and there is no linear ordering because the temporal distances lack the property of additivity in a metric space. Namely, what Gödel wants to show is that our intuitive concept of “time” cannot be elucidated in an R-world.²

² See Yourgrau (2005).

He observes that if two events are simultaneous for one observer, it can be the case that they are not simultaneous for another observer, if he is in motion relative to the first observer. Hence, Gödel remarks that the statement: “‘A before B by t seconds’ is certainly something not inherent in the events, nor does it allow time to be a substratum in which the events are lying” (Gödel 1946/49, p. 250).

Notice that in the R-worlds, differently from the theory of relativity, it is not possible that a first signal (as presented in Reichenbach [1928] 1958, Section 22) can occur. To clarify the concept of *first-signal*, imagine the following situation. We want to send a signal from P to P' . We call E_1 the event of the departure of the signal from P , while E' is the event of its arrival on P' . Simultaneously at the arrival of the signal, another signal is sent from P' to P . Let the event of the arrival of the signal from P' to P be called E_2 . Therefore, the order of E' relative to E_1 and E_2 is the following: $E' > E_1$ and $E' < E_2$. Consider now an event E on P for which $E_1 < E < E_2$. We could ask which is the position of E relative to E' . If the first signal is sent from P then $E > E'$, while if the first signal is sent from P' then $E' > E$. Thus, E and E' cannot be correlated, namely they are indeterminate in respect to the time order. Of course, this requires the velocity of light to be limited, since if one assumes the Newtonian framework within which the velocity of light is infinite, then the interval $E_1 - E_2$ will be reduced to a point E_0 . In addition, if we accept that the velocity of the signal is finite, then the first signal determines a finite interval and not a pointlike event, e.g., we can provide a topological definition such that “any two events which are indeterminate as to their time order may be called simultaneous” (*ibid.*, p. 145) and therefore any event which occurs in P between E_1 and E_2 is indeterminate and simultaneous with E' .



In case of an R-world, we cannot have a first signal, since from every space-time point a signal can be sent to every other point, which goes back to its de-

parture point, since the trajectory is a circle. This implies that every observer has his own “now”, which cannot represent the objective flow of time. A similar point on the role that the concept of “now” plays was addressed by Einstein in some conversations with Rudolf Carnap, which have been reported in his intellectual autobiography.³

Gödel observes that another important feature is that his solution of Einstein’s field equations seems to be not coherent with *Mach’s principle*, which is, in any case, a principle logically independent of the theory of relativity. Mach’s principle greatly inspired Einstein in the development of the theory of relativity. This principle states that: “mass there influences inertia here”, i.e., position and velocity are relative concepts and this fact implies that even our concept of inertia needs to be relativized.⁴ Mach presented this principle in his *Mechanics* (but the principle is nowadays known after that Einstein popularized it in 1918) where he attacked the Newtonian idea of the absolute space and the absolute time, because he argues that all the principles of mechanics are based on our experience of relative locations and relative motions, hence also the concept of inertia should be revisited.⁵ Thus, according to Mach, in-

³ “Once Einstein said that the problem of the Now worried him seriously. He explained that the experience of the Now means something special for man, something essentially different from the past and the future, but that this important difference does not and cannot occur within physics. That this experience cannot be grasped by science seemed to him a matter of painful but inevitable resignation. I remarked that all that occurs objectively can be described in science; on the one hand the temporal sequence of events is described in physics; and, on the other hand, the peculiarities of man’s experiences with respect to time, including his different attitude towards past, present, and future, can be described and (in principle) explained in psychology. But Einstein thought that these scientific descriptions cannot possibly satisfy our human needs; that there is something essential about the Now which is just outside the realm of science. We both agreed that this was not a question of a defect for which science could be blamed, as Bergson thought. I did not wish to press the point, because I wanted primarily to understand his personal attitude to the problem rather than to clarify the theoretical situation. But I definitely had the impression that Einstein’s thinking on this point involved a lack of distinction between experience and knowledge. Since science in principle can say all that can be said, there is no unanswerable question left. But though there is no theoretical question left, there is still the common human emotional experience, which is sometimes disturbing for special psychological reasons” (Carnap in Schilpp 1963, pp. 37-38).

⁴ In fact, in general relativity, the *principle of equivalence* – between the “freely falling trajectories in a *gravitational field* and the trajectories described by the free particles affected by no force at all – replaces the classical law of inertia holding in both Newtonian mechanics and special relativity” (Friedman 2008, p. 108).

⁵ Isaac Newton presented his well-known “rotating bucket argument” in order to show the difference between an absolute motion and a relative one and between the absolute space and time and their relative counterparts. Mach would interpret that experiment without any reference

ertia is not an internal property of a single body anymore, but rather the consequence of the influence of all the other bodies of the universe. But in the R-worlds the inertial field is independent of the state of motion of matter, hence Mach's principle does not hold in such kind of world (see Gödel 1949c, p. 271).

Another feature of the R-worlds is that we could travel *in principle* in every region of the past, the present and the future with a space-shuttle, but actually we cannot travel in time since the fuel needed by the shuttle would be too many times greater than the weight of the shuttle. Nevertheless, Gödel argues that the impossibility which occurs in practice can be discovered later as a kind of impossibility which occurs also *in principle*⁶ as in the case of the Uncertainty principle.

2. Gödel, Kant and the concept of time

Gödel maintains that it is possible to assign to the concept of time both a negative and a positive definition. According to the negative definition, time is not a characteristic or ordering inherent in the objects, while according to the positive definition, time is a characteristic concerning the relation of the objects with something else and time is the a priori condition of possibility of our sensibility. Furthermore, Gödel maintains that our intuitive conception of time implies that time is a system of one-dimensional points, isomorphic with a straight line and that time flows.

Gödel considers the theory of relativity as a step further with respect to the Kantian philosophy towards a more objective grasp of the concept of time. Nevertheless, he argues that the idealistic account of time cannot be falsified by the theory of relativity itself, because of the difficulties to compare time in physics with time in our everyday experience. He writes:

Change becomes possible only through the lapse of time. The existence of an objective lapse of time, however, means (or, at least, is equivalent to the fact) that reality consists of an infinity of layers of "now", which comes into existence successively. But, if simultaneity is something relative in the sense just explained, reality cannot be split up into such lay-

towards absolute concepts and would modify Newton's concept of inertia. For the history of the concept of Mach's principle, see Lichtenegger and Mashhoon (2007).

⁶ On the distinction between "effective" and "in principle", see Pataut (2002) which is an interview to Dummett.

ers in an objectively determined way. Each observer has his own set of “nows”, and none of these various systems of layers can claim the prerogative of representing the objective lapse of time (Gödel 1949b, pp. 202-203).

It seems that Gödel considers the hypothesis for which the concept of “present” can be not a point-like dimension of the objective time, but a subjective dimension of time with duration which splits up the events in time relative to an observer. This latter view resembles the construction of the intuitionistic continuum and the *ur-intuition* (or *two-ity*) of mathematics.⁷ But Gödel claims that such subjective (intuitionistic) dimension of time cannot stand for any temporal objective level, because it is relative to the position of the observer and because the concept of existence cannot be relativized to a particular observer in a particular context.⁸ Using a modern terminology, we can say that Gödel maintains an endurantist view in ontology, since objects exist without any reference to temporal parts and temporal dimensions.⁹ I will come back on the issue of Gödel’s account of the concept of existence in Section 4, because

⁷ “[...] the ur-intuition of mathematics (and of every intellectual activity) is the substratum, divested of all quality, of any perception of change, a unity of continuity and discreteness, a possibility of thinking together several entities, connected by a ‘between’, which is never exhausted by the insertion of new entities. Since continuity and discreteness occur as inseparable complements, both having equal rights and being equally clear, it is impossible to avoid one of them as a primitive entity, trying to construe it from the other one, the latter being put forward as self-sufficient; in fact it is impossible to consider it as self-sufficient. Having recognized that the intuition of continuity, of ‘fluidity’, is as primitive as that of several things conceived as forming together a unit, the latter being at the basis of every mathematical construction, we are able to state properties of the continuum as a ‘matrix of points to be thought of as a whole’” (Brouwer 1907, p. 17).

⁸ On the constitutive dimension of the intuitionistic intervals of time (which have a duration and need an intuitionistic *continuum*) as opposed to the realistic views on time, see Dummett (2000).

⁹ Recently, some analytical philosophers have started being interested in Gödel’s R-worlds, because of their importance in ontology. Ted Sider (2001) maintains that endurantism cannot hold in an R-world. He endorses Lewis’ view for which temporary intrinsic properties (such as “being charged”) exist as dyadic relations between objects and times. But in the R-world it is impossible to define time, therefore endurantism does not hold, because there is no dyadic relation with time. Effingham and Melia (2007) have replied to Sider that the endurantist position can hold in the R-worlds if we analyse the property of “being charged” as a relation between individuals and points of a manifold, which can stand for times, or temporal parts or just points in a structure without any reference to times when, for instance, they are interpreted in the R-worlds. In any case, Gödel’s account of the concept of “existence” of an object does not require any dyadic temporal relation, because it is an absolute concept.

firstly I want to analyse the Kantian features of time, which turn out to be important for the Gödelian remarks on the concept of time.

According to Kant, we cannot experience “Time itself” (B.219) and we cannot place the objects in this absolute time. Hence, we need an indirect way to determine the temporal ordering of the objects through the distinction between a subjective and an objective dimension of time which is very important in Kant’s philosophy, i.e., the cognitive dimension of time does not coincide with the objective account of it.¹⁰ This distinction is based on the further dichotomy between the constitutive figures of experience and the regulative ones (which rule a pre-existing state of affairs). Consider the structure of the *Critique of Pure Reason*. In the Analytic of Principles we apply every particular category to the sensible intuition, while this function was done in the transcendental deduction a priori and in a more general way. In this way, Kant makes a distinction between the Axioms of Intuition and the Anticipations of Perception, which are constitutive features of the experience and the Analogies and the Postulates of the Empirical Thought, which are regulative. The second and the third analogy of experience present the dimension of causality and the connections among objects through their mutual interaction. By contrast, in the Transcendental Aesthetic, the constitutive element of space and time is almost merely subjective and it is an *empty* form of sensibility, while with the analogies we achieve a more objective framework about the nature of the objects, where space and time are conceived in their physical fullness.¹¹ Moreover, Kant refers to the *modi of time* (namely Persistence – time as an enduring magnitude –, Succession and Simultaneity), because he wants to present a unitary framework for time, which can allow for an adequate interplay between the subjective (and constitutive) dimension of time and the objective (and regulative) dimension of it. Persistence refers to the concept of time intended as a magnitude or a force, while Succession and Simultaneity alone cannot achieve to present the “unity of time” without a persisting time.¹²

¹⁰ See the example of the ship and the house in the *Critique of Pure Reason*, B 232.

¹¹ See Watkins (2005). The Kantian distinction between the *regulative* and the *constitutive* features of knowledge was taken into consideration by many philosophers. See Quine (1948), where this distinction is presented making a comparison between the Medieval dispute over universals and the philosophy of mathematics of the first half of the last century. Quine noticed that the constitutive aspects of our knowledge correspond to an intuitionistic account of mathematics, while classical mathematics deals with the entities which pre-exist to any definition, namely a regulative role. This distinction plays also an important role in philosophy of language. See Searle (1969, pp. 33-42).

¹² Notice that there are some formal counterparts of these Kantian *modi of time*, e.g. Succession and Simultaneity are almost equivalent to the relations *P* (precedes) and *O* (overlaps) in

Gödel's remarks on Kant's conception of time do not take into account the dimension of the objective time. In Gödel (1946/49) there is no discussion about the Analogies of the Experience, because he bases his Kantian remarks especially on the *Prolegomena* and the Transcendental Aesthetic. As a matter of fact, the Analogies of Experience are concerned with the existence of appearances, the Transcendental Aesthetic with the mere form of their sensible intuition (see Longuenesse 1998, chapter 11). Therefore, he interprets Kant's concept of time in a subjectivist way, because it is in the Analogies of Experience that time has a more objective dimension. Nevertheless, some Gödelian remarks on the Kantian structure of cognition could result interesting, even when they are compared with the theory of relativity and the non-Euclidean geometry. Consider the following quote:

The fact that the physical bodies surrounding us move by the laws of a non-Euclidean geometry does not exclude in the least that we should have a Euclidean "form of sense perception", i.e. that we should possess an a priori representation of Euclidean space and be able to form images of outer objects only by projecting our sensations on this representation of space, so that, even if we were born in some strongly non-Euclidean world, we would nevertheless invariably imagine space to be Euclidean, but material objects to change their size and shape in a certain regular manner when they move with respect to us or we with respect to them (1946/49, p. 255).

Gödel wants to show that our Euclidean intuition would be a priori valid even if the structure of the world were non-Euclidean. Gödel states that our Euclidean intuition is innate and a priori valid even in a non-Euclidean world. Actually, Klein proposed in 1871 a simple projective model of hyperbolic geometry, in which the fundamental entities are the same entities of the Euclidean geometry and in which the points of the geometry are in an n -dimensional disk, or ball, and the lines of the geometry are line segments contained in the disk. In this way it is possible to map non-Euclidean objects with Euclidean ones.¹³

Gödel argues that the temporal relations represent relations between the perceiving subject and the objects. Namely, *Schein* stands for the appearance of the things in themselves. *Erscheinung* stands for the appearance of the

the Russell-Walker calculus (see Russell 1914, 1936; Walker 1947), while Persistence is equivalent with the concept of fluent in the event-calculus (see van Lambalgen and Hamm 2004).

¹³ A similar idea was envisaged in 1884 by Edwin Abbott in his novel *Flatland: A Romance of Many Dimensions*.

things with the sensibility of the subject. Appearances are directed towards objects of thought which have the properties of the objects of our sensibility and they are neither mere illusions nor properties of the things in themselves, but their possibility is justified by the existence of the things in themselves. Having this framework in mind, we can analyse Gödel's argument of the a priori validity for the Euclidean intuition. The apriority of space and time refers to the objective counterpart of these ideas but, through the *Erscheinung*, the appearances of the objective space and time are connected with the innate schema of sensibility of the subject. Therefore, even the subjective space and time are respectively applicable in the forms of Newtonian time and Euclidean space to the physical reality (as a kind of relative space and time) as we saw in the case of Klein's geometrical translation. Notice that the arguments about the apriority of space and time hold if one maintains that there is an innate Euclidean intuition, which is merely assumed by Gödel. When Gödel refers to the Kantian objectivity of time and space, he does not take into account the penetrating analysis of these concepts in the *Transcendental Analytic* (as I have stressed above). Gödel's epistemic views support the idea that we can construct our knowledge with different levels of objectivity of the phenomena that can lead us towards the very essence of the things in themselves (noumena). Hence, the Kantian world of appearances (with "many subjective elements of the world of the sensations" according to Gödel) is considered to be the first step of the objectivation of reality if it is compared with the theory of relativity, which is a further step in this process with the aim to go beyond the appearances and approach the things in themselves partially and step by step.

3. Ideality of time

Gödel presented an argument to show the idealistic features of time in his R-worlds and in our actual world. Apparently, such claim does not seem to fit with Gödel's Platonistic views, but a further analysis can explain this apparent contrast between his ontology and his account of the concept of time. Gödel's argument shows the ideality of time in the R-worlds and from that he infers the ideality of time in our world.¹⁴ The complete argument runs in the following way:

¹⁴ Gödel partially assumes a Leibnizian point of view on the nature of the individuals and the physical laws. In fact, Leibniz believed that an individual is an inhabitant of one single world, and every other world with different physical laws would determine a different individ-

- (0) Time is real only if change is real;
- (1) Change is real only if there exists an objective lapse of time;
- (2) Time is real only if there exists an objective lapse of time [from (0) and (1)];
- (3) “The existence of an objective lapse of time means or at least is equivalent to the fact, that reality consists of an infinity of layers of ‘now’ which come into existence successively” (Gödel 1949b, p. 202);
- (4) Reality consists of an infinity of layers of “now” which come into existence successively only if space-time admits of a global time function (*cosmic time*);
- (5) Time is real only if space-time admits of a global time function [from (2), (3) (4)];
- (6) Gödel’s rotating-model M is a solution to Einstein’s field equations, hence it is a physically *possible* model;
- (7) Since for every x in M , x chronologically precedes itself, M does not possess a global time function;
- (8) In the physically possible world M , time is ideal [from (5) (6) (7)];
- (9) The main, contingent, non-lawlike difference between M and our universe is given by the (probable) absence of a net rotation of matter, which implies the existence of cosmic time in our world;
- (10) ?;

therefore, time is ideal also in our universe (this rigorous presentation of the argument is due to Dorato 2001).

Gödel assumes that change and time are real if an objective lapse of time exists. Then, he assumes that reality consists of layers of now that come to exist. In Savitt (2006) it is argued that this fact can lead, in principle, towards two different views on the metaphysics of time: *presentism* (the existence of the objects is relativized to the present, but they cease to exist after) and *possibilism* (the existence of the objects occurs in the present but it persists in the future). I maintain that Savitt’s analysis does not handle correctly Gödel’s views on the concepts of existence and time. Gödel has an absolute account of the concept of existence and he wants to determine the nature of time making saved his ontology. Moreover, Gödel could not use any metaphysical determi-

ual. In fact, if an individual X is determined by the set of his properties, then X can not inhabit two worlds w_1 and w_2 with different physical laws, since X in w_1 has at least different spatio-temporal properties in respect to X in w_2 .

nations of the objects in time, since the aim of his argument is to show the nature of time itself. Thus, Savitt's claim cannot be referred to Gödel's argument. In any case, the most critical point in Gödel's argument is that to infer the impossibility of a global time function in the R-worlds, which leads to an antirealist account of time in the R-worlds and in our universe, too.

Many attempts to understand the conclusion of this argument have been presented and some of them have been focused on the Gödelian ontology, notably on the role of the concept of "possible world". Notice that in 1949 there was no Kripke semantics for modal logic, but only a semantics for Lewis' modal logic system S_5 due by Carnap with the concept of "state description" (Carnap 1947). A state description is a class of sentences, which contains for every atomic sentence either this sentence or its negation, but not both, and no other sentences such that the *intension* of a term is a function from state descriptions to objects, while a predicate is a function from state descriptions to sets of objects. Thus, the concept of (logical) "necessity" is defined as true in every state description, while the concept of "possible" is defined as true at least in one state description. Notice that Carnap's framework does not deal with the different accessibility relations among state descriptions¹⁵ (or possible worlds) as we do in Kripke's frames for modal logic. Hence, the issue of the accessibility among different worlds was still very problematic in 1949, when Gödel wrote his argument for the ideality on time.

4. Possible worlds. Ontology and cognition

Jaakko Hintikka claims that Gödel maintains a "one-world assumption", since his concept of truth does not refer to the concept of "true in all possible worlds", but "true in the actual world". He writes that "one central idea of Leibniz's was never taken up by Gödel: the idea of possible worlds. This is one of the many indications of Gödel's actualism" (Hintikka 2000, p. 47; see also Hintikka 1998). Hence, logical and abstract entities are placed in an abstract region of the actual world, not in different possible worlds. Our epistemic access to this abstract region is provided by the mathematical intuition, which extends the powers of ordinary sensible perception. Thus, if the domain of discourse is our single world, then there is a collapse of modalities, viz. what is possible is necessary and vice versa.

¹⁵ We can say that a state description is the syntactical counterpart of the concept of "possible world" (or of Wittgenstein's concept of "state of affairs") in the modal system S_5 .

Van Atten (2001) criticises Hintikka's interpretation, since he notices that Gödel uses a possible-world reasoning in the argument for the "ideality of time" in one of his essays on the rotating universes. It seems that Gödel is assuming that the concept of possibility involved in his above argument is the *logical* one, which allows for an Euclidean relation as in Lewis' modal system S_5 . Notice that, in his argument about the ideality of time, the concept of *physical* possibility occurs, while it seems that in point 10 of the argument the concept of logical possibility is involved. Writes Gödel:

The mere compatibility with the laws of nature of worlds in which there is no distinguished absolute time, and, therefore, no objective lapse of time can exist, throws some light on the meaning of the time also in those worlds in which an absolute time can be defined (1949b, p. 562).¹⁶

Van Atten maintains that there is no collapse between *possibility* and *necessity* in Gödel's modal views, but between the concept of *possibility* and the concept of *actuality*.¹⁷ Namely, if something is possible, then it is actual, i.e., it can occur in our world.

I argue that we need to consider the absolute account of the Gödelian concept of existence again: "The concept of existence [...] cannot be relativized without destroying its meaning completely" (1949b, p. 559). Hence, my view is that one cannot say (like Hintikka) neither that the concept of "possible world" does not play any role in Gödel nor that Gödel is an a holder of a mere actualism in ontology. My claim is that Gödel possesses the concept of "possible world", but the Gödelian possible worlds cannot be completely alternative to our actual world (in time), because of his account of the concept of "existence" and also because the accessibility relations among possible worlds did

¹⁶ For some Gödelian observations on the nature of time written in his philosophical notebooks, see the last chapter of Cassou-Nougues (2007).

¹⁷ Van Atten notes that the philosopher Nicolai Hartmann held an ontological view similar to Gödel's one, in the case of the ideal objects of our experience (see Hartmann 1937). The Kantian views on necessity, possibility and actuality are very different. In fact, according to Kant in the Postulates of the Empirical Thought we have that: "What agrees (in terms of intuitions and concepts) with the formal conditions of experience is *possible*"; "That whose coherence with the actual is determined according to universal conditions of experience is *necessary*"; and "What coheres with the material conditions of experience (with sensation) is *actual*" (A218/B266, my italics). Thus, in the Kantian views on the modalities there is an interplay between the ontological features of reality and our structure of the cognition (and intuition). Moreover, the Kantian modalities are *regulative* aspects of reality, while, according to Hartmann, modalities are the *constitutive* aspects of being.

not receive a proper formal treatment at the time. Namely, there cannot exist worlds which are completely independent of the actual world. Hence, if time turns out to be ideal in an R-world, then even in our actual world time cannot be real. Notice that this fact does not imply a form of actualism and does not coincide with the idea that the concept of possibility collapses into the concept of actuality, since the R-worlds are distinct from the actual world but their structure can influence our world.

A better understanding of Gödel's ontology can be achieved if one focuses on his ontology in the field of mathematics. As a matter of fact, the absolute Gödelian ontology entails that only static and atemporal objects can exist and this idea is completely different from the intuitionistic views on the nature of the mathematical objects which are dynamic and temporal. If an object is temporal, then it can be either *intratemporal* (namely, static and always existing) or *omnitemporal* (namely, an object comes to exist in a moment and ceases to exist in another moment). Van Atten observes that at first Husserl held an intratemporal view on the nature of the ideal objects, but after 1917 he realized that the temporal flow is a condition for the identity in their process of constitution.¹⁸ Husserl's later turned out to be not too much different from Brouwer's ones, in particular if one analyses the intuitionistic (law-like or free) choice sequences, they seem to be mathematical objects which behave in a very particular way, i.e., they are open only in respect to the future. Hence, one cannot say that choice sequences are completely *omnitemporal*, and, according to Gödel, they are not pure mathematical objects but objects which belong to applied mathematics, because of the role that time plays in their construction (see van Atten 2006). Namely, Gödel maintains that time cannot be a *constitutive* feature of the mathematical objects as in intuitionism. He wants to show with his argument of the ideality of time that our intuitive concept of time cannot be a *constitutive* element of the physical external world, too. Both claims are supported by his absolute ontology (which he calls "Platonism" or "conceptual realism"), in which time seems to play a relative role, while, from an epistemic point of view, he maintains that temporality is essentially connected with the mathematical intuition and the sense perception. In the 50's he writes:

The similarity between mathematical intuition and a physical sense is very striking. It is arbitrary to consider the proposition "This is red" an im-

¹⁸ This turn in Husserl's works coincides with a switch from the *static phenomenology* to the *genetic phenomenology*.

mediate datum, but not so to consider the proposition expressing *modus ponens* or complete induction (or perhaps some simpler proposition from which the latter follows). For the difference, as far as it is relevant here, consists solely in the fact that in the first case a relationship between a concept and a particular object is perceived, while in the second case it is a relationship between concepts. [...] There is no substantial difference between mathematics and other sciences (1953/59, p. 359).¹⁹

And in another Gödelian essay we can read:

That something besides the sensations actually is immediately given follows (independently of mathematics) from the fact that even our ideas referring to physical objects contain constituents qualitatively different from sensations or mere combinations of sensations, e.g. the idea of the object itself (1964, p. 271).²⁰

Thus, according to Gödel there is no difference – from a cognitive point of view – between the way in which we can grasp the meaning of the formal objects and the objects which we perceive in reality. Hence, there is also no cognitive distinction between the actual universe and his (abstract) possible R-worlds and that is why he can claim for the ideality of the time from the R-worlds to our actual world. In both worlds our Euclidean intuition is always the same and it allows to interpret the non-Euclidean worlds in accordance with some geometrical translations such as the one from Klein. In any case, the importance of the (visual and Euclidean) mathematical intuition was firmly criticised by the neopositivists, who argued that not everything which can be *conceived* has a visual and intuitive counterpart as in the case of a space with *n*-dimensions or a chiliagon.

5. Concluding remarks

The relation between ontology and cognition is the fundamental point of the Gödelian views on time. I have showed in the previous sections that it is not too easy to conciliate the absoluteness of Gödel's timeless ontology with the temporality of his idea of cognition and intuition. Moreover, his subjectivist (and

¹⁹ The quotation is taken from the fifth version of the essay “Is Mathematics a Syntax of Language?”; the second and the sixth version of the essay are published in Rodríguez-Consuegra (1995).

²⁰ This quote shows also Gödel's critics towards the associationist models of mind.

incomplete) reading of Kant's works does not provide any systematic framework within which to handle the problem of the nature of time, which can be ideal in the R-worlds but there is no evidence to support that such kind of ideality holds in our world in time. It is true that there is the *logical possibility* that if time is ideal in an R-world, then it is ideal in our world, even if the R-worlds are coherent with the theory of relativity. This means that not all the solutions of Einstein's field equations are *physically possible*, but there are some solutions, like the ones which allow for the R-worlds, which are just *logically possible*. One of the many reasonable ways to discern the solutions which are logically possible from the ones which are physically possible can be to consider all the solutions which violate Mach's principle as not physically possible, but only logically possible.²¹ In this way, the Kantian idea of causality as a relational *magnitudo*²² (or interaction between substances) of the third Analogy of Experience can be vindicated as a concept which turns out to be particularly connected with the subsequent Mach's principle. Thus, the Kantian causality is a heuristic principle either in the modern theory of relativity and in the structure of human cognition. Kant maintains in the third Analogy that each object is considered to be the cause (as well as the effect) of the specific determinations of all the things perceived as simultaneous along with it (Longuenesse 1998, chapter 11). But this interaction between objects is Mach's principle *in nuce* and this is a very important connection between Kant and Mach (and the early formulations of the theory of relativity) on this point.²³

REFERENCES

VAN ATTEN, Mark (2001): "Gödel, Mathematics, and Possible Worlds", *Axiomates*, 12, pp. 355-363.

²¹ See Horák (1979).

²² Kant observes that the knowledge of the *magnitudo* of the universe can only be *conditionally* analyzed and grasped also because we cannot intuit an infinitely large *magnitudo*. He also states: "we cannot therefore say anything at all in regard to the magnitude of the world, not even that there is in it a regress *in infinitum*. All that we can do is to seek for the concept of its magnitude according to the rule which determines the empirical regress in it. This rule says no more than that however far we may have attained in the series of empirical conditions, we should never assume an absolute limit, but should subordinate every appearance, as conditioned, to another as its condition, and that we must advance to this condition" (Kant [1781] 1965, B547-548).

²³ I am grateful to Arcangelo Rossi, Michiel van Lambalgen, Carlo Dalla Pozza and Mario Castellana for useful comments and suggestions.

- (2006): *Brouwer meets Husserl: On the Phenomenology of Choice Sequences*, Dordrecht: Springer.
- BROUWER, Luitzen Egbertus Jan (1907): *On the Foundations of Mathematics*, in *Brouwer Collected Works. I: Philosophy and Foundations of Mathematics*, edited by A. Heyting, Amsterdam: North-Holland, 1975, pp. 11-101.
- CARNAP, Rudolf (1947): *Meaning and Necessity. A Study in Semantics and Modal Logic*, Chicago: The University of Chicago Press.
- CASSOU-NOGUÈS, Pierre (2007): *Les démons de Gödel. Logique et folie*, Paris: Editions du Seuil.
- DORATO, Mauro (2001): “On Becoming, Cosmic Time and Rotating Universes”, in C. Callender (ed.), *Time, Reality and Experience*, Cambridge: Cambridge University Press, pp. 253-276.
- DUMMETT, Michael (2000): “Is Time a Continuum of Instants?”, *Philosophy*, 75, pp. 497-515.
- EFFINGHAM, Nikk and MELIA, Joseph (2007): “Endurantism and Timeless Worlds”, *Analysis*, 67, pp. 140-147.
- FRIEDMAN, Michael (2008): “Einstein, Kant, and the A Priori”, in M. Massimi (ed.), *Kant and Philosophy of Science Today, Royal Institute of Philosophy Supplement*, 63, pp. 95-112.
- GÖDEL, Kurt (1946/49): “Some Observations about the Relationship between the Theory of Relativity and Kantian Philosophy”, in *Collected Works*, edited by S. Feferman *et al.*, Oxford: Oxford University Press, 1995, vol. III, pp. 247-259.
- (1949a): “An Example of a New Type of Cosmological Solutions of Einstein’s Field Equations of Gravitation”, *Reviews of Modern Physics*, 21, pp. 447-450; reprinted in *Collected Works*, edited by S. Feferman *et al.*, Oxford: Oxford University Press, 1990, vol. II, pp. 190-198.
- (1949b): “A Remark about the Relationship between Relativity Theory and Idealistic Philosophy”, in P. A. Schilpp (ed.) (1963, pp. 555-562); reprinted in *Collected Works*, edited by S. Feferman *et al.*, Oxford: Oxford University Press, 1990, vol. II, pp. 202-207.
- (1949c): “Lecture on Rotating Universes”, in *Collected Works*, edited by S. Feferman *et al.*, Oxford: Oxford University Press, 1995, vol. III, pp. 269-287.
- (1953/59): “Is Mathematics a Syntax of Language?”, in *Collected Works*, edited by S. Feferman *et al.*, Oxford: Oxford University Press, 1995, vol. III, pp. 356-362.
- (1964): “What is Cantor’s Continuum Problem?”, in P. Benacerraf and H. Putnam (eds.), *Philosophy of Mathematics: Selected Readings*, 2nd ed., Cambridge: Cambridge University Press, 1983; reprinted in *Collected Works*, edit-

- ed by S. Feferman *et al.*, Oxford: Oxford University Press, 1990, vol. II, pp. 254-270.
- HARTMANN, Nicolai (1938): *Möglichkeit und Wirklichkeit*, Berlin: Walter de Gruyter & Co.
- HEISENBERG, Werner Karl (1958): *Physics and Philosophy: The Revolution in Modern Science*, London: Allen & Unwin.
- HINTIKKA, Jaakko (1998): "On Gödel's Philosophical Assumptions", *Synthese*, 114, pp. 13-23.
- (2000): *On Gödel*, Belmont (Calif.): Wadsworth.
- HORÁK, Zdeněk (1979): "Elementary Physical Approach to Mach's Principle and Its Observational Basis", *General Relativity and Gravitation*, 11, pp. 261-279.
- KANT, Immanuel ([1781] 1965): *Critique of Pure Reason*, trans. by N. Kemp Smith, New York: St. Martin's Press.
- LICHTENEGGER, Herbert and MASHHOON, Bahram (2007): "Mach's Principle", in L. Iorio (ed.), *The Measurement of Gravitomagnetism: A Challenging Enterprise*, New York: Nova Science, pp. 13-25.
- LONGUENESSE, Beatrice (1998): *Kant and the Capacity to Judge*, Princeton (N.J.): Princeton University Press.
- PATAUT, Fabrice (2002): "Truth, Meaning, Modalities and Ethics (A Second Interview with Micheal Dummett)", *Philosophical Investigations*, 25, pp. 225-271.
- PUTNAM, Hilary (1967): "Time and Physical Geometry", *Journal of Philosophy*, 64, pp. 240-247.
- QUINE, Willard Van Orman (1948): "On What There Is", *Review of Metaphysics*, 2, pp. 21-38.
- REICHENBACH, Hans ([1928] 1958): *The Philosophy of Space and Time*, New York: Dover Publications.
- RODRÍGUEZ-CONSUEGRA, Francisco A. (1995): *Kurt Gödel. Unpublished Philosophical Essays*, Basel-Boston-Berlin: Birkäuser Verlag.
- RUSSELL, Bertrand (1914): *Our Knowledge of the External World*, London: Allen and Unwin.
- (1936): "On Order in Time", *Proceedings of Cambridge Philosophical Society*, 32, pp. 216-228.
- SAVITT, Steven (2006): "Being and Becoming in Modern Physics", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/spacetime-bebecome>.
- SCHILPP, Paul Arthur (ed.) (1963): *The Philosophy of Rudolf Carnap*, LaSalle (Ill.): The Open Court Publishing Co.
- SEARLE, John (1969): *Speech Acts. An Essay in the Philosophy of Language*, Cambridge: Cambridge University Press.
- SIDER, Ted (2001): *Four-Dimensionalism*, Oxford: Oxford University Press.

- VAN LAMBALGEN, Michiel and HAMM, Fritz (2004): *The Proper Treatment of Events*, London: Routledge.
- WALKER, Arthur Geoffrey (1947): “Durées et instants”, *Revue Scientifique*, 85, pp. 131-134.
- WANG, Hao (1996): *A Logical Journey. From Gödel to Philosophy*, Cambridge: MIT Press.
- WATKINS, Eric (2005): *Kant and the Metaphysics of Causality*, Cambridge: Cambridge University Press.
- YOURGRAU, Palle (2005): *A World Without Time. The Forgotten Legacy of Gödel and Einstein*, New York: Basic Books.

Criteria for the characterization of token causation

Tyler J. VanderWeele

Departments of Epidemiology and Biostatistics, Harvard University

e-mail: tvanderw@hsph.harvard.edu

1. Introduction
2. The Definition of Actual Cause in Halpern and Pearl
3. The Implications of Two Examples
4. The Claim of the Relativity of Actual Causation
5. Criteria for the Characterization of Token Causation

ABSTRACT. It is argued that the definition of an “actual cause” given by Halpern and Pearl (2005) is neither necessary nor sufficient for some event to constitute an actual cause. Only when Halpern and Pearl sufficiently elaborate, in a manner of their own choosing, a structural model does their definition in fact identify the actual cause. It is proposed that a characterization of actual or token causation might be considered adequate if one could articulate either (1) a definition for the actual cause which supplies a necessary condition for an actual cause and is such that an elaborated structural model can always be constructed which will only identify the event or events constituting the actual cause; or (2) a definition for the actual cause which supplies a sufficient condition for an actual cause and is such that an elaborated structural model can always be constructed which will identify the event or events constituting the actual cause.

KEYWORDS: actual causation, causal diagrams, structural models, token causation

1. Introduction

The problem of providing a characterization of token causation or of an “actual cause” has long eluded philosophers (Hall and Paul, 2003; Collins et al., 2004). Although in every-day conversation we can often agree that one event was the cause of another, trying to articulate precisely the rules by which we come to this agreement proves difficult.

Halpern and Pearl (2005) recently proposed a definition for an “actual cause” which can be applied to causal diagrams which represent causal structural models as formulated by Pearl (1995, 2000) and Halpern (2000). They demonstrate that their definition, in conjunction with sufficiently elaborated structural models, handles well a number of examples which are often considered problematic in the philosophical literature, such as those described by Bennett (1993) and Hall (2004). Halpern and Pearl (2005) use “actual cause” to make clear that they are interested in statements of the form “ X caused Y (in this particular case)” as opposed to statements of the form “ X is a cause of Y (in general, but perhaps not in this particular case).” Statements of the former type are often referred to as instances of “token causation” and have been the subject of interest in many philosophical analyses as well as in legal decisions; statements of the latter type are often referred to as instances of “type causation” and often are of interest in the social and biomedical sciences.

In the examples that Halpern and Pearl consider, their definition can be used to correctly identify the event which intuition suggests is the actual cause. Often the crudest causal diagram or structural model to describe one of these examples is insufficient to identify correctly the actual cause of a particular event but an elaborated structural model represented by a causal diagram with additional nodes does allow the use of their definition to identify the actual cause. In this paper two examples are used to argue that the definition given by Halpern and Pearl is neither necessary nor sufficient for some event to constitute an actual cause. It is not the goal of this paper to show that every possible account of actual causation using structural equations will fail but merely that the account given by Halpern and Pearl is not adequate. It is furthermore proposed that a characterization of actual or token causation might be considered adequate if one could articulate either (1) a definition for the actual cause which supplies a necessary condition for an actual cause and is such that an elaborated structural model can always be constructed which will only identify the event

or events constituting the actual cause; or (2) a definition for the actual cause which supplies a sufficient condition for an actual cause and is such that an elaborated structural model can always be constructed which will identify the event or events constituting the actual cause.

We will begin with a brief description of the framework and the definition of actual cause proposed by Halpern and Pearl (2005). We will then consider two examples, one from Halpern and Pearl (2005) and one introduced here, to argue that the definition given by Halpern and Pearl is neither necessary nor sufficient for some event to constitute an actual cause. We then discuss Halpern and Pearl's position, implied by their definition, that whether an event is the actual cause of another event is relative to a particular causal model. Finally, we propose criteria for an adequate characterization of token causation. We do not give a definition that satisfies these criteria but merely indicate what criteria it seems a definition of "actual cause" ought to satisfy in order to be considered an adequate characterization of token causation.

2. The Definition of Actual Cause in Halpern and Pearl

The definition of an "actual cause" given by Halpern and Pearl (2005) makes use of structural models as proposed by Pearl (1995, 2000) and Halpern (2000) which are sometimes graphically represented by causal diagrams. Halpern and Pearl (2005) define a causal model as a set of exogenous variables \mathcal{U} determined outside the model, a set of endogenous variables \mathcal{V} determined inside the model, a set \mathcal{R} which for each variable $Y \in \mathcal{U} \cup \mathcal{V}$ specifies a range of possible values for Y and a collection of functions \mathcal{F} which for each $X \in \mathcal{V}$ contains a function F_X which maps a value in \mathcal{U} and a value in $\mathcal{V} \setminus X$ to a value in X . If we let z be any value of the variables in $\mathcal{V} \setminus X$, then the function $F_X(u, z)$ can be thought of as the value of X that would be obtained if the variables in \mathcal{U} took on the values u and the variables in $\mathcal{V} \setminus X$ were set to z . The causal relations amongst the endogenous variables in \mathcal{V} are sometimes represented on a causal diagram in which there is an arrow from $X_1 \in \mathcal{V}$ to $X_2 \in \mathcal{V}$ if F_{X_2} depends on the value X_1 obtains.

For illustration we consider an example given in by Halpern and Pearl (2005, Example 4.2) and adapted from Hall (2004) which we will return to below. The example concerns two rocks thrown at a bottle, one thrown by Suzy and one

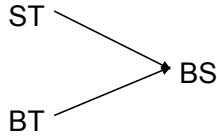


Figure 1: Example 4.2 from Halpern and Pearl (2005).

thrown by Billy; Suzy and Billy have perfect aim so that if either of them throw the rock the bottle will shatter. In the example of Halpern and Pearl, Suzy’s rock hits the bottle first and is the actual cause of the bottle’s shattering; however, if Suzy’s rock had not been thrown, then Billy’s rock would have hit the bottle and shattered it. Halpern and Pearl initially define three variables ST , BT and BS . The variable $ST = 1$ if Sally throws and 0 otherwise; the variable $BT = 1$ if Billy throws and 0 otherwise; the variable $BS = 1$ if the bottle shatters and 0 otherwise; all three variables are taken as endogenous. The variable $BS = 1$ if either $ST = 1$ or $BT = 1$ and 0 otherwise; the variable ST , however, does not depend on BT or BS and likewise the variable BT does not depend on ST or BS . The structural equations for the example are given by:

$$\begin{aligned}
 ST &= 1 \\
 BT &= 1 \\
 BS &= ST \vee BT.
 \end{aligned}$$

We thus could graphically represent the causal relations as given in Figure 1. Halpern and Pearl (2005) make the following proposal for a definition of an actual cause.

Definition 1 (Halpern and Pearl 2005) $X = x$ is the actual cause of some event ϕ if the following three conditions hold:

(AC1) Both $X = x$ and ϕ actually occur.

(AC2) There exists a partition (Z, W) of \mathcal{V} with $X \subseteq Z$ and some setting (z', w') of (Z, W) such that if z^* denotes the actual value of Z then both

- (a) changing (X, W) from its actual value (x, w) to (x', w') changes ϕ from true to false and
- (b) for all subsets W' of W and all subsets Z' of Z , setting W' to its value in w' and setting Z' to its value in z^* will leave ϕ true provided X is set to its actual value x .

(AC3) If X consists of multiple variables then no subset of X is such that conditions (AC1) and (AC2) can be satisfied for that subset.

If $X = x$ and ϕ satisfy conditions (AC1)–(AC3) then Halpern and Pearl (2005) say that $X = x$ is the actual cause of ϕ . Halpern and Pearl (2005) go on to discuss a possible refinement of the definition just given which restricts the endogenous variables to a set of “allowable settings.” The definition just given above is then a special case of their refined definition. However, the comments that we make below apply also to their refined definition and for simplicity we will thus consider the definition of actual cause just given.

3. The Implications of Two Examples

As noted above, in their example 4.2, Halpern and Pearl (2005) consider two rocks thrown at a bottle, one by Suzy and one by Billy; Suzy’s rocks hits the bottle first and is the actual cause of the bottle’s shattering; however, if Suzy’s rock had not been thrown, then Billy’s rock would have hit the bottle and shattered it. In the simplest causal diagram that represents these causal relationships as given in Figure 1, the definition of an actual cause given by Halpern and Pearl identifies the throwing of both rocks as actual causes of the bottle’s shattering. If ϕ is taken as $BS = 1$ then for Sally’s throwing $X = ST = 1$ if we let $W = BT$ and $(x', w') = (0, 0)$ we have (AC1) $ST = 1$ and $BS = 1$, (AC2a) $BS = 0$ if $ST = 0$ and $BT = 0$, (AC2b) $BS = 1$ regardless of whether $W = BT$ takes the value 0 or 1, provided $ST = 1$ and (AC3) holds trivially. Thus the definition of Halpern and Pearl implies that Sally’s throwing is the actual cause of the bottle’s breaking. However, applying the same reasoning to Billy’s throwing $X = BT = 1$, if we let $W = ST$ and $(x', w') = (0, 0)$ we have (AC1) $BT = 1$ and $BS = 1$, (AC2a) $BS = 0$ if $BT = 0$ and $ST = 0$, (AC2b) $BS = 1$ regardless of whether $W = ST$ takes the value 0 or 1, provided $BT = 1$ and (AC3) holds trivially. Thus the definition of Halpern and Pearl applied to Figure 1 implies also that Billy’s throwing is the actual cause of the bottle’s breaking; this is contrary to intuition. To remedy this problem, Halpern and Pearl consider a more elaborate causal diagram which, when used in conjunction with their definition, correctly identifies Suzy’s rock and not Billy’s rock as the actual cause of the bottle’s shattering. Consider the causal diagram given in Figure 2, for example, where $SH = 1$ if Sally’s rock hits

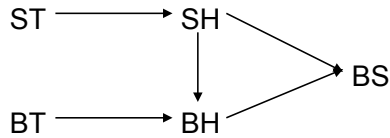


Figure 2: Elaborated causal diagram for Example 4.2 of Halpern and Pearl (2005).

the bottle and $SH = 0$ otherwise and $BH = 1$ if Billy’s rock hits the bottle and $BH = 0$ otherwise.

The structural equations for the elaborated description of the causal relationships amongst the variables are given by

$$\begin{aligned}
 ST &= 1 \\
 BT &= 1 \\
 SH &= ST \\
 BH &= BT \ \& \ (\text{not } SH) \\
 BS &= SH \vee BH.
 \end{aligned}$$

Halpern and Pearl go on to show that if their definition is applied to Figure 2 then it correctly identifies Suzy’s throwing the rock and not Billy’s throwing the rock as the actual cause of the bottle’s shattering. In any case, the simplest causal diagram for this example with two rocks demonstrates that the definition of Halpern and Pearl is not *sufficient* for a particular event to be an actual cause because, for the simplest causal diagram given in Figure 1, their definition implies that Billy’s throwing the rock was an actual cause of the bottle’s shattering. In defending their definition, Halpern and Pearl state (p. 845) that “the truth of every claim must be evaluated relative to a particular model of the world.” We will return below to the question of whether it is reasonable to claim actual causation is relative in this way.

Note that there are certain cases in which two events in some way interact so that the effect comes about if and only if both events occur; in such cases, it is appropriate to speak of two actual causes and an adequate characterization of what constitutes an actual cause will be able to handle these cases with two or more actual causes. In the case of the bottle’s shattering, however, we would

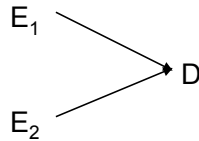


Figure 3: Example showing the definition of Halpern and Pearl (2005) is not a necessary condition for an event to be an actual cause.

ordinarily only say that Suzy’s throwing her rock was the actual cause of the bottle’s shattering.

We now consider a second example that demonstrates that the definition of Halpern and Pearl is also not a necessary condition for an event to be an actual cause. Suppose that a man is suffocating and if left untreated he will die at time t . A physician is present and has two possible treatments he can administer: injection 1 and injection 2. If injection 1 is given the man will be able to breathe; however, he is allergic to the chemical compounds constituting injection 1 and will thus die of heart failure, again at time t , resulting from an allergic reaction to injection 1. If injection 2 is given the man will once again be able to breathe but he is also allergic to the chemical compounds constituting injection 2 and thus if he is given injection 2 he will once again die of heart failure at time t . If he is given both injections 1 and 2, the two sets of chemicals interact and although the man is once more able to breathe, he will die of a failure of the nervous system at time t . The physician knows that the man is allergic to the chemical compounds constituting injection 2 but he does not know that the man is allergic to the chemical compounds constituting injection 1; the physician thus gives the man the first injection and the latter dies, at time t , of heart failure from injection 1. Quite clearly, the actual cause of this man’s death is his receiving injection 1. A causal diagram could be constructed involving only the two injections, E_1 and E_2 respectively, and the outcome D with $D = 1$ indicating death at time t and $D = 0$ otherwise, with arrows from E_1 and E_2 pointing to D as in Figure 3.

The structural equations could be written as

$$E_1 = 1$$

$$E_2 = 0$$

$$D = (E_1 \ \& \ E_2) \vee (E_1 \ \& \ \text{not } E_2) \vee (\text{not } E_1 \ \& \ E_2) \vee (\text{not } E_1 \ \& \ \text{not } E_2) = 1$$

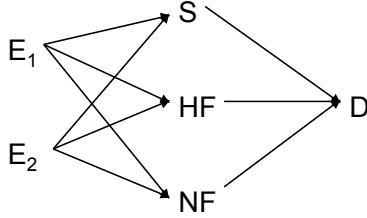


Figure 4: Elaborated causal diagram from Figure 3 showing the possible causes of death.

The definition of an actual cause given by Halpern and Pearl applied to Figure 3, will not identify either of the injections as the actual cause of death nor the combination of these injections nor the absence of them both. This is because the definition of Halpern and Pearl would require that under some set of values for the variables on the graph $D = 0$; however, in the scenario just given D will be 1 regardless of the values of E_1 and E_2 . In this case, the definition of Halpern and Pearl delivers the peculiar conclusion that for this model there is no actual cause of the man’s death.

It is possible to construct an elaborated causal diagram for this example such that the definition of Halpern and Pearl does correctly identify injection 1 as the actual cause of death. Consider for instance the causal diagram given in Figure 4 where S indicates suffocation, HF indicates heart failure and NF indicates failure of the nervous system.

The structural equations corresponding to the elaborated diagram can be written as

$$\begin{aligned}
 E_1 &= 1 \\
 E_2 &= 0 \\
 S &= \text{not } E_1 \ \& \ \text{not } E_2 \\
 HF &= (E_1 \ \& \ \text{not } E_2) \vee (\text{not } E_1 \ \& \ E_2) \\
 NF &= E_1 \ \& \ E_2 \\
 D &= S \vee HF \vee NF.
 \end{aligned}$$

In actual fact, we have $E_1 = 1$, $E_2 = 0$, $S = NF = 0$ and $HF = D = 1$. Now let $X = E_1$ and $x = 1$, $Z = \{E_1, E_2, HF, D\}$, $W = \{S, NF\}$ and $(x', w') = (0, 0, 0)$ we then have (AC1) $E_1 = 1$ and $D = 1$ actually occur, (AC2a) $D = 0$ if E_1

is set to 0 and S and NF are both set to 0 with E_2 at its original value of 0, (AC2b) provided E_1 is set to 1, then $D = 1$ regardless of whether S and NF are set to 0 and regardless of whether E_2 or HF are set to their original values 0 and 1 respectively; also (AC3) holds trivially. Thus the definition of Halpern and Pearl applied to Figure 4 implies that injection 1 is the actual cause of the individual's death. However, the simplest causal diagram, given in Figure 3, and the corresponding structural causal model for this example, demonstrate that the conditions given in the definition of Halpern and Pearl are not *necessary* for an event to be an actual cause.

We note that the critique offered here differs from that offered by Menzies (2004). Menzies considers an earlier proposal of Halpern and Pearl (2001) and points out that their definition does not handle well cases in which token causation is not transitive. Menzies gives an example in which an assassin's poisoning the king's coffee causes a guard to put in an antidote into the coffee; the antidote without the poison would also be lethal but the antidote with the poison neutralizes the poison; the guard's putting the antidote into the coffee thus causes the king's survival. We would not ordinarily say that the assassin's poisoning the coffee caused the king's survival but Menzies shows that the definition given in Halpern and Pearl (2001) would identify the assassin's poisoning the coffee as the actual cause of the king's survival. Halpern and Pearl, in their (2005) paper, give some attention to such examples using their refined definition of actual cause restricting the endogenous variables to a set of "allowable settings" but they do not explicitly discuss Menzies' example.

4. The Claim of the Relativity of Actual Causation

The two examples in the preceding section indicate that the definition of an "actual cause" given by Halpern and Pearl is neither necessary nor sufficient for an event to constitute an actual cause. Pearl and Halpern seem to attempt to get around this issue by proposing that whether an event is an actual cause of another event is relative to the structural model under consideration. They are unconcerned by the fact that their definition, in the structural model represented in Figure 1, identifies Billy's stone as an actual cause of the bottle's shattering because their position is that (p. 845), "according to our definition, the truth of every claim must be evaluated relative to a particular model of the world; that is

our definition allows us to claim only that C causes E in a (particular context in a) particular structural model.” They further note that, “It is possible to construct two closely related structural models such that C causes E in one and C does not cause E in the other.” They consider this “a feature of our model, not a bug.”

Their position is problematic for a number of reasons. First, it seems as though the definition of Halpern and Pearl is missing an important feature of our discourse concerning actual causation. When we are presented with insufficient information to determine what the actual cause is, our judgement is usually one of agnosticism. For example, if, along the lines of Figure 1, we were only told that the bottle shatters if and only if either Sally or Billy throw a rock and in fact they both threw a rock, we would ordinarily conclude that we didn’t have enough information to determine whether Sally’s throwing or Billy’s throwing was the actual cause of the bottle’s shattering. We would respond agnostically; we would not generally conclude that the throwing of both were actual causes and then revise our assessment once further information was available. It would seem that an adequate definition of an actual cause would allow for such agnosticism and would in fact deliver a judgement of agnosticism precisely when we ordinarily do so in every-day reasoning. As we have seen above, Halpern and Pearl’s definition does not provide this judgement of agnosticism, at least not in the case considered.

Second, it is not at all clear that the implications of the relativity of Halpern and Pearl’s definition of actual causation cohere with the sorts of conclusions we ordinarily come to concerning the use of expressions like “ C caused E “. In every-day discussion, we generally achieve near universal agreement concerning which event or events constitute an actual cause; certainly, simple examples such as those presented above would not by themselves lead us to think that the notion of “actual cause” is relative in the way that Halpern and Pearl claim. It might be possible come up with examples in which intuitions about which event is the actual cause are not clear so that it is not possible to achieve consensus i.e. examples in which, if looked at one way, some event seems as though it were the actual cause whereas if looked at another way, a different event seems like the actual cause rather than the first. However, the examples considered above are not of this variety; in these cases, the intuition is clear as to what the actual cause in fact is. It seems that a definition should not have relativistic implications for examples in which the language and reasoning we in fact use about actual causation is clear.

Finally, as a related point, it seems that if our ordinary use of the expressions like “actual cause” and “*C* caused *E*” are not obviously relative, then a case would need to be made that they are in fact so. Again it might be possible to construct examples in which the relativity of the use of the expressions “actual cause” and “*C* caused *E*” is made clear; however, this has not been accomplished in their paper. Halpern and Pearl seem to take the relativity of actual causation as their premise rather than their conclusion.

In the end it is not wholly clear what their definition contributes to reasoning. The approach Halpern and Pearl seem to take in making use of their definition is the following: an event may be considered an actual cause if one can construct a structural model such that their definition identifies the actual cause which agrees with intuition. However, as a characterization of what is meant by “actual cause” it is not clear what their approach contributes beyond relying entirely on intuition from the start.

5. Criteria for the Characterization of Token Causation

An advance beyond Halpern and Pearl would perhaps be made if it were possible to develop a definition of an actual cause that satisfied one of two sets of criteria. On the one hand, it might be possible to devise a definition for an actual cause which, (1a) when applied to any structural model corresponding to the description of the causal relationships, however crude, will always identify some set of events amongst which is included the event or events constituting the actual cause (i.e. for which the definition is a necessary condition for the actual cause) and for which (1b) an elaborated structural model can always be constructed which will identify only the event or events constituting the actual cause. On the other hand, it might be possible to devise a definition for the actual cause which, (2a) when applied to any structural model corresponding to the description of the causal relationships, however crude, would be satisfied only for events which were actual causes (i.e. for which the definition is a sufficient condition for the actual cause) and for which (2b) an elaborated structural model can always be constructed which will in fact identify the event or events constituting the actual cause. The example given above concerning the man suffocating demonstrates that the definition of Halpern and Pearl does not satisfy the first of these two sets of criteria. The example concerning the two rocks and

the bottle's shattering demonstrates the definition of Halpern and Pearl does not satisfy the second of these two sets of criteria.

Definitions which satisfied either the first or the second set of criteria would have the potential to give non-relative answers in accordance with intuition in cases in which such intuition was clear. However, whether they in fact did so would of course have to be examined once a definition were proposed. A definition satisfying the second set of criteria above would also have the potential to deliver judgements of agnosticism, discussed above, in that in simple models in which there is insufficient information to determine the actual cause, it might then not identify any event as the actual cause; however, since (2a) requires the definition is simply a sufficient condition, the conclusion would not then be that there is no actual cause but simply that nothing can be said in this case without further information (e.g. without an elaborated structural model). If definitions satisfying the first or the second set of criteria were proposed, it would remain to be determined whether examples could be constructed in which the event or events constituting the actual cause were relative, both in intuitive reasoning and as assessed by the definition.

A definition which satisfied one of the two sets of the criteria above would constitute a considerable advance in the problem of characterizing what constitutes token causation or an "actual cause." If two definitions could be provided, each satisfying one of the two sets of criteria given above, the concept of token causation could perhaps then be considered adequately characterized. Whether or not this is possible remains to be seen.

REFERENCES

- BENNETT, J. (1993). "Event causation: the counterfactual analysis". In E. Sosa and M. Tooley (eds.), *Causation*. Oxford: Oxford University Press, pp. 217–33.
- COLLINS, J., HALL, N. AND PAUL, L. A. (2004). "Counterfactual and causation: history, problems and prospects". In J. Collins, N. Hall and L. A. Paul (eds.), *Causation and Counterfactuals*, Cambridge, MA: MIT Press, pp. 1–58.
- HALL, N. (2004). "Two concepts of causation". In J. Collins, N. Hall and L. A. Paul (eds.), *Causation and Counterfactuals*, Cambridge, MA: MIT Press, pp. 225–76.

- HALL, N. AND PAUL, L. A. (2003). "Causation and preemption". In P. Clark and K. Hawley (eds.), *Philosophy of Science Today*, Oxford, Oxford University Press, pp. 100–29.
- HALPERN, J. Y. (2000). "Axiomatizing causal reasoning". *Journal of Artificial Intelligence* 12, pp. 317–37.
- HALPERN, J. Y. AND PEARL, J. (2001). "Causes and explanations: a structural model approach. Part I: Causes". Technical Report R-266, Cognitive Science Laboratory, Los Angeles: University of California.
- HALPERN, J. Y. AND PEARL, J. (2005). "Causes and explanations: a structural-model approach. Part I: Causes". *British Journal of the Philosophy of Science* 56, pp. 843–87.
- MENZIES, P. (2004). "Causal models, token causation, and processes". *Philosophy of Science* 71 pp. 820–832.
- PEARL, J. (1995). "Causal diagrams for empirical research". *Biometrika* 82 pp. 669–688.
- PEARL, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge, Cambridge University Press.

L&PS – Logic & Philosophy of Science

Information on the Journal

AIMS AND CONTENTS

L&PS – Logic and Philosophy of Science is an on-line philosophical journal sponsored by the Department of Philosophy of the University of Trieste (Italy). The journal promotes both theoretical and historical research in the philosophy of science and logic, without excluding any particular cultural perspective.

Topics welcomed by the journal include:

- the theory of scientific knowledge and the analysis of the general methodological problems of science (such as scientific discovery, causation, scientific inference, induction and probability, the structure of scientific theories and their relations with empirical data);
- the methodological and foundational problems of the different sciences, from the natural, to the biomedical, to the social sciences;
- the problems related to the historical development of logic, in all its branches, and to the role of logical methods both in the general methodology of science and in the foundations of empirical and mathematical sciences;
- the philosophical problems raised by the development of the cognitive sciences and the philosophy of mind, with particular attention to those results that are relevant for the analysis of scientific practice;
- the epistemological problems related to Artificial Intelligence, robotics, virtual reality, and artificial life;
- the problems in the sociology and the history of science that are relevant to the philosophical investigation of science;
- the problems related to the ethics of science;
- the questions related to the historical and conceptual development of the philosophy of science and logic;
- the problems of the philosophy of language, with particular attention to those results that are relevant for logic and philosophy of science.

The journal will appear twice a year (in March and in September), and one of these issues will be usually devoted to a special topic.

INFORMATION FOR THE AUTHORS

Papers submitted to the journal must be written either in Italian or in English, and must be accompanied by a short summary in English (and also in Italian for the articles written in Italian). All papers will be evaluated by anonymous referees.

In order to promote critical discussion and exchange among scholars, the journal is willing to publish reports on work in progress, to be submitted and evaluated according to the criteria already mentioned above.

The copyright is left to the authors, provided that any reprint of the paper explicitly mentions the version previously published in L&PS.

EDITORIAL BOARD

Gilberto Corbellini (Roma) gilberto.corbellini@uniroma1.it

Mauro Dorato (Roma) dorato@uniroma3.it

Roberto Festa (Trieste) festa@units.it

Marco Giunti (Cagliari) giunti@unica.it

Roberto Giuntini (Cagliari) giuntini@unica.it

Simone Gozzano (L'Aquila) simone.gozzano@cc.univaq.it

Federico Laudisa (Milano) federico.laudisa@unimib.it

Francesco Paoli (Cagliari) paoli@unica.it

Mario Piazza (Chieti) m.piazza@unich.it

Guglielmo Tamburrini (Pisa) gugt@fls.unipi.it

EDITORS IN CHIEF

Mauro Dorato

Roberto Festa

Roberto Giuntini

ASSISTANT EDITORS

Marco Giunti

Francesco Paoli

EDITORIAL ADISORY BOARD

Vito Michele Abrusci, *Roma*; Dario Antiseri, *Roma*; Giovanni Boniolo, *Padova*; Andrea Cantini, *Firenze*; Mirella Capozzi, *Roma*; Martin Carrier, *Bielefeld*; Arturo Carsetti, *Roma*; Ettore Casari, *Pisa*; Carlo Cellucci, *Roma*; Roberto Cordeschi, *Salerno*; Giorgio De Rossi, *Trieste*; Giuliano Di Bernardo, *Trento*; Rosaria Egidi, *Roma*; Maurizio Ferriani, *Bologna*; Maria Carla Galavotti, *Bologna*; Sergio Galvan, *Milano*; Pierdaniele Giaretta, *Padova*; Gurol Irzik, *Istanbul*; Theo A.F. Kuipers, *Groningen*; Diego Marconi, *Vercelli*; Enrico Moriconi, *Pisa*; Ilkka Niiniluoto, *Helsinki*; Francesco Orilia, *Macerata*; Paolo Parrini, *Firenze*; Angelo Maria Petroni, *Bologna*; Huw Price, *Sydney*; Giorgio Sandri, *Bologna*; Marina Sbisà, *Trieste*; Silvano Tagliagambe, *Sassari*; Nicola Vassallo, *Genova*; Achille C. Varzi, *New York*; Alberto Voltolini, *Vercelli*; Geeron Wolters, *Konstanz*; Giancarlo Zanier, *Trieste*.

TYPESETTING, GRAPHICAL ADVICE AND TECHNICAL ASSISTANCE

Gustavo Cevolani (g.cevolani@gmail.com)

Luca Tambolo (l_tambolo@hotmail.com)

WEBMASTER

Gustavo Cevolani (g.cevolani@gmail.com)