

Francesco Beccuti

La disgiunzione di Gödel

La disgiunzione di Gödel è la tesi filosofica secondo cui o l'abilità matematica umana non è meccanizzabile oppure esistono problemi matematici assolutamente insolubili. Dopo aver brevemente delineato il contesto storico e teoretico della disgiunzione, ne sottolineiamo la rilevanza e per la filosofia della matematica e per la filosofia della mente. Procediamo poi a presentarne una possibile dimostrazione sottolineando gli assunti filosofici e i problemi d'idealizzazione che ne stanno alla base. Presentiamo inoltre in dettaglio i due corni della disgiunzione, soffermandoci su alcuni tentativi di stabilirne la validità e sulle critiche relative.

INDICE

1. INTRODUZIONE
 2. LA DISGIUNZIONE DI GÖDEL E LA SUA RILEVANZA FILOSOFICA
 3. DIMOSTRAZIONE E PROBLEMI D'IDEALIZZAZIONE
 4. IL PRIMO DISGIUNTO
 5. IL SECONDO DISGIUNTO
- BIBLIOGRAFIA

1. Introduzione

Nel 1929 Gödel prova che ogni teoria del prim'ordine coerente ha un modello, stabilendo una fondamentale connessione fra la teoria dei modelli e la teoria della dimostrazione. Questo risultato, il teorema di completezza, viene in quel momento percepito come il primo passo verso una conferma della particolare filosofia della matematica della scuola di Hilbert: tutti i problemi matematici possono essere risolti per mezzo di dimostrazioni sintattiche. Tuttavia, i successivi teoremi d'incompletezza, dimostrati da Gödel stesso nel 1931, danno per contro conferma negativa delle tesi di Hilbert. Gödel dimostra infatti che ogni teoria coerente T capace di dimostrare gli assiomi dell'aritmetica di Peano è incompleta: esiste un enunciato vero ϕ , nel linguaggio dell'aritmetica, tale che $T \not\vdash \phi$ e $T \not\vdash \neg\phi$ (ossia T non dimostra né ϕ né la sua negazione).

In particolare, la dimostrazione del secondo teorema d'incompletezza si ottiene mediante una tecnica detta di "aritmetizzazione" della sintassi di T , che procede attraverso l'assegnazione di numeri ai simboli, alle formule e alle sequenze di formule di T . Dato T , allora, si può costruire in modo canonico una relazione $Prov_T(x, y)$, che esprime intuitivamente il fatto che x è il numero di Gödel di una dimostrazione in T della formula con numero di Gödel y . Ora, se si fissa un numero n che rappresenta nel linguaggio considerato un qualsiasi enunciato falso (per esempio $0 = 1$), allora per quanto detto la formula

$$\forall x \neg Prov_T(x, n)$$

esprime il fatto che non c'è alcun x che è numero di Gödel di una dimostrazione in T di $0 = 1$, ossia tale formula esprime la coerenza di T .

Nel contesto della discussione filosofica sulla disgiunzione di Gödel è importante notare che, indipendentemente dalla nostra fiducia nella verità degli assiomi, il problema se da essi derivi tramite una sequenza finita una contraddizione è un problema a proposito di un fatto del tutto oggettivo: la sequenza di formule che porta dagli assiomi della teoria ad una contraddizione è una sequenza finita e derivata tramite regole sintattiche precise. Dunque essa, oggettivamente, esiste oppure non esiste. Questo fatto è messo bene in luce dalla possibilità di trasformare l'enunciato che esprime la coerenza del sistema di cui sopra in un problema aritmetico ben definito. Segue infatti da alcuni risultati dello stesso Gödel e dal lavoro congiunto di Putnam, Robinson e Matiyasevich sul decimo problema di Hilbert che la coerenza di un sistema di assiomi T è sempre (dimostrabilmente) equivalente a un problema nella forma

$$\forall x_1 \dots x_n P(x_1, \dots, x_n) \neq 0,$$

per P un polinomio diofanteo, ossia un polinomio a coefficienti e variabili intere (Feferman 2006, 139).

Nel 1936 Turing sviluppa una nozione di macchina automatica (ora detta macchina di Turing) per descrivere matematicamente il concetto intuitivo di procedura effettiva. Una macchina di Turing è un modello matematico di una macchina molto semplice che opera (computa) meccanicamente su un nastro, su cui può leggere, cancellare o scrivere simboli. La sequenza di operazioni di una macchina di Turing è determinata da un numero finito di istruzioni elementari, da un numero finito di stati della macchina e da una sequenza finita di simboli, detta "input" della computazione, che si trova inizialmente sul nastro. Tale nastro è invece potenzialmente infinito, come potenzialmente infinita è la durata della computazione stessa. Se essa non è infinita, allora la sequenza di simboli che si trova sul nastro a computazione terminata è detta "output" della computazione.

La macchina di Turing è comunque solo uno dei particolari tentativi di definire rigorosamente il concetto di "procedura meccanica" o "procedura effettiva" o semplicemente di "algoritmo"¹. Molti altri tentativi sono stati avanzati in letteratura prima e

¹ Tuttavia le nozioni di "procedura meccanica" e di "procedura algoritmica" non sono esattamente coincidenti, essendo la prima più generale della seconda (vedi Gandy 1988).

dopo il fondamentale intervento di Turing (per esempio il λ -calcolo di Church o la teoria delle funzioni ricorsive di Kleene). Tuttavia, tutti i particolari modelli di computazione proposti fino ad oggi sono stati provati essere meno o al più ugualmente espressivi rispetto al modello di Turing. Questo ha portato alla formulazione della seguente tesi:

Tesi di Church-Turing: la nozione di macchina di Turing cattura completamente il concetto di procedura effettiva.

Si noti che è possibile determinare attraverso una procedura effettiva se un dato numero è il numero di Gödel di una formula di una teoria T . Dato che, poi, le dimostrazioni in T non sono altro che sequenze finite di formule, è ancora possibile determinare effettivamente se un dato numero è numero di Gödel di una dimostrazione in T e, quando lo è, possiamo effettivamente determinare il numero dell'ultima formula della sequenza, ossia del teorema stabilito dalla dimostrazione. Ne segue allora che, data una teoria T , si può costruire una procedura meccanica che controlli, per ogni numero naturale, se esso è il numero di una dimostrazione di T e che, quando lo è, produca il numero del teorema stabilito dalla dimostrazione. Si può dunque associare ad ogni teoria T una macchina di Turing M che enumeri tutti e soli i teoremi di T . Viceversa, data una macchina di Turing M , e specificato un opportuno linguaggio formale, si può far corrispondere a M un'equivalente teoria T nel linguaggio specificato, semplicemente scegliendo una opportuna numerazione di Gödel tale per cui tutti i numeri che M enumera siano numeri di Gödel di formule del linguaggio² e infine definendo i teoremi di T come la chiusura deduttiva delle formule via via enumerate da M secondo tale numerazione.³ Abbiamo allora il seguente (meta-)teorema che enuncia un fatto stabilito matematicamente: l'identità fra macchine di Turing e teorie logiche.

Isomorfismo teorie-macchine: Data una qualsiasi teoria logica T , esiste una macchina di Turing i cui input e regole di funzionamento corrispondono agli assiomi e alle regole di T e il cui output consiste di tutti e soli i teoremi di T . Viceversa, data una qualsiasi macchina di Turing M , esiste una teoria i cui assiomi e regole logiche corrispondono agli input e alle regole della macchina e i cui teoremi sono tutti e soli gli output di M .

Ne segue che parlare di macchine di Turing e parlare di teorie logiche è la stessa cosa: tutto ciò che può esser fatto attraverso una macchina automatica può esser fatto tramite una corrispondente teoria logica e viceversa. In questo senso il problema dell'intelligenza artificiale è un problema della logica. L'isomorfismo fra sistemi formali e macchine di Turing ci permette anche, nel contesto della discussione sulla disgiunzione di Gödel e sul meccanicismo, di non tenere in nessun conto la definizione formale di macchina di Turing e di parlare liberamente di macchina, intendendo un qualsiasi dispositivo capace di manipolare i dati simboli (quelli dell'aritmetica) e di dimostrare sintatticamente teoremi veri (dell'aritmetica). Ci si può permettere inoltre di parlare liberamente di assiomi o di coerenza di una macchina, intendendo sempre gli assiomi e la coerenza della teoria corrispondente alla macchina considerata in base all'isomorfismo di cui sopra. Allo stesso modo si può parlare di applicare il teorema d'incompletezza alle macchine. Data una

² Si può in effetti definire un'assegnazione di numeri di Gödel che sia, a differenza di quella originariamente introdotta da Gödel, non solo iniettiva, ma anche suriettiva (vedi Soare 2016, xxxii).

³ Vedi Feferman (2006, 137-138)

macchina coerente, esiste un enunciato (quello che esprime la coerenza della macchina stessa), che tale macchina non può provare.

2. La disgiunzione di Gödel e la sua rilevanza filosofica

Nel 1951 Gödel presenta per la prima volta esplicitamente in una conferenza alla American Mathematical Society le sorprendenti conseguenze filosofiche dei suoi celebrati teoremi d'incompletezza, che

[rendono] impossibile il fatto che qualcuno, considerando un sistema d'assiomi e di regole logiche determinato e ben-definito, possa a proposito di esso coerentemente pronunciare la seguente asserzione: Io percepisco (con certezza matematica) che tutti questi assiomi e queste regole sono corretti, e inoltre, credo che contengano tutta la matematica. Se qualcuno pronunciasse tale asserzione si contraddirebbe. Perché se egli percepisse gli assiomi considerati come corretti, allora percepirebbe anche (con la stessa certezza) la loro coerenza. Dunque egli avrebbe un'intuizione matematica non derivabile da tali assiomi. (Gödel 1951, 309)

E allora questo significa che non è possibile comprendere tutta la matematica in un solo sistema di assiomi? Molto dipende da che cosa intendiamo per "matematica". Nel corso della conferenza, infatti, Gödel introduce una suggestiva distinzione fra la matematica nel suo senso *soggettivo* (l'insieme delle proposizioni dimostrabili a partire da un qualche sistema di assiomi) e la matematica nel suo senso *oggettivo* (l'insieme delle proposizioni vere in senso assoluto). Tali insiemi coincidono, si chiede Gödel, oppure no? Se sì, allora non c'è speranza, per quanto detto, di poter comprendere tutta la matematica in un unico sistema assiomatico, perché se tale sistema esiste, allora l'enunciato che esprime la coerenza del sistema non può, per il secondo teorema d'incompletezza, esser dimostrabile in tale sistema, e ciò contraddirebbe l'assunto iniziale. Se invece la matematica oggettiva è distinta dalla matematica soggettiva, allora da un lato la matematica soggettiva potrebbe esser passibile d'esser compresa in unico sistema assiomatico, ma, d'altra parte, rimarrebbe il problema di spiegare filosoficamente (e matematicamente) l'esistenza di proposizioni matematiche vere, ma non accessibili mediante dimostrazione in un sistema formale. Gödel giunge dunque a sostenere la seguente tesi disgiuntiva: o la matematica soggettiva sorpassa qualsiasi tentativo di formalizzazione, oppure la matematica oggettiva sorpassa la matematica soggettiva. In altre parole,

[...] la seguente conclusione disgiuntiva è inevitabile: O la matematica è inesauribile nel senso che i suoi assiomi evidenti non potranno mai essere compresi per mezzo di una regola finita, cioè la mente umana (anche solo nel campo della matematica pura) sorpassa infinitamente le capacità di ogni macchina finita, oppure esistono problemi diofantei assolutamente irrisolvibili [...] (dove il caso in cui ambedue i termini della disgiunzione sono veri non è escluso [...]) Ed è questo fatto, stabilito matematicamente, che mi sembra essere di grande interesse filosofico. (Gödel 1951, 310).

La tesi di Gödel è dunque la seguente: o il meccanicismo è falso oppure esistono problemi matematici a cui non si può sperare di trovare soluzione. Questa è, scriverà poi Gödel, «forse la prima proposizione rigorosamente provata a proposito di un concetto filosofico» (Gödel in Cassou-Noguès 2008, 2). Secondo Gödel, infatti, ambedue le alternative si pongono in netta opposizione rispetto al materialismo in filosofia (della mente nel caso del primo corno della disgiunzione, della matematica nel caso del secondo corno):

Se vale la prima alternativa, questo sembra implicare che il funzionamento della mente umana non può essere ridotto al funzionamento del cervello che, sotto ogni apparenza, è una macchina finita con un numero finito di parti, ossia i neuroni e le loro connessioni (Gödel 1951, 311)

La seconda alternativa, invece,

Sembra confutare il fatto che la matematica sia solo una nostra creazione. Questa alternativa sembra implicare che gli oggetti e i fatti matematici (o almeno qualcosa in essi) esistono oggettivamente e indipendentemente dai nostri atti e dalle nostre decisioni mentali, ossia che deve valere qualche forma di platonismo o di “realismo” in riferimento agli oggetti matematici. (Gödel 1951, 311)

In altre parole, o il meccanicismo è falso oppure il Platonismo è vero, o entrambe le cose (Wang 1996, 186).

La rilevanza dei teoremi d'incompletezza rispetto a una teoria scientifica della mente umana, pur discussa da molti autori di formazione logica, rimane tuttavia ancora largamente ignorata dalla gran parte del dibattito filosofico e psicologico contemporaneo, che sembra non averne ancora recepito l'importanza. Come ha osservato Benacerraf, infatti,

[...] se esiste una macchina di Turing che può provare tutto ciò che la mia chiusura del prim'ordine può provare, allora io non posso provare che [...] tale macchina è ad un tempo adatta a fare aritmetica e che può provare tutto ciò che io posso provare. [...] Una persona a cui ho spiegato [questo risultato] ne ha concluso che la psicologia come la conosciamo è impossibile. Perché se noi non siamo al più macchine di Turing, allora essa è impossibile, ma se siamo macchine di Turing, allora ci sono determinati fatti a proposito di noi stessi che non possiamo conoscere [...] Non prenderò posizione in proposito. Tuttavia possiamo riformulare come segue la portata [filosofica] dei teoremi di Gödel: se io sono una macchina di Turing, allora la mia stessa natura mi impedisce di obbedire all'imperativo socratico: CONOSCI TE STESSO. (Benacerraf 1967, 30)

3. Dimostrazione e problemi di idealizzazione.

Un abbozzo di dimostrazione del disgiunto gödeliano potrebbe essere il seguente.⁴ Si assuma che la mente umana possa essere simulata algebricamente. Per la tesi di Church-Turing, esiste allora una macchina di Turing in grado di simulare la mente umana nella sua capacità di produrre teoremi veri dell'aritmetica. Ne segue che la mente può essere simulata tramite un sistema formale di assiomi. E allora, se la mente è coerente, il secondo teorema d'incompletezza si applica a tale sistema di assiomi. Quindi esiste un enunciato vero (quello che esprime la coerenza della mente) che il sistema stesso (la mente) non può provare. Per l'assunto iniziale, allora, tale enunciato è umanamente (ossia assolutamente) indimostrabile.

Questo argomento, evidentemente, non è una dimostrazione matematica *strictu sensu*. In effetti, essa comprende concetti informali come "mente", "macchina", o "problema assolutamente insolubile", mentre le dimostrazioni matematiche coinvolgono solamente concetti precisamente definiti. La letteratura sulla disgiunzione è in ogni caso concorde sul fatto che Gödel abbia stabilito la sua tesi al di là di ogni dubbio ragionevole e comunque «non meno conclusivamente di qualsiasi altro argomento filosofico» (vedi Horsten e Welch 2016, 2). Persistono tuttavia, in merito alla formulazione della disgiunzione, alcune questioni irrisolte di natura epistemologica, linguistica e d'idealizzazione che la letteratura non cessa di mettere in luce, di cui la più stringente è di certo che il contenuto della tesi meccanicistica non è completamente chiaro (come non è chiaro il concetto di dimostrabilità assoluta). Che cosa s'intende infatti per "la mente umana nel campo della matematica pura", che cosa significa che essa è meccanizzabile, e in che senso?

In primo luogo, la questione della meccanizzabilità della mente (il suo esser "compresa in una regola finita") è forse quella che meno si presta a dubbie interpretazioni, dato che da tempo disponiamo di definizioni precise di "procedura effettiva" o di "algoritmo", ossia di processi che possono essere portati a termine in un numero finito di passaggi da un agente idealizzato a partire da un numero finito di istruzioni. Come detto nell'introduzione, tutti i tentativi storici di definire una procedura effettiva risultano essere equivalenti. Quindi, tenendo conto della tesi di Church-Turing, si può esprimere senza perdere di generalità la tesi meccanicistica nel contesto della disgiunzione di Gödel: la mente umana nel campo della matematica pura è una macchina di Turing. Nel seguito quando parleremo genericamente di "macchine" intenderemo sempre "macchine di Turing", a meno di non specificare altrimenti. Dunque la meccanizzabilità della mente, se è possibile, deve essere possibile tramite una macchina di Turing.

Rimane al contrario ancora del tutto oscuro che cosa s'intenda per "mente umana" anche solo nel contesto "ristretto" della matematica pura. Significa la mente di un singolo matematico oppure si sta parlando più in generale della comunità degli studiosi di

⁴ Una dimostrazione alternativa della disgiunzione che fa uso dei concetti di matematica oggettiva e soggettiva introdotti in precedenza, oltreché del teorema d'inddefinibilità della verità di Tarski, è la seguente. Sia S l'insieme delle verità matematiche riconoscibili come tali dalla mente umana (matematica soggettiva), e sia V l'insieme delle verità matematiche (matematica oggettiva). Per definizione (o per ipotesi) si ha che $S \subseteq V$. Inoltre, per il teorema d'inddefinibilità della verità di Tarski, V non è definibile nel linguaggio dell'aritmetica e dunque *a fortiori* V non è ricorsivamente enumerabile. Ne segue che, se $S = V$ allora neanche S è ricorsivamente enumerabile, e dunque il meccanicismo è falso. Segue allora che, se il meccanicismo è vero $S \neq V$, e in tal caso dunque esiste una proposizione $\phi \in V$ tale che $\phi \notin S$, cioè ϕ è vera, ma non umanamente dimostrabile (Shapiro 2016, 190-191). Vedi anche l'argomento modale delineato in Bruni (2006, 41-43).

matematica? La maggior parte dei commentatori odierni tende in questo contesto a identificare la mente umana con la più larga categoria della comunità degli studiosi, ponendo l'enfasi sul fatto che la matematica è attività cooperativa e sociale.⁵ Viceversa Gödel, nelle sue conversazioni con Wang, tende invece a porre l'accento sul carattere solitario della matematica e definisce la mente umana nel contesto della sua disgiunzione come il potere di provare teoremi da parte di un singolo essere umano scevro da limitazioni di spazio e di tempo.⁶

Si noti infatti che, banalmente, il numero di teoremi che un singolo essere umano può provare nel corso della sua vita è un numero finito. Allo stesso modo la quantità di teoremi che l'umanità nel suo complesso può provare prima che il sole si spenga (nella suggestiva metafora di Shapiro 2016, 192) è anch'essa finita. Dato che l'insieme delle conseguenze logiche di un insieme finito è ricorsivamente enumerabile,⁷ allora è chiaro che in questo senso la matematica è banalmente meccanizzabile. È evidente allora che non è questo il senso in cui s'intende (almeno da parte degli anti-meccanicisti) la questione della meccanizzabilità della matematica. Non si sta infatti qui trattando della matematica nel senso dei teoremi che sono e saranno dimostrati dagli esseri umani, bensì della loro capacità *potenziale* di dimostrare teoremi. Inoltre, lo stesso discorso si potrebbe fare per la dimostrabilità assoluta, anch'essa intesa come capacità potenziale di provare teoremi. In ambedue i casi, allora, è necessario fissare un agente idealizzato in grado di dimostrare teoremi e di dimostrarli correttamente. Solo a questo punto il teorema d'incompletezza può essere applicato: dato l'insieme dei teoremi che è possibile provare, chiuso per conseguenze logiche, esiste un teorema vero (quello che esprime la coerenza del sistema iniziale) che è fuori da tale insieme. Entrano dunque in gioco difficili nozioni temporali, modali ed epistemologiche: si tratta di meccanizzare la totalità dei teoremi che un essere umano (o la comunità matematica umana) potrebbe produrre. Horsten (2005, 18), nel contesto della dimostrabilità assoluta, rifacendosi ad Anderson (1993), prova a discutere esplicitamente tale riferimento alla modalità, concludendo che, nel caso di un agente idealizzato in grado di provare teoremi, la dimensione modale tende a eclissare la dimensione temporale (se le teniamo entrambe in conto). Inoltre, data la debolezza della nostra intuizione della dimensione controfattuale in riferimento alla dimostrabilità, Horsten si contenta di tener in conto solo un'infinita dimensione temporale: l'agente idealizzato in grado di provare teoremi è supposto avere tempo infinito e spazio infinito con dimostrazioni strutturate come alberi finiti con gli assiomi come foglie dell'albero e il teorema come radice dell'albero avente i nodi più bassi ottenuti da quelli più alti tramite le regole d'inferenza.⁸ Horsten e Leitgeb scelgono allora di condurre lo studio della dimostrabilità da parte di un agente idealizzato (ancora non definito in maniera sufficientemente chiara) nel contesto dell'aritmetica epistemica: un'estensione dell'aritmetica di Peano arricchita di un operatore primitivo denotante la nozione di provabilità intuitiva che si comporta come un operatore modale ed è conseguentemente assiomaticizzato, secondo la linea di ricerca inaugurata da Myhill (1960) e ripresa indipendentemente da Shapiro (1985).

⁵ Ad esempio Leitgeb (2009, 282).

⁶ Ossia immortale e con una provvista infinita di carta e inchiostro.

⁷ Cioè esiste una macchina di Turing in grado di enumerare uno per uno tali teoremi.

⁸ Vedi anche Leitgeb (2009, 265)

In conclusione, comunque, la maggior parte degli autori si contenta di adottare lo spirito suggerito, pur con molte riserve,⁹ da Feferman (2006, 144) e da Shapiro (2016, 203), e approssicare le questioni filosofiche del meccanicismo (e dell'esistenza di problemi assolutamente insolubili) per così dire alla lettera, assumendo che le idealizzazioni siano fornite di senso.¹⁰ Con questo spirito, già a partire dalla dimostrazione sopra accennata possiamo identificare le principali assunzioni (matematiche e filosofiche) sottese alla dimostrazione della disgiunzione di Gödel:

- 1) Il secondo teorema d'incompletezza
- 2) L'isomorfismo teorie-macchine
- 3) La tesi di Church-Turing
- 4) La coerenza della mente umana

Il carattere (meta)matematico del primo e del secondo assunto, fa sì che essi siano al di là di ogni dubbio, mentre la tesi di Church-Turing (nonostante le sue giustificazioni siano di natura filosofica o empirica) è universalmente accettata dalla comunità degli esperti. Ne segue che colui il quale, accettando le succitate idealizzazioni riguardo la mente umana e la sua coerenza, volesse comunque mettere seriamente in dubbio la verità della disgiunzione dovrebbe cercare di mostrare che tale mente è incoerente o, al più, para-coerente. Hofstadter (1975), ad esempio, sostiene con forza la tesi dell'incoerenza sostanziale della mente umana, mentre Lucas (1970) si è di molto speso a difesa della possibilità di argomentare tale coerenza.

4. Il primo disgiunto

In effetti, è proprio Lucas (1961, 112) a proclamare che «il meccanicismo è falso», fornendo un famoso argomento per il primo disgiunto basato sul teorema d'incompletezza.¹¹ L'argomento di Lucas viene poi riformulato a più riprese da Penrose (1989 e 1994) che lo include in un ragionamento più generale volto a dischiudere la natura della coscienza umana anche in considerazione dei più recenti sviluppi della meccanica quantistica e della biologia neuronale. Dirighiamo il lettore a Lindström (2001) e a Chalmers (1995) per le formulazioni più compatta (e le critiche più cogenti) agli argomenti di Penrose. Ci concentreremo invece qui sulla versione di Lucas dell'argomento contro il meccanicismo.

⁹ La posizione di Shapiro (2016, 205-206), condivisa da Feferman, è che non esiste ad oggi una tesi meccanicistica sufficientemente ben definita.

¹⁰ Vedi anche Fano e Graziani (2011) per una discussione delle premesse filosofiche della disgiunzione e dei relativi problemi d'idealizzazione. L'articolo, inoltre, contiene un nuovo argomento per la disgiunzione che migliora un precedente argomento di Chihara (1972).

¹¹ Già Turing (1950) considera «l'obiezione matematica» al meccanicismo, ossia il fatto che il secondo teorema d'incompletezza possa essere usato per l'escludere la possibilità dell'intelligenza meccanica. La risposta di Turing all'obiezione matematica è che, sebbene sia vero che ci sono limitazioni ai poteri delle macchine, non è stato però dimostrato che l'intelletto umano non sia esso stesso soggetto alle medesime limitazioni. Tuttavia, Turing si rende conto che tale sommaria risposta non è del tutto convincente e dunque propone ai sostenitori dell'obiezione matematica di procedere nel dibattito prendendo in considerazione il «gioco dell'imitazione» (descritto nel medesimo articolo e passato alla storia con il nome di «test di Turing») come base della discussione sulla possibilità di meccanizzare l'intelligenza (Turing 1950, 445).

Si assuma per assurdo che la mente umana sia (o sia passibile di essere simulata da)¹² una macchina di Turing coerente. Allora esisterebbe un sistema di assiomi che descrive il funzionamento della mente umana. La coerenza di tale sistema di assiomi è una proposizione matematica che tale sistema (tale macchina) non può provare (né asserire). Al contrario, dato che la mente può asserire (e provare, secondo Lucas) la coerenza della macchina che dovrebbe rappresentarla, segue l'assurdo.

Oltre alle ipotesi di cui sopra, tale argomento, si può far vedere, poggia su due assunzioni filosofiche ulteriori:

- 5) la conoscibilità della coerenza della mente umana.
- 6) la conoscibilità della macchina che dovrebbe rappresentare la mente umana.

Per una rassegna delle più famose critiche dell'argomento di Lucas si consulti Labinaz (2016) oltreché Fano e Graziani (2013). Le obiezioni più pertinenti all'argomento di Lucas sono basate sull'invalidità di uno o più degli assunti 4) 5) e 6). In effetti è possibile che la mente umana sia una macchina, ma non sia una macchina coerente. Oppure è possibile che la mente sia una macchina coerente ma che non sappia di essere tale. Infine è possibile che la mente umana sia una macchina coerente, che sa di esserlo, ma che non è in grado di stabilire con precisione la sua natura di macchina (ossia, intuitivamente, che sappia di essere *una* macchina, ma che non sappia *quale* macchina essa stessa sia).

In merito alla questione della coerenza della mente, come detto, Lucas sostiene che siamo coerenti e che di tale coerenza possiamo dare una dimostrazione informale basata sul fatto che non siamo disposti ad accettare qualsiasi enunciato come vero¹³ e sull'impossibilità di sostenere una posizione di completo agnosticismo rispetto alla questione della verità matematica:¹⁴

[la coerenza della mente deve] essere assunta come premessa necessaria del pensiero [...] Non è possibile [di ciò] alcuna dimostrazione formale: ma assumere il contrario è auto-contraddittorio [...] Non potrei convincermi di non essere coerente senza allo stesso tempo abbandonare la mia convinzione dell'esistenza di una differenza fra il vero e il falso [...] Dunque, anche se la possibilità che io sia essenzialmente incoerente non può essere confutata con argomenti puramente formali [...] questa è una posizione di assoluto nichilismo concettuale. (Lucas 1970, 162-163)

Tale argomento, secondo Lucas, costituisce il modo per confermare le ipotesi 4) e 5): siamo coerenti e possiamo sapere di esserlo. Se tuttavia non accettiamo la coerenza della mente

¹² Si noti che v'è una distinzione importante fra le due tesi meccanicistiche: a) la mente umana è (o è simulabile attraverso) una macchina e b) è possibile costruire tale macchina. Si può dire che la letteratura, da Lucas in poi, si è concentrata piuttosto sulla prima tesi che sulla seconda.

¹³ Mentre invece, come detto nell'introduzione, è proprietà di una teoria incoerente la capacità di provare qualsiasi enunciato (vero o falso). Questo argomento per la coerenza della mente è però giudicato debole in Wang (1974, 319).

¹⁴ Bruni (2005, 222-223) evidenzia infatti la dipendenza dell'argomento di Lucas da una certa concezione (filosoficamente opinabile, secondo Bruni) del concetto di verità in matematica.

oppure non accettiamo la possibilità di servirci di argomenti informali come quello di Lucas per provare la coerenza della mente, allora non possiamo accettare neppure l'argomento di Lucas. Infatti se 4) è falsa, allora il meccanicismo potrebbe essere vero perché la mente potrebbe essere una macchina incoerente. In tal caso, allora, non esiste un enunciato gödeliano per tale macchina perché il teorema d'incompletezza non vi si applica punto (e anzi tale macchina è in grado di provare tutte le proposizioni, anche quelle false!). Se invece 4) fosse vera, ma 5) falsa allora non vi sarebbe modo per la mente di sapere di essere coerente: la mente allora potrebbe essere una macchina coerente che non sa di esser tale, come ogni altra macchina coerente a cui, per definizione, si può applicare il teorema d'incompletezza.

Come già accennato, inoltre, l'argomento di Lucas non esclude il fatto che la mente possa essere una macchina, ma che non sia umanamente possibile sapere *quale* macchina essa sia. Tale obiezione, originariamente posta da Benacerraf (1967), e basata sul rifiuto dell'ipotesi 6), costituisce forse il colpo più forte inferto agli argomenti anti-meccanicistici fondati sui teoremi d'incompletezza. Un interessante filone formale di ricerca, purtroppo ancora poco battuto, si è infatti sviluppato intorno all'osservazione di Benacerraf. In particolare, gli studi di Reinhardt (1986) e di Carlson (2000) nel contesto della succitata aritmetica epistemica, pongono in luce l'esistenza di macchine di Turing coerenti che, intuitivamente, sanno di essere una qualche macchina di Turing coerente ma non sanno con esattezza *quale* macchina di Turing esse stesse sono. Questi studi, si noti bene, escludono la possibilità che argomenti basati sui teoremi d'incompletezza possano essere usati per negare il meccanicismo, a meno di non postulare da parte della mente umana una totale e (forse inverosimile) auto-conoscenza: l'ipotesi 6), appunto.¹⁵

Esiste comunque un dibattito contro-argomento anche per quel meccanicista che decidesse di accettare tutte le ipotesi sottostanti all'argomento di Lucas. Il meccanicista in tal caso, potrebbe ancora controbattere a Lucas che, sebbene ogni data macchina M non può asserire l'enunciato ϕ che esprime la sua coerenza, esiste sempre un'altra macchina M_1 capace di asserire ϕ : è sufficiente "aggiungere" ϕ agli assiomi di M .¹⁶ Tuttavia in tal caso, per il secondo teorema d'incompletezza, esisterebbe un nuovo enunciato ϕ_1 che esprime la coerenza di M_1 e che M_1 non sarebbe comunque in grado di asserire. Ma allora il meccanicista potrebbe di nuovo allo stesso modo concepire un'altra macchina M_2 capace di asserire ϕ_1 . In tal caso, però, sempre per il secondo teorema d'incompletezza, esisterebbe comunque un nuovo enunciato ϕ_2 che esprime la coerenza di M_2 e che M_2 non può asserire, e così via all'infinito. Nel caso in cui accettiamo tutte le ipotesi di Lucas, allora si configura allora quella che Benacerraf ha chiamato una vera e propria «partita di tennis» fra Lucas e un ipotetico meccanicista che accetti le sue ipotesi: data una qualsiasi macchina, Lucas può "gödelizzare" e quindi pare poter provare più verità aritmetiche di ogni

¹⁵ Vedi anche, su questa direzione di ricerca, Alexander (2014) il quale mostra che, in generale, una macchina corretta (cioè una macchina in grado di provare solo teoremi veri, condizione necessaria per la coerenza) può conoscere esattamente il suo proprio programma, ma non può sapere allo stesso tempo di essere corretta. Aldini, Fano e Graziani (2016), mostrano invece che, almeno per particolari input iniziali, ci sono macchine corrette in grado di conoscere il proprio programma e contemporaneamente la propria correttezza.

¹⁶ Come è stato notato da Putnam e Benacerraf, infatti, Lucas non può provare formalmente la coerenza di una qualsiasi macchina, ma solo fornire per essa un argomento informale. Tale argomento, dal punto di vista formale, si riduce all'assunzione dell'enunciato che esprime la coerenza. Ma anche una macchina può assumere un qualsiasi enunciato, semplicemente aggiungendolo ai propri assiomi.

macchina data; al contrario però esiste sempre una macchina più potente capace di superare le limitazioni della precedente.¹⁷ Neanche un “salto nel transfinito” può aiutare a risolvere il gioco in favore di un contendente o dell’altro. In effetti si può costruire, tramite opportuni artifici di notazione ordinale, una macchina definita come

$$M_\omega := \bigcup_{n < \omega} M_n,$$

ossia una macchina che ingloba tutte le precedenti ed è dunque capace di asserire tutte le formule di Gödel $\phi_1, \phi_2, \phi_3, \dots$. Tuttavia tale macchina è ancora incompleta: esiste un’altra formula ϕ_ω che tale macchina non può provare, e dunque la partita di tennis fra Lucas e il meccanicista può continuare anche oltre il primo ordinale infinito.¹⁸ Hofstadter (1979, 475) in merito sostiene, sulla base di un teorema dell’analisi ordinale di Church e Kleene,¹⁹ che, a un certo punto nella catena degli ordinali, Lucas sicuramente avrà la peggio sul meccanicista.²⁰ Lucas (1996, 111), invece, argomenta a partire dal medesimo teorema per mostrare che, al contrario, il meccanicista avrà la peggio. Più di recente, e con tutta la sua autorevolezza, Shapiro (2016, 200), criticando sia Lucas sia Hofstadter, sostiene che le tecniche dell’analisi ordinale non sono di alcun aiuto per stabilire la verità o la falsità del meccanicismo.

Si noti infine che Lucas – sostanzialmente solo nel tentativo di confutare il meccanicismo con argomenti puramente matematici o filosofici²¹ – gradualmente recepisce le critiche mossegli negli anni e alleggerisce la sua posizione originaria. Lucas (1996, 105) non sostiene più di poter confutare il meccanicismo *tout court*, ma reputa di poter fornire un argomento capace di confutare ogni possibile istanza particolare della tesi meccanicistica, ossia ogni possibile macchina che un ipotetico meccanicista²² indichi come il modello della mente umana. L’argomento di Lucas, allora, nell’accettazione delle sue ipotesi, costituirebbe non una dimostrazione dell’impossibilità di meccanizzare la mente umana, ma uno schema di dimostrazione per ogni possibilità effettiva di meccanizzare il ragionamento umano: ogniquale volta un meccanicista sosterrà l’ipotesi che una certa macchina coerente è quella che dovrebbe costituire il modello meccanico della mente umana, allora il secondo teorema d’incompletezza si potrà applicare a tale macchina, così confutando l’ipotesi iniziale del meccanicista

¹⁷ Vedi Turing (1950, 445): «Possono esistere uomini più intelligenti di ogni data macchina, e tuttavia possono esistere altre macchine ancora più intelligenti, e così via»

¹⁸ Dal livello $\omega + 1$ in poi, tuttavia, entrano in gioco difficili problemi di notazione ordinale e d’intensionalità delle teorie aritmetiche considerate. Feferman (1962) costituisce la guida indiscussa a tali problemi, a cui Franzen (2004) si può considerare un’introduzione.

¹⁹ Vedi Rogers (1967, 206-207).

²⁰ Vedi anche l’esposizione informale in Berto (2008, 214-218).

²¹ Vedi però Galvan (2004), che critica l’argomento di Lucas, ma fornisce una parziale difesa di alcuni tratti degli argomenti anti-meccanicistici basati sui teoremi d’incompletezza: «Allo stato attuale della ricerca non è possibile pronunciarsi con un grado sufficiente di certezza sulla attendibilità di questo o quest’altro modello globale della mente. Alla luce, poi, della riflessione metamatematica, non ci sono ragioni sufficienti per ritenere giustificato, neppure a lungo termine, un modello computazionale della mente (nel suo complesso) fondato sulla analogia tra mente e programma di funzionamento (sistema formale) del cervello. Anzi le difficoltà messe in luce da una corretta considerazione dei teoremi di limitazione sembrano puntare in direzioni diverse» (Galvan 2004, 171-172).

²² Che accetti, si noti bene, tutte le ipotesi alla base dell’argomento di Lucas.

In definitiva, comunque, non possediamo oggi una dimostrazione universalmente accettata dell'assunto meccanicista, né una sua confutazione. Ci si può chiedere però speculativamente se è comunque possibile una dimostrazione alternativa della falsità del meccanicismo. Dall'enunciato della disgiunzione appare subito evidente che, se non esistono problemi assolutamente insolubili per la mente, allora essa non può essere simulata correttamente da una macchina di Turing.

Se il mio risultato è preso insieme con l'atteggiamento razionalistico di Hilbert che non è stato refutato dal mio risultato, allora [...] la mente non è meccanica. (Gödel in Wang 1996, 186-187)

L'atteggiamento razionalistico di cui Gödel parla è lo stesso espresso dal famoso proclama di Hilbert (1902, 445): «non v'è *ignorabimus* in matematica», ossia non esiste problema matematico che, in linea di principio, l'umanità non possa risolvere..

5. Il secondo disgiunto

In merito alla questione dell'esistenza di problemi assolutamente insolubili, la letteratura è meno generosa e non esiste neanche una definizione condivisa di dimostrabilità assoluta o di dimostrabilità intuitiva.²³ Si è dunque tentato, come detto, a partire da una proposta dello stesso Gödel ripresa da Myhill (1960) – e indipendentemente da Shapiro (1985) – di studiare le proprietà della dimostrabilità assoluta (e, più in generale, della conoscibilità assoluta) delle proposizioni matematiche attraverso una teoria dell'aritmetica arricchita di un operatore modale atomico di dimostrabilità, opportunamente assiomaticizzato. Su questa linea, allora, sono notevoli gli interventi di Tharp e Fitch, che tentano due argomenti sintattici relativamente semplici per mostrare opposte conclusioni: l'uno, a partire da ipotesi moderatamente deboli, l'inesistenza di proposizioni indimostrabili e l'altro, a partire da ipotesi ancora più deboli, l'esistenza di enunciati assolutamente inconoscibili (o indimostrabili). L'argomento di Tharp, ricostruibile da Wang (1989, 147) e da Reinhardt (1986, 433-434), risente però della non completa plausibilità di un'ipotesi alla base del ragionamento: che sia indimostrabile l'indimostrabilità di ogni proposizione vera.²⁴ D'altra parte neanche la proposta di Fitch (1963) è del tutto soddisfacente. In questo caso infatti, la proposizione indimostrabile di cui Fitch mostrerebbe l'esistenza deve necessariamente contenere il succitato operatore modale, ed è dunque una proposizione solo semi-matematica, non un enunciato aritmetico (un problema diofanteo) come nella disgiunzione di Gödel.²⁵

²³ Una direzione di ricerca importante è quella della "filosofia matematica" di Leitgeb (2013) e Horsten (2014), che riprendono la nozione di "rigore informale" espressa in Kreisel (1972). Nella stessa direzione vedi anche Antonutti-Marfori (2011), che, a partire da una prospettiva epistemica dal sapore anti-formalista, difende la necessità di analizzare il concetto di provabilità informale dal punto di vista della pratica matematica.

²⁴ Se B è l'operatore che esprime la dimostrabilità, allora l'argomento di Tharp si basa sull'assunto

$$\forall p (p \rightarrow \neg B \neg B p),$$

per ogni p appartenente al linguaggio arricchito del succitato operatore modale.

²⁵ Se K è l'operatore che esprime la conoscenza, allora il testimone addotto da Fitch come proposizione inconoscibile è della forma

$$p \wedge \neg K p.$$

Altri tentativi non del tutto soddisfacenti son quelli messi in luce da Feferman e Solovay (1990) che evidenziano l'esistenza di enunciati indecidibili dal punto di vista pratico, ossia di enunciati decidibili *a priori*, ma troppo complessi per essere decisi in un tempo ragionevole sia da un essere umano sia da un computer.²⁶ Boolos (1982) invece mostra che ci sono enunciati aritmetici veri "estremamente indecidibili", ossia caratterizzati (all'interno dell'aritmetica di Peano) da proprietà aritmetico-modali possedute anche da tutti gli altri enunciati dell'aritmetica di Peano. Nonostante il nome suggestivo, Boolos non dimostra che gli enunciati estremamente indecidibili siano indimostrabili in senso assoluto, ma solo all'interno dell'aritmetica di Peano.

Più di recente, Williamson (2016) avanza un originale argomento metafisico per la conclusione che ogni enunciato vero è assolutamente dimostrabile basato sulla plausibilità dell'esistenza controfattuale di esseri capaci di dimostrarlo. Williamson, dopo aver argomentato la non-analiticità delle verità matematiche, sostiene che la conoscenza matematica non deriva solamente dalle dimostrazioni, ma, per quel che riguarda gli assiomi, anche dalla storia evolutiva degli esseri umani, per cui alcuni enunciati ci sembrano «primitivamente convincenti» in forza delle specifiche caratteristiche ereditarie del nostro cervello (Williamson 2016, 246). Ne segue che, dato un qualsiasi enunciato vero ϕ , è plausibile ritenere che potrebbero evolversi creature capaci di trovare ϕ primitivamente convincente e che, dato che ϕ sarebbe dimostrabile da tali creature,²⁷ allora ne sarebbe stabilita la dimostrabilità in senso assoluto. Se accettiamo che ogni enunciato aritmetico sia vero o falso e *tertium non datur*, allora, sostiene Williamson, ogni enunciato aritmetico è assolutamente dimostrabile (o refutabile).²⁸

Ulteriori tentativi si dirigono infine nella direzione di considerare enunciati indipendenti della teoria degli insiemi come l'assioma di scelta o l'ipotesi del continuo come possibili candidati per essere enunciati assolutamente indecidibili. Come ben spiegato da Koellner (2006), questa è l'opinione di Gödel fino al 1946, quando esprime la speranza che si possa giungere ad un teorema generalizzato di completezza per la teoria degli insiemi basato sulla nozione di computabilità di Turing che stabilisca l'impossibilità di enunciati assolutamente indecidibili. Gödel giunge poi alla visione più matura espressa della conferenza del 1951. Koellner conclude che, nel contesto della teoria degli insiemi,

non esiste ad oggi alcun argomento solido per mostrare che un dato enunciato è assolutamente indimostrabile. Non possediamo neanche un chiaro contesto entro il quale un tale argomento potrebbe procedere. (Koellner 2006, 188)

²⁶ Ad esempio l'enunciato "il valore della $10^{10^{100}}$ -esima cifra decimale di π è 7" può, in linea di principio, essere verificato o falsificato da un semplice procedimento meccanico. In pratica, però, la verifica di tale enunciato è del tutto al di fuori delle reali capacità di qualsiasi calcolatore.

²⁷ Ovviamente qui "dimostrabile" è usato nel senso allargato per cui se ϕ è un assioma allora ϕ è dimostrabile tramite una dimostrazione di una riga che riconosce appunto che ϕ è un assioma.

²⁸ Si noti tuttavia che Williamson non crede che il suo argomento possa essere usato per sostenere posizioni anti-meccanicistiche: anche assumendo che le succitate creature siano possibili (futuri) esseri umani allora si ha che, per ogni verità aritmetica è possibile che esista un matematico che può (o potrà) provarla. Questo è diverso dal dire che è possibile che ci sia un matematico che può (o potrà) provare tutte le verità aritmetiche. In altre parole l'argomento di Williamson non stabilisce alcuna asimmetria fra le capacità degli esseri umani e le capacità delle macchine di Turing.

Bibliografia

- Aldini A., Fano V. e Graziani P., 2016, «Theory of Knowing Machines: Revisiting Gödel and the Mechanistic Thesis», *History and Philosophy of Computing*, pp. 57-70.
- Alexander S., 2014, «A machine that knows its own code», *Studia Logica*, CII, pp. 567-576.
- Anderson C.A., 1993, «Towards a logic of a priori knowledge», *Philosophical Topics*, XXI, 2, pp. 1-20.
- Antonutti-Marfori M., 2010, «Informal proofs and Informal Rigour», *Studia Logica* XCVI, 2, pp. 261-272.
- Benacerraf P., 1967, «God, the devil and Gödel», *The Monist*, LI, pp. 9-32.
- Berto F., 2008, *Tutti pazzi per Gödel!*, Bari, Laterza.
- Boolos G., 1982, «Extremely Undecidable Sentences», *The Journal of Symbolic Logic*, XLVII, 1, pp. 191-196.
- Bruni R., 2005, «Su alcune ricadute concettuali dei teoremi d'incompletezza di Kurt Gödel», *Annali del Dipartimento di Filosofia 2003-2004*, Firenze University Press, pp. 209-226.
- Bruni R., 2006, «Gödel, Turing, the Undecidability Results and the Nature of Human Mind», *Logical Approaches to Computational Barriers: Proceedings of the Second Conference on Computability in Europe*, A. Beckmann et alii (ed.), pp. 37-46.
- Carlson T., 2000, «Knowledge, machines and the consistency of Reinhardt's strong mechanistic thesis», *Annals of Pure and Applied Logic*, CV, pp. 51-82.
- Cassou-Noguès P., 2008, *I demoni di Gödel*, Milano, Bruno Mondadori.
- Chalmers, D. J., 1996, «Minds, Machines and Mathematics», *Psyche*, II, pp. 11-20.
- Chihara C. S., 1972, «On alleged refutations of mechanism using Gödel's incompleteness results», *The Journal of Philosophy* LXIX, pp. 507-526.
- Fano V. e Graziani P., 2013, «Mechanical Intelligence and Gödelian arguments» in *The Legacy of A. M. Turing*, E. Agazzi (ed.), Franco Angeli, pp. 48-71.
- Fano V. e Graziani P., 2011, «Gödel and the fundamental incompleteness of human self-knowledge» *Logic and Philosophy of Science*, IX, 1, pp. 263-274.
- Feferman S., 1962, «Transfinite recursive progressions of axiomatic theories», *Journal of Symbolic Logic* XXVII, pp. 259-316.

- Feferman S., 2006, «Are There Absolutely Unsolvable Problems? Gödel's Dichotomy», *Philosophia Mathematica*, XIV, 2, pp. 134–152.
- Feferman S. e Solovay R., 1990, «Introductory Note to 1972a» in (Gödel 1990, pp. 281-304).
- Fitch, F., 1963, «A Logical Analysis of Some Value Concepts», *Journal of Symbolic Logic* I, pp. 204-217.
- Franzen, T., 2004, *Inexhaustibility. A non-exhaustive treatment*, Lecture Notes in Logic XXVIII, A. K. Peters.
- Galvan S., 2004, «Gödel e il modello computazionale della mente», *Rivista di Filosofia Neo-Scolastica*, XCVI, 1, pp. 145-174.
- Gandy R., 1988, «The confluence of ideas in 1936», *The Universal Turing Machine*, R. Henke ed., New York, Oxford University Press, pp. 55-111.
- Gödel K., 1951, «Some basic theorems on the foundations of mathematics and their implications», in (Gödel 1995, pp. 304-323).
- Gödel K., 1990 *Collected Works, Volume II*, S. Feferman, et alii ed., New York: Oxford University Press.
- Gödel K., 1995, *Collected Works: Volume III*, S. Feferman et alii ed., New York, Oxford University Press.
- Hilbert D., 1902, «Mathematical problems: Lecture delivered before the International Congress of Mathematicians at Paris in 1900», *Bulletin of the American Mathematical Society*, VIII, pp. 437-479.
- Hofstadter D., 1975, *Gödel, Escher Bach*, New York, Basic Books.
- Horsten L., 2005, «Remarks on the Content and Extension of the Notion of Provability», *Logique et Analyse*, XLVIII, 189/192, pp. 15-32.
- Horsten L., 2013, «Mathematical Philosophy?» in *The Philosophy of Science in a European Perspective*, IV, Springer, pp. 73-86.
- Horsten L. e Welch P., 2016, *Gödel's Disjunction*, Oxford, Oxford University Press.
- Labinaz P., 2016, «Introduzione» in *J. R. Lucas Against Mechanism*, P. Labinaz (ed.), Milano, Mimesis.
- Leitgeb H., 2009, «On Formal and Informal Provability» in *New Waves in Philosophy of Mathematics*, O. Bueno e Ø. Linnebo (eds.), Basingstoke, Palgrave Macmillan.

- Leitgeb H., 2013, «Scientific Philosophy, Mathematical Philosophy, and All That», *Metaphilosophy*, XLIV, pp. 267-275.
- Lindström P., 2001, «Penrose's new argument», *Journal of Philosophical Logic*, XXX, pp. 241-250.
- Lucas J. R., 1961, «Minds, Machines and Gödel», *Philosophy*, XXXVI, pp. 112-137.
- Lucas J. R., 1968, «Satan Stultified: A Rejoinder to Paul Benacerraf», *The Monist*, LII, pp. 145-158.
- Lucas J. R., 1970, *The Freedom of the Will*, Oxford, Clarendon Press.
- Lucas J. R., 1996, «Minds, Machines and Gödel: A retrospect», *Machines and Thoughts: The Legacy of Alan Turing*, P. Millican and A. Clarks ed., Oxford, Oxford University Press, pp. 103-124.
- Koelnerr P., 2006, «On the Question of Absolute Undecidability», *Philosophia Mathematica*, XIV, pp. 153–188.
- Kreisel G., 1972, «Informal Rigour and Completeness Proofs» in *Problems in the Philosophy of Mathematics*, I. Lakatos (ed.), Amsterdam, North Holland.
- Penrose R., 1989, *The Emperor's New Mind*, Oxford, Oxford University Press.
- Penrose R., 1994, *Shadows of the Mind*, Oxford, Oxford University Press.
- Reinhardt W. N., 1986, «Epistemic theories and the interpretation of Gödel's incompleteness theorems», *Journal of Philosophical Logic*, XV, pp. 427-474.
- Rogers, H. (1967) *Theory of Recursive Functions and Effective Computability*, New York, McGraw-Hill.
- Shapiro S., 2016, «Idealization, Mechanism and Knowability» in (Horsten and Welch 2016, pp. 189-206).
- Shapiro S., 1985, «Epistemic and Intuitionistic Arithmetic» in *Intensional Mathematics*, Amsterdam, North Holland, pp. 11-46.
- Soare R. I., 2016, *Turing Computability: Theory and Applications*, New York, Springer-Verlag.
- Turing A. M., 1939, «Systems of logic based on ordinals», *Proceedings of the London Mathematical Society*, XLV, pp. 161-228.
- Turing A. M., 1950, «Computing Machinery and Intelligence», *Mind*, LIX, pp. 433-460.
- Wang H., 1974, *From Mathematics to Philosophy*, London, Routledge and Kegan Paul.

Wang H., 1989, «Tharp and Conceptual Logic», *Synthese*, LXXXI, 2, pp. 141-152.

Wang H., 1996, *A Logical Journey*, Cambridge (MA), The MIT Press.

Williamson T., 2016, «Absolute Provability and safe Knowledge of Axioms», in (Horsten e Welch 2016, pp. 243-253).