

Wittgenstein e Putnam: meravigliosa sintonia. Il caso dei cervelli in una vasca

Massimo Dell'Utri

Università di Sassari
dellutri@uniss.it

Abstract The argument that Hilary Putnam presented in 1981 against the Brains in Vat Hypothesis (BIV) is based on a semantic conception that has close affinities with some of Wittgenstein's reflections on the meaning of linguistic expressions and the nature of mental phenomena presented in the *Philosophical Investigations*. This has been explicitly acknowledged by Putnam, but there is a fundamental aspect of Putnam's argument against the BIV (an aspect of which Putnam himself was not fully aware) whose origin can also be traced in Wittgenstein – in the latter Wittgenstein, that of *Of Certainty*. It is in fact centred on the notion of “hinge-proposition”. In the article I provide a reconstruction of Putnam's proof against the BIV, I emphasise the links with some passages in the *Philosophical Investigations*, I highlight the aspect of *Of Certainty* that makes Putnam's argument work, and I point out how this aspect is already present in the notion of “quasi-necessary truths” developed by Putnam in the 1960s. On the basis of all this, I conclude that the argument against the BIV presented by Putnam in 1981 is correct, despite the criticism to which it has been repeatedly subjected.

Keywords: Brains in a vat, Hilary Putnam, Ludwig Wittgenstein, quasi-necessity, hinge-propositions

Received 01 04 2024; accepted 24 06 2024.

0. Introduzione

L'argomento che Hilary Putnam presentò nel 1981 contro l'ipotesi dei cervelli in una vasca (CIV) risente di alcune riflessioni di Ludwig Wittgenstein sul significato delle espressioni linguistiche e sulla natura dei fenomeni mentali presentate nelle *Ricerche filosofiche* (cfr. Putnam 1981: Cap. 1). Questo legame con Wittgenstein è stato espressamente riconosciuto da Putnam, ma c'è un fondamentale aspetto dell'argomento putnamiano contro i CIV non messo in luce da Putnam all'atto di presentare il proprio argomento e che ha a che vedere con un caposaldo del pensiero di Wittgenstein – dell'ultimissimo Wittgenstein, quello di *Della certezza*.

In quel che segue mi propongo di fornire una ricostruzione dell'argomento di Putnam contro i CIV, di evidenziare i tratti essenziali delle argomentazioni delle *Ricerche filosofiche* su cui Putnam si è basato, e di portare alla luce l'aspetto di *Della certezza* che rende possibile all'argomento putnamiano di funzionare.

Ciò che l'ipotesi ci invita a immaginare è presto detto:

Suppose evil scientists removed your brain from your body while you slept, and set it up in a life-support system in a vat. Suppose they then set out to trick you into believing that you were not just a brain in a vat, but still up and about, engaging in a normally embodied round of activities in the real world (Dennett 1991: 3).

A questo scenario immaginario Putnam ha aggiunto un computer che, collegato alle sinapsi del cervello con una serie di fili e alimentato da una moltitudine di software diversi, fa sperimentare al cervello le numerose situazioni di una vita “normale”. La domanda che sorge immediatamente è: “Come puoi dire di non essere un cervello in una vasca?”. L’idea è che non sia possibile, cioè che non si possa dimostrare al di là di ogni dubbio di esistere insieme al resto del mondo. Pertanto, nulla di ciò che pensiamo o diciamo sembra essere certo, mentre tutto sembra essere soggetto a dubbi inquietanti – come sosterebbe uno scettico radicale.

La particolarità dell’ipotesi dei CIV è che essa solleva diverse questioni, non solo quella scettica: è un’ipotesi che riguarda la metafisica, l’epistemologia, la semantica, la filosofia del linguaggio, la filosofia della mente, la neurofisiologia e altro ancora. Nei paragrafi seguenti affronterò la questione neurofisiologica e quella riguardante la validità della confutazione dell’ipotesi presentata da Putnam. Prima di fare ciò, è necessario fare un breve accenno al pensiero di Hilary Putnam.

1. I principali contributi di Hilary Putnam alla filosofia

Immagino che se vi chiedessero quali sono i principali contributi alla filosofia che Putnam ha dato nel corso della sua lunga carriera, probabilmente menzionereste questioni come la concezione causale del significato e del riferimento, o l’ipotesi funzionalista in filosofia della mente, o la critica della dicotomia fatto/valore, o la discussione del realismo metafisico, o l’ipotesi della Terra Gemella, o la stessa ipotesi dei CIV, per citarne solo alcune. Si tratta di questioni innegabilmente molto importanti su cui Putnam ha lasciato un’impronta duratura e sono certamente d’accordo con ognuna di queste scelte; ma, se questa domanda venisse posta a me, preferirei dare una risposta diversa: sottolineerei l’importanza della sua nozione di *verità quasi-necessarie*.

Questa nozione ha lo scopo di caratterizzare lo status di un tipo speciale di enunciati: enunciati la cui negazione è inconcepibile. Tradizionalmente, gli enunciati di questo tipo sono stati definiti “necessari” in un senso assoluto e sono stati considerati epistemicamente *eterni*, ma gli sviluppi della riflessione filosofica a partire almeno dall’ultimo decennio del XIX secolo – in parte guidata dalla critica degli empiristi, da un lato, e dalla prospettiva fallibilista dei pragmatisti, dall’altro – hanno rivelato l’insostenibilità di tale interpretazione della necessità, evidenziando come nessun enunciato possa godere per sempre di uno status privilegiato. Abbracciando il fallibilismo e sentendosi profondamente insoddisfatto delle conclusioni dei positivisti logici, Putnam ha sostenuto che, sebbene nessun enunciato sia immune da revisione, alcuni enunciati sono tali che la loro negazione appare insensata: essi *sono* immuni da revisione, quindi *necessari*, ma – e questo è il punto – solo *contestualmente*, cioè rispetto allo schema concettuale in vigore in una determinata epoca. Inoltre, per lo stesso motivo il loro contenuto può essere conosciuto indipendentemente dall’esperienza, il che li qualifica come *a priori*. Un esempio può chiarire il punto.

Consideriamo l’enunciato «Due rette parallele si estendono all’infinito senza mai incontrarsi» – chiamiamolo *S*. Prima dell’avvento delle cosiddette geometrie non euclidee – come quella sviluppata da Bernhard Riemann – per secoli gli assiomi e i postulati forniti da Euclide hanno rappresentato le *leggi* a cui i movimenti dei corpi nello spazio obbedivano necessariamente – leggi considerate certe, indistruttibili ed eterne, la cui falsità era letteralmente inconcepibile. Per questo motivo, l’enunciato *S* non solo era considerato vero, ma il suo *status* era quello di una verità necessaria e a priori. Tuttavia, dopo l’irruzione sulla scena delle teorie di Einstein, accade l’inconcepibile. Qualcosa che ha lo status di verità necessaria e a priori si rivela essere una verità contingente e a

posteriori. Questo però ancora non inquadra lo status *effettivo* di S secondo Putnam: la seguente citazione ci farà capire cosa intende.

Although modern physics does not *yet* say that space is Riemannian, it does say that our space has variable curvature. This means that if two light rays stay a constant distance apart for a long time and then come closer together after passing the sun, we do not *say* that these two light rays are following curved paths through space, but we say rather that they follow straight paths and that two straight paths may have a constant distance from each other for a long time and then later have a decreasing distance from each other (Putnam 1962: 242).

Ciò su cui Putnam richiama la nostra attenzione è che è la presenza di una nuova teoria fisica, alternativa rispetto a quella accettata finora, che può determinare un cambiamento nello status di un enunciato come S . Ma, in assenza di tale teoria, S continua a mantenere il suo status di affermazione necessaria a priori: se a qualcuno venisse in mente di sottoporlo a test, esso non potrebbe in alcun modo essere dimostrato falso. *Mutatis mutandis*, lo stesso vale per la geometria euclidea nel suo insieme.

Perciò, ci dice Putnam, la geometria euclidea è una teoria sintetica, ma in quanto tale molto *sui generis*: prima dell'avvento delle geometrie non euclidee, nessuna prova empirica avrebbe potuto dimostrare la sua inadeguatezza. Essa serviva come *paradigma*, come sfondo concettuale a partire dal quale era possibile osservare, spiegare e prevedere i fenomeni che si verificano nello spazio fisico. Nessun esperimento poteva dimostrarne l'inadeguatezza, perché ogni esperimento che si riusciva a concepire si basava inevitabilmente su ipotesi dedotte da essa e, quindi, ne presupponeva la validità. Ed è proprio perché sembrava intoccabile da qualsiasi esperimento che essa veniva considerata come un tipico esempio di teoria necessaria e a priori. La situazione cambia radicalmente quando vengono presentate teorie alternative e valide, le uniche in grado di fornire le premesse teoriche sulla base delle quali si possono formulare esperimenti volti a confutare la geometria euclidea: il suo carattere di teoria sintetica diventa allora chiaro. Cosa dire allora dell'enunciato S , un enunciato derivato da una teoria che ha goduto di uno status *paradigmatico* fino all'avvento di una o più teorie alternative? La risposta di Putnam è che S ha lo status di verità necessaria e a priori finché non ci sono alternative al paradigma su cui poggia, e lo status di enunciato contingente e a posteriori quando queste alternative appaiono e si dimostrano valide. Come ho accennato sopra, per enunciati di questo tipo egli ha coniato il nome di "quasi-necessari rispetto a uno schema concettuale" (Putnam 1994: 251). Si tratta dei più importanti elementi del nostro sistema concettuale: ad esempio i principi della matematica e della logica, le leggi della scienza naturale e gli enunciati che fungono da principi quadro all'interno del sistema, come la tesi che esiste un passato o la tesi secondo cui si può dire che una persona possiede una certa conoscenza solo se è in possesso di una giustificazione adeguata.

Putnam propone di riservare a questi elementi l'etichetta di "sintetici", ossia enunciati rivedibili in linea di principio, ma non sulla base dei risultati dell'esperienza, a meno che questi risultati non vengano interpretati per mezzo di una teoria alternativa a quella attuale. Finché tale teoria non compare e non viene convalidata, gli enunciati quasi-necessari non possono essere toccati dall'esperienza e sono quindi "non empirici", "non a posteriori". In una parola, sono a priori. Ecco quindi come, un paio di secoli dopo Kant, si può parlare secondo Putnam di enunciati *sintetici a priori*.

La cosa che vorrei sottolineare qui è la strettissima affinità tra la nozione di enunciato quasi-necessario e la nozione wittgensteiniana di "proposizione cardine", una nozione che caratterizza quelle proposizioni che «sono esenti da dubbio, come se fossero i perni sui quali si muovono» (Wittgenstein 1969: § 341) tutte le altre proposizioni del nostro

linguaggio. Si tratta di perni assai solidi – molto più solidi di quanto non siano le altre proposizioni: «Se voglio che la porta si apra, i perni devo essere saldi» (Wittgenstein 1969: § 343). E la loro solidità è data dal fatto che queste proposizioni cardine non vengono messe in discussione, non creano perplessità o dubbi di qualche genere, rendono sicuro l'uso del linguaggio e ci fanno sentire perciò appagati: «La mia *vita* consiste in questo: che sono appagato di alcune cose» (Wittgenstein 1969: § 344).

E l'affinità tra enunciati quasi-necessari e proposizioni cardine non si ferma qui. Anche queste ultime non possono infatti essere considerate *empiriche*, nonostante l'apparenza: «Del fondamento di tutto l'operare con i pensieri (con il linguaggio) fanno parte non soltanto le proposizioni della logica, ma anche certe proposizioni che hanno la *forma* di proposizioni empiriche» (Wittgenstein 1969: § 401; corsivo mio) e invece non lo sono. Da sottolineare anche che il carattere particolare di questo tipo di proposizioni non è a esse intrinseco, non ha a che fare con una loro presunta natura essenziale e dunque immodificabile: semplicemente, è così che accade, senza che vi sia una qualche ragione metafisica di ordine superiore:

Le proposizioni che per me sono incontestabili non le imparo esplicitamente. Posso forse *trovarle* in seguito così come si trova l'asse di rotazione di un corpo rotante. Quest'asse non è stabile nel senso che sia mantenuto stabile, ma nel senso che è il movimento intorno ad esso a determinarne l'immobilità (Wittgenstein 1969: § 152).

Movimento che perciò potrebbe cambiare, e magari anche in modo drastico, capovolgendo la relazione tra proposizioni cardine e proposizioni empiriche:

Ci si potrebbe immaginare che certe proposizioni che hanno forma di proposizioni empiriche vengano irrigidite e funzionino come una rotaia per le proposizioni empiriche non rigide, fluide; e che questo rapporto cambi col tempo, in quanto le proposizioni fluide si solidificano e le proposizioni rigide diventano fluide (Wittgenstein 1969: § 96).

Tornerò più avanti sugli enunciati quasi-necessari che funzionano come una rotaia. Volgiamoci ora alla questione neurofisiologica sotto cui è possibile discutere l'ipotesi dei CIV e, subito dopo, alla dimostrazione putnamiana della falsità dell'ipotesi.

2. La questione neurofisiologica

La domanda centrale in questo contesto è la seguente: «È possibile che un cervello viva in una vasca di sostanze nutritive?». Mario Alai pone una domanda ulteriore: «Potrebbe essere collegato a un computer che, inviando impulsi adeguati, simuli le esperienze di una vita normale?» (Alai 1994: 3). Queste domande configurano possibilità diverse, in quanto la prima si interroga sulla possibilità che una vasca di sostanze nutritive sia sufficiente a far vivere il cervello, mentre la seconda aggiunge la possibilità che in tale condizione il cervello abbia una coscienza.

Per quanto riguarda quest'ultima domanda, sono propenso a condividere il punto di vista di Robert Hanna e a rispondere negativamente. Ecco cosa sostiene:

I hold that the supposed consciousness of a causally detached brain – say, a living brain floating listlessly in a vat, as in Hilary Putnam's famous thought-experiment – even though it seems both conceivable and logically possible, just would not be a consciousness like ours. On our view, a consciousness necessarily involves a brain that is causally-dynamically coupled with all the other vital systems, organs, and processes of our living body (Hanna 2011: 21).

Se Hanna ha ragione, un cervello ha bisogno di un intero corpo per essere cosciente, insieme alla complessa rete di interconnessioni con il sistema vitale del corpo. Un CIV staccato dal corpo, invece, è solo un pezzo di carne, anche se è vivo. Ma è vivo? Questo ci riporta alla prima domanda.

Secondo Alai, bisognerebbe chiederlo a neurofisiologi e informatici, perché solo chi possiede conoscenze e competenze specifiche in questi campi può dirlo. Tuttavia, aggiunge, per un profano «sembrerebbe che, sebbene un'operazione così sofisticata sia attualmente irrealizzabile, sia comunque possibile almeno in linea di principio» (Alai 1994: 3), sottintendendo che un profano e uno scienziato risponderebbero in modo diverso. E nella stessa posizione del profano pone il filosofo, se è questo che intende quando afferma che «la questione filosofica non ha nulla a che fare con quella neurofisiologica: i CIV sono possibili» (Alai 1994: 4).

A differenza di Alai, penso che neurofisiologi, profani e filosofi siano tutti sulla stessa barca nel pensare che sia in linea di principio possibile per un cervello vivere in una vasca di nutrienti. Tuttavia, mi sembra che questa possibilità di principio possa facilmente ridursi a un'*effettiva impossibilità*. Ecco cosa intendo, basandomi su un'argomentazione che Putnam ha avanzato nel 1976 durante le sue John Locke Lectures a Oxford.

In genere, quando si parla di “possibilità in linea di principio” si intende almeno la possibilità *logica* – nessun conflitto con i principi logici – o la possibilità *fisica* – nessun conflitto con le leggi naturali – o anche la possibilità *tecnica* – nessun conflitto con lo stato attuale della nostra tecnologia. Esiste tuttavia una nozione ulteriore di possibilità che è rilevante per la discussione di argomenti come il presente e che si colloca in posizione intermedia tra queste due ultime nozioni: la questione che è interessante indagare non è se sia logicamente possibile per un cervello vivere in una vasca di nutrienti, né se sia fisicamente o tecnologicamente possibile,

but rather whether it is physically possible to do this *fast enough to count as a prediction*, whether it is possible to make [all this] ‘in real time’, as computer people say. And while it may follow from the ‘hypothesis of physicalism’ that it must be possible for Logically Omniscient Jones to predict [that a brain lives in a vat of nutrients], and it may even follow that it must be physically possible [...], there is no reason at all to believe, on that basis or on any other, that it must be possible to do it *in real time* (Putnam 1978: 64).

Possiamo certamente concludere che è logicamente e fisicamente possibile che un cervello viva in una vasca, e possiamo anche aggiungere che in linea di principio è tecnicamente possibile ottenere un simile stato di cose, ma in pratica ciò richiederebbe molto probabilmente «so much ‘real time’ that the human race will have long ceased to exist before such [operation] is carried out!» (Putnam 1978: 65). Pertanto, ciò che a prima vista appare come una possibilità finisce per rivelarsi un'*impossibilità pratica*. Se tutto questo è corretto, abbiamo che un CIV rappresenta un'*impossibilità* qualsiasi cosa se ne dica dal punto di vista del neurofisiologo, o del profano, o del filosofo. Con questo possiamo chiudere la nostra discussione della questione neurofisiologica – ci ritorneremo brevemente più sotto – e passare all'argomento putnamiano contro l'ipotesi dei CIV.

3. La dimostrazione di Putnam

In questo paragrafo mostrerò due modi rigorosi per ricostruire l'argomento putnamiano volto a dimostrare che non siamo CIV, discuterò alcune obiezioni e aggiungerò alcuni commenti generali. Prima di farlo voglio brevemente sottolineare il debito che Putnam

ha esplicitamente riconosciuto nei confronti di Wittgenstein, considerato tra gli ispiratori della sua argomentazione.

Quest'ultima, infatti,

first occurred to me when I was thinking about a theorem in modern logic, the “Skolem-Löwenheim Theorem”, and I suddenly saw a connection between this theorem and some arguments in Wittgenstein’s *Philosophical Investigation*” (Putnam 1981: 7).

Non è difficile capire a quali argomenti di Wittgenstein si stia riferendo qui Putnam: sono gli argomenti contro l'ipotesi dell'esistenza di un linguaggio privato, quelli che gli hanno fatto affermare cose come: «Il *concetto* ‘dolore’ l’hai imparato con il linguaggio» (Wittgenstein 1953: § 384); oppure: «Un ‘processo interno’ abbisogna di criteri esterni» (Wittgenstein 1953: § 580); e anche: «Il pensare non è un processo incorporeo che presti vita e senso al parlare, e che sia possibile staccare dal parlare, come il maligno stacca l’ombra di Schlemiehl dal suolo» (Wittgenstein 1953: § 339). In buona sostanza, non sarebbe possibile parlare un qualsiasi linguaggio storico-naturale le cui espressioni siano correttamente dotate di significato e di riferimento se non si fosse all’interno di un tessuto sociale e materiale regolato da norme e convenzioni. Proprio Wittgenstein, più efficacemente di altri, ha a più riprese sottolineato «that without such public norms, norms shared by a group and constituting a ‘form of life’, language and even thought itself would be impossible» (Putnam 1981: 107).

Visto cosa si muove sullo sfondo della dimostrazione della falsità dell’ipotesi dei CIV, veniamo alle sue due ricostruzioni.

La prima è quella usata talvolta da Putnam nei suoi corsi. Consiste in quattro passi. Per apprezzarla, abbiamo bisogno di due premesse e di una definizione. La prima premessa è il *vincolo causale sul riferimento*, secondo il quale non ci si può riferire a certi oggetti usando un certo termine se non si sono avute interazioni causali cariche di informazioni con questi oggetti (o con cose o proprietà con le quali si può formulare una descrizione di questi oggetti). Un corollario ne è che le spiegazioni del riferimento fornite dalle teorie che non presuppongono tale vincolo finiscono per basarsi su qualcosa di analogo a un atto di magia, magari perché postulano delle facoltà mentali difficilmente collocabili in un quadro naturalista – sono quelle che Putnam chiama “teorie magiche del riferimento”. La seconda premessa è la *proprietà devirgolettante del riferimento*, secondo la quale nella mia lingua qualsiasi termine “*t*” si riferisce a tutte e solo le *t* (senza virgolette), quando possiede effettivamente un riferimento: ad esempio, “finestra” si riferisce a tutte e sole le finestre. Segue la definizione di *italiano da vasca*: l’italiano da vasca è la lingua che parlerei se fossi un CIV. Ed ecco la dimostrazione nella sua prima ricostruzione:

- (1) Nella mia lingua, “cavolo” si riferisce ai cavoli [seconda premessa]
- (2) In italiano da vasca, “cavolo” non si riferisce ai cavoli [prima premessa]
- (3) Pertanto, la mia lingua non è l’italiano da vasca [1, 2]
- (4) Pertanto, non sono un CIV [3, definizione di italiano da vasca]

La seconda ricostruzione della dimostrazione è la seguente:

- (1) Io posso porre la domanda: Sono da sempre un CIV?
- (2) Se fossi da sempre un CIV, non potrei porre questa domanda
- (3) Pertanto, non sono da sempre un CIV (Ebbs 2016: 34)

La differenza con la precedente sta nel fatto che qui c'è una sola premessa (quella *causale*), non si parla (esplicitamente) di lingue, e il piano su cui si colloca la dimostrazione è *epistemologico*: la domanda citata è come quelle che si pongono nel contesto dell'indagine scientifica e non.

Si noti che per Putnam la mancanza di un riferimento genuino in enunciati e pensieri annulla per ciò stesso la possibilità di pensare e parlare: proprio questo sarebbe il difetto principale dei CIV, che dunque non potrebbero nemmeno autenticamente dire o pensare «Io sono un CIV».

Cerchiamo ora di dare una valutazione equilibrata della dimostrazione putnamiana.

4. Come valutare la dimostrazione?

Entrambe le ricostruzioni dell'argomento putnamiano sono valide. Tuttavia, devo ammettere che esse lasciano una strana sensazione: dimostrano davvero che siamo esseri umani reali? Si potrebbe essere tentati dal seguente ragionamento:

[...] the argument does not prove that I am not a brain in a vat; it makes the weaker point that if I were a brain in a vat I could not coherently think [or say] that I were. But from this it does not follow that the skeptical scenario that I might be a brain in a vat could not be true. This would only follow if the brain-in-a-vat scenario violated a physical law or contained a logical contradiction (Bernecker 2016: 58-59).

Equivalentemente:

If I accept the argument, I must conclude that a brain in a vat can't think truly that it is a brain in a vat, even though others can think this about it. What follows? Only that I cannot express my skepticism by saying "Perhaps I am a brain in a vat". Instead I must say "Perhaps I can't even *think* the truth about what I am, because I lack the necessary concepts and my circumstances make it impossible for me to acquire them?". If this doesn't qualify as skepticism, I don't know what does (Nagel 1986: 73).

Questa sembrerebbe la cosa giusta da dire sulla dimostrazione di Putnam, ma ad un attento esame non credo che sia la morale corretta. Credo che dietro questo ragionamento si possano individuare due presupposti nascosti: uno è che contro un'ipotesi scettica come quella dei CIV si debba necessariamente avanzare un argomento *diretto* – ossia che la confuti direttamente e chiaramente, senza passare da ulteriori considerazioni –, e l'altro è che un argomento diretto contro lo scetticismo sia destinato a essere una *petizione di principio* – ossia a non poter evitare di assumere implicitamente la stessa tesi che è invece da dimostrare. La penso così perché mi sembra si possa ritenere che la dimostrazione di Putnam stabilisce unicamente la tesi più debole menzionata da Bernecker e Nagel solo se si è convinti, da un lato, che per stabilire la tesi più forte sia necessario *dimostrare* al di là di ogni dubbio la falsità dell'ipotesi dei CIV e la verità dell'ipotesi dell'effettiva esistenza del mondo esterno – ossia, è necessario fornire una dimostrazione diretta – e, dall'altro lato, che qualsiasi dimostrazione diretta di questo tipo non può che “assumere” che gli esseri umani facciano un uso referenziale dei termini del loro linguaggio – che è esattamente ciò che doveva essere dimostrato. L'ultimo punto rappresenta un dubbio ricorrente nei confronti della dimostrazione di Putnam. Diversi autori hanno basato il loro rifiuto della dimostrazione sulla sua presunta fallacia di essere appunto una petizione di principio (cfr. Dell'Utri 1990):

If it is agreed that you couldn't refer to BIVs if you were in the VAT scenario, don't you have to know that you are *not* in the VAT scenario before you can know that you

can refer to BIVs – and thus know exactly the thing that the VAT argument is supposed to prove? (Thorpe & Wright 2022: 66).

Per commentare questo punto propongo di esaminare brevemente un altro dubbio sollevato contro la dimostrazione di Putnam. Eccolo:

One may suspect that there is something wrong with this argument already when Putnam claims that the BIV hypothesis, although *physically* possible, may be ruled out by philosophy, through an *a priori* reasoning. For it is not clear how something can be physically (i.e., *a posteriori*) possible if it is already *a priori* impossible (Alai 1994: 23).

È la risposta a questo dubbio che ci metterà sulla strada giusta per comprendere il senso della dimostrazione di Putnam. Quando l'ha presentato per la prima volta Putnam ha chiarito che stava cercando di studiare

the *preconditions* for *thinking about, representing, referring to*, etc. [...] by *reasoning a priori*. Not in the old 'absolute' sense (since we don't claim that magical theories of reference are *a priori* wrong), but in the sense of inquiring into what is *reasonably* possible *assuming* certain general premisses, or making certain very broad theoretical assumptions. Such a procedure is neither 'empirical' nor quite 'a priori', but has elements of both ways of investigating (Putnam 1981: 16).

Si tratta di un'affermazione rivelatrice in quanto dimostra, da un lato, che la sua dimostrazione non si basa su un ragionamento a priori di tipo tradizionale – privo di elementi empirici – e, dall'altro, che Putnam ci invita a riflettere sulla nostra situazione epistemologica – sui nostri *standard epistemologici*, che secondo Putnam non sono fissi e dati una volta per tutte: sono fallibili e rivedibili, persino eliminabili, ma finché restano in campo hanno forza sufficiente per svolgere il loro lavoro. In breve, esemplificano le nozioni contestuali di necessario e di a priori che abbiamo visto sopra: non possiamo *al momento* dare un senso alla loro falsità perché sono una sorta di *proposizioni cardine* wittgensteiniane che danno forma al nostro modo di fare ricerca in generale:

In our inquiries we can do no better than to start in the middle, relying on already established beliefs and inferences, and applying our best methods for re-evaluating particular beliefs and inferences and arriving at new ones (Ebbs 2016: 31).

E tra le credenze consolidate con cui iniziamo “in the middle” ci sono quelle secondo cui non siamo CIV, o che sul riferimento ci devono essere vincoli causali, o che le teorie magiche del riferimento sono sbagliate, o – come detto sopra – che esiste un passato. Si tratta di «what we do not doubt in our hearts» (Peirce 1955: 229), di «default entitlements» (Williams 2000: 611), di «entitlements *without* reasons [that] are like citizens under the laws, innocent until proven guilty» (Macarthur 2019: 279-80). Possono sembrare le solite credenze empiriche suscettibili di essere più o meno facilmente falsificate, ma in realtà sono sintetiche a priori – credenze, enunciati, giudizi immaginare la cui falsità risulta per noi insensato: «To think that [brains in vats] are a possibility, I would have to assume that my entire view of the world is wrong in a way I am utterly unable to make cohere with anything I believe» (Putnam 2012: 487)¹. L'impossibilità “pratica” dell'esistenza di CIV che abbiamo incontrato alla fine del paragrafo 2 assume ora una coloritura più spiccatamente epistemologica: date le caratteristiche fondamentali del nostro schema concettuale complessivo, la possibilità dei CIV si rivela *inintelligibile*. E questo rafforza il legame con Wittgenstein.

¹ Nella citazione ho sostituito “brains in vats” a “zombies”.

Infatti, la tesi dell'inintelligibilità di ipotesi come quella dei CIV viene rinvenuta da Putnam – nell'articolo del 2012 da cui è tratta la precedente citazione – nella discussione riportata in *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*, dove Wittgenstein sostiene che il problema della comprensione di un enunciato come «Credo in un Giudizio Universale» fatta da un credente non è di tipo *linguistico*: non è che il non credente non afferri il significato delle parole dell'enunciato in questione. Piuttosto, il fatto è che, nonostante il chiaro significato delle parole, l'enunciato appare *inintelligibile* al non credente (Wittgenstein stesso) a causa del contrasto tra l'enunciato e le proposizioni cardine appartenenti alla visione del mondo del non credente. Ecco Wittgenstein:

If you ask me whether or not I believe in a Judgement Day, in the sense in which religious people have belief in it, I wouldn't say: "No. I don't believe there will be such a thing." It would seem to me utterly crazy to say this. [...]
In one sense, I understand all he says—the English words "God", "separate", etc. I understand. I could say: "I don't believe in this," and this would be true, meaning I haven't got these thoughts or anything that hangs together with them. But not that I could contradict the thing.
You might say: "Well, if you can't contradict him, that means you don't understand him. If you did understand him, then you might." That again is Greek to me. My normal technique of language leaves me. I don't know whether to say they understand one another or not (Wittgenstein 1966: 55).

Allo stesso modo, secondo Putnam non possiamo davvero dare un senso a un'ipotesi come quella dei CIV perché «I haven't got these thoughts or anything that hangs together with them»: l'ipotesi stessa appare *inintelligibile*. È quindi l'appello a questa caratteristica wittgensteiniana – implicitamente contenuta nella nozione di "quasi-necessità" sviluppata da Putnam a partire dagli anni Sessanta e via via perfezionata – che si aggiunge all'argomentazione degli anni Ottanta basata sul concetto di significato e, insieme a essa, decreta lo sgretolamento dell'ipotesi dei CIV. Perché, se le cose stanno così, allora «I cannot express or understand the statement that I am always a brain in a vat without presupposing substantive beliefs and principles from which I may infer that I am not always a brain in a vat. I therefore conclude that the statement is self-undermining» (Ebbs 2016: 35). Che è proprio la conclusione tratta da Putnam negli anni Ottanta:

Sometimes a thesis is called 'self-refuting' if it is *the supposition that the thesis is entertained or enunciated* that implies its falsity. For example, 'I do not exist' is self-refuting if thought by *me* (for any '*me*'). So one can be certain that one oneself exists, if one thinks about it (as Descartes argued).
What I shall show is that the supposition that we are brains in a vat has just this property. If we can consider whether it is true or false, then it is not true (I shall show). Hence it is not true (Putnam 1981: 7-8).

Se perciò guardiamo all'ipotesi dei CIV dal punto di vista del nostro schema concettuale complessivo e delle nostre reali pratiche di indagine, e se ci convinciamo che cercare di rispondere al sostenitore dell'ipotesi dei CIV con una dimostrazione diretta significa cercare l'impossibile, possiamo renderci conto non solo che criticare la dimostrazione di Putnam in quanto rappresentante una petizione di principio è fuori luogo, ma anche che la sua dimostrazione è un bel modo di mettere a tacere quel sostenitore.

5. La morale conclusiva

Luigi Perissinotto ha osservato acutamente (in un dibattito) che considerare il ruolo delle verità quasi-necessarie dietro alla dimostrazione presentata da Putnam negli anni

Ottanta rende ridondante, e quindi tutto sommato superflua, la dimostrazione stessa. Sono d'accordo. Per escludere la possibilità dei CIV basterebbe far notare come essa confligge con le verità quasi-necessarie alla base del nostro sistema concettuale complessivo, e come quindi rappresenti un'impossibilità concettuale – almeno fino a quando gli elementi fondamentali di tale sistema concettuale rimangono invariati. In altre parole, per criticare l'ipotesi dei CIV basta un opportuno argomento epistemologico e metafisico – basta cioè la filosofia, i cui argomenti sono però notoriamente rivedibili –, mentre la logica può al più aiutare a rendere rigorosi i passaggi del ragionamento – non a conferire a quest'ultimo il carattere della dimostrazione logico-matematica inconfutabile (se corretta).

Perché dunque Putnam ha negli anni Ottanta presentato la dimostrazione contro l'ipotesi dei CIV, visto che già da una ventina d'anni gli era evidente il ruolo che le verità quasi-necessarie hanno nella nostra cognizione? Io credo che nel formulare la dimostrazione Putnam non vedesse in modo chiaro e consapevole il legame tra la sua dimostrazione e il ruolo delle verità quasi-necessarie – per quanto strano possa sembrare –, e che non lo vedesse perché la sua anima di matematico dimostratore di teoremi aveva preso il sopravvento su quella di filosofo. Questo può apparire ancora più strano, ma non si deve dimenticare che, da un lato, Putnam ha portato avanti la sua attività di filosofo e quella di matematico in modo parallelo per più di un trentennio, ottenendo notevoli risultati in entrambi i campi, e, dall'altro, che l'idea della dimostrazione contro l'ipotesi dei CIV gli era venuta ragionando sul teorema di Skolem-Löwenheim, come abbiamo visto sopra, dunque in un contesto logico-matematico.

Non solo. Mi spingerei fino a dire che una chiara consapevolezza del legame tra dimostrazione e ruolo delle verità quasi-necessarie Putnam non l'abbia *mai* avuta, dando luogo a una oscillazione tra i due poli del legame – oscillazione tra considerare la propria dimostrazione come un argomento che confuta definitivamente l'avversario, e considerarla invece opinabile (a torto o a ragione). Questo si può evincere da una lezione tenuta appena tre anni prima della sua morte, e pubblicata poi con il titolo *Sixty-Five Years of Philosophy*², in cui egli ritorna sulla sua dimostrazione dichiarando che

The purpose of this argument was to show that a scientifically up-to-date and less solipsistic version of the Cartesian claim that there might be no external world at all, because I might be a disembodied mind deceived by an evil demon, could be *refuted* (Putnam 2016: 219; corsivo mio),

ossia eliminata una volta per tutte grazie a una dimostrazione diretta e schiacciante del tipo di una dimostrazione logica corretta, per poi affermare alla fine della propria ricognizione dell'argomento che quest'ultimo

is not meant to convince a skeptic. Skeptics can always reject the premise; they can reject semantic externalism. On the other hand, if the skeptic assumes that semantic externalism is false, the skeptic assumes something I do not believe and do not have to believe. Thus, even if the argument does not show that Cartesian skepticism is wrong, it does show that Descartes, or his imagined contemporary successor who talks about Brains in a Vat rather than disembodied minds, does not have a proof I must accept that skepticism is *right*, or at least not a proof based on the possibility that we are all Brains in a Vat, because I don't have to accept the possibility that *I* am a Brain in a Vat (Putnam 2016: 221),

² Ringrazio uno dei due referee per aver richiamato la mia attenzione su questo scritto.

dove, da un lato, è chiaro che “io” non devo accettare questa possibilità perché cozza con le assunzioni centrali del mio sistema concettuale, e, dall’altro, che la dimostrazione non viene intesa come un argomento “diretto” e volto a confutare l’avversario una volta per tutte, ma più debolmente a “convincere” – a convincere un avversario che non si convincerà mai³.

Una morale appare chiara da tutto ciò, ed è che, quando le dimostrazioni logiche investono la nostra dimensione metafisica ed epistemologica, non sono solo questioni di applicare correttamente le regole della logica. E in questo, di nuovo, l’affinità tra gli elementi centrali del pensiero del secondo Wittgenstein e quelli del pensiero di Putnam è notevole, come il seguente passo che Wittgenstein scrisse a pochi giorni dalla morte mostra:

«Ma anche se in casi del genere non posso sbagliarmi – non è possibile che io sia sotto narcosi?» Se sono sotto narcosi, e se la narcosi mi priva della mia coscienza, allora in questo momento non sto realmente *parlando* e *pensando*. Non posso seriamente supporre che in questo momento sto sognando. Chi, sognando, dice: «Io sogno», anche se parlasse in modo da essere udito, non avrebbe più ragione di quanta non ne avrebbe se, in sogno, dicesse «piove», mentre piove davvero. Anche se il suo sogno fosse realmente connesso con il rumore della pioggia (Wittgenstein 1969: § 676; corsivo mio).

Una meravigliosa salutare sintonia⁴.

Bibliografia

Alai, Mario (1994), «Brains in Vat and Their Minds: A Wrong Impossibility Proof» in *European Review of Philosophy*, 1, pp. 3-18.

Bernecker, Sven (2016), *Extended Minds in Vats*, in Goldberg, Sanford C., edited by (2016), *The Brain in a Vat*, Cambridge, Cambridge University Press, pp. 54-72.

Coliva, Annalisa (2022), *Wittgenstein Rebinged: The Relevance of On Certainty for Contemporary Epistemology*, London, Anthem Press.

³ Una non chiara consapevolezza dei limiti dell’uso dei risultati della logica a fini filosofici traspare anche dal modo in cui Putnam spiega come mai riusciamo a fare sì che le parole che usiamo si riferiscano agli oggetti a cui si presuppone si riferiscano. Ci riusciamo, abbiamo visto nel §3, grazie al *principio della devirgolettatura* (*disquotation principle*) di Tarski (ossia la famosa *Convenzione V*): «in my language, the word “cabbage” refers to cabbages. How do I know that? There is a fundamental logical principle, basically stated by Tarski, according to which *in one’s own language* any term *T* refers to, precisely, *Ts*» (Putnam 2016: 220). Tuttavia, il fatto è che il principio della devirgolettatura non è che “logico”, e deve perciò essere associato a una posizione metafisica specifica (e a un resoconto causale). *Da solo* non funziona come Putnam vorrebbe. Uno spirito disincarnato potrebbe infatti fornire esempi perfettamente corretti del principio di devirgolettatura, ma difficilmente diremmo che si riferisce a qualcosa, giacché mancherebbe di connessioni causali.

⁴ I commenti di entrambi i referees sono stati per me molto utili: li ringrazio molto.

Ebbs, Gary (2016), *How to Think about Whether We Are Brains in Vats*, in Goldberg, Sanford C., edited by (2016), *The Brain in a Vat*, Cambridge, Cambridge University Press, pp. 27-36.

Dell'Utri, Massimo (1990), «Choosing Conceptions of Realism: The Case of the Brains in a Vat» in *Mind*, 99, pp. 79–90.

Hanna, Robert (2011), «Minding the Body» in *Philosophical Topics*, 39, pp. 15-40.

Macarthur, David (2019), *Pragmatic Naturalism: The Authority of Reason, the Agrippan Trilemma and the Significance of Philosophising in medias res*, in Giladi, Paul, edited by (2019), *Responses to Naturalism: Critical Perspectives From Idealism and Pragmatism*, New York, Routledge, pp. 271-289.

Moyal-Sharrock, Danièle (2004), *Understanding Wittgenstein's On Certainty*, New York, Palgrave Macmillan.

Moyal-Sharrock, Danièle (2021), *Certainty in Action: Wittgenstein on Language, Mind and Epistemology*, London, Bloomsbury Academic.

Nagel, Thomas (1986), *The View from Nowhere*, Oxford, Oxford University Press.

Peirce, Charles S. (1955), *Some Consequences of Four Incapacities*, in Buchler, Justus, edited by (1955), *Philosophical Writings of Peirce*, New York, Dover, pp. 228-250.

Perissinotto, Luigi (2010), “...to begin at the beginning”: *The Grammar of Doubt in Wittgenstein's On Certainty*, in Perissinotto, Luigi & Sanf elix, Vicente, edited by (2010), *Doubt, Ethics and Religion: Wittgenstein and the Counter-Enlightenment*, Frankfurt a.M., Ontos Verlag, pp. 151-178.

Perissinotto, Luigi (2016), «Miracles, Hinges, and Grammar in Wittgenstein's On Certainty» in *International Journal for the Study of Skepticism*, 6, pp. 143-164.

Putnam, Hilary (1978), *Meaning and the Moral Sciences*, London, Routledge & Kegan Paul.

Putnam, Hilary (1981), *Reason, Truth and History*, Cambridge, Cambridge University Press.

Putnam, Hilary (1994), *Rethinking Mathematical Necessity*, in Putnam, Hilary (1994), *Words and Life*, Conant, James, edited by, Cambridge, MA, Harvard University Press, pp. 245-263.

Putnam, Hilary (2012), *Wittgenstein: A Reappraisal*, in Putnam, Hilary (2012), *Philosophy in an Age of Science: Physics, Mathematics, and Skepticism*, De Caro, Mario & Macarthur, David edited by, Cambridge, MA, Harvard University Press, pp. 482-492.

Putnam, Hilary (2016), *Sixty-Five Years of Philosophy: A Participant's Thoughts and Experiences*, in Putnam, Hilary (2016), *Naturalism, Realism, and Normativity*, De Caro, Mario, edited by, Cambridge, MA, Harvard University Press, pp. 213-226.

Thorpe, Joshua R. & Wright, Crispin (2022), *Putnam's Proof Revisited*, in Conant, James e Chakraborty, Sanjit, edited by (2022), *Engaging Putnam*, Berlin, de Gruyter, pp. 63-88.

Williams, Michael (2000), «Epistemic Entitlement», in *Philosophy and Phenomenological Research*, 60, pp. 607-612.

Wittgenstein, Ludwig (1953), *Philosophische Untersuchungen*, Oxford, Basil Blackwell (*Ricerche filosofiche*, trad. it. di R. Piovesan e M. Trinchero, Torino, Einaudi 1967).

Wittgenstein, Ludwig (1966), *Lectures and Conversations on Aesthetics, Psychology, and Religious Belief*, Barrett, Cyril, edited by, Berkeley, University of California Press.

Wittgenstein, Ludwig (1969), *On Certainty*, Oxford, Basil Blackwell (*Della Certezza*, trad. it. di M. Trinchero, Torino, Einaudi 1978).