

QUADERNS DE FILOSOFIA VOL. II NÚM. 2 (2015): 57-88

ISSN: 2341-1414 eISSN: 2341-3042 DOI: 10.7203/QFIA.2.2.7083

JORDI VALOR

Universitat de València

Les paradoxes i la filosofia: tres visions contemporànies

Rebut: 23/9/15. Acceptat: 28/9/15

Resum: Habitualment les paradoxes són caracteritzades com a enunciats aparentment falsos recolzats per arguments aparentment bons que parteixen de premisses que la majoria de la gent consideraria trivialment vertaderes. Aquest article ofereix una visió panoràmica de tres perspectives contemporànies de les paradoxes. Segons la concepció epistèmica, les paradoxes juguen un paper crucial en el desenvolupament de la ciència i mai no constitueixen proves correctes d'enunciats falsos. D'acord amb la concepció dialetheista, la conclusió d'alguns raonaments paradoxals és alhora vertadera i falsa, havent-hi, per tant, algunes contradiccions vertaderes (i falses). Finalment, la concepció mística entén les paradoxes recalitrants (antinòmies) com a problemes irresolubles que mostren els límits del pensament humà. La darrera secció de l'article discuteix les paradoxes d'autoreferència, una família de paradoxes que comparteixen certs trets estructurals i que han tingut un impacte profund en el desenvolupament de la ciència al llarg del segle xx.

Abstract: A paradox is usually described as an apparently false statement supported by an apparently good argument which departs from premises that most people would find trivially true. This survey article presents a brief overview of three different contemporary perspectives on paradoxes. According to the epistemic view, paradoxes play a crucial role in the progress of science and cannot be regarded as sound proofs of a false statement. According to the dialetheist view, the conclusion of some paradoxical reasonings is both, true and false, what means that there must be some true (and false) contradictions. Finally, the mystic view understands recalitrant paradoxes (antinomies) as unsolvable problems which show the limits of human thought and cognition. The last section of this paper focuses on the paradoxes

of self-reference, a family of paradoxes (sharing some structural features) which have had a deep impact in the development of science during the twentieth century.

Paraules clau: paradoxa, paradoxa verídica, paradoxa falsídica, antinòmia, dialetheisme, autoreferència, circularitat, *inclosure*.

Keywords: paradox, veridic paradox, falsidic paradox, antinomy, dialetheism, self-reference, circularity, inclosure schema.

I. LA NOCIÓ DE PARADOXA

SI HOM CERCA EL MOT 'PARADOXA' en un diccionari, hi trobarà que la principal accepció del terme —molt pròxima al seu origen etimològic— descriu una paradoxa com una idea o opinió que sentim espontàniament com quelcom inacceptable o contrari a l'opinió comuna (deixant oberta la possibilitat que l'opinió correcta siga la paradoxal, i no pas la que ens resulta més natural acceptar). Tanmateix, aquesta caracterització del terme no fa justícia a les paradoxes. Qui opine que algú sense diners és ric, o que $2 = 1$, o que els gossos són boles de sucre voladores, tindrà certament una creença contrària a l'opinió comuna, però no trobem cap paradoxa en el fet que algú crega quelcom fals. Com va observar Quine (1966, 3), les paradoxes estan estretament lligades a un raonament: una paradoxa és una tesi o una opinió que trobem inicialment absurda però que ve avalada per un argument aparentment acceptable del qual és la conclusió. Un exemple il·lustrarà millor aquesta idea.

És una trivialitat dir que algú sense diners és pobre i que un multimilionari és ric. Per altra banda, és obvi que un ric no deixarà de ser-ho pel mer fet de perdre un euro. En general, pensem que el següent principi és correcte:

(PT_R) Si algú amb n euros és ric, algú amb $n-1$ (o $n+1$) euros també ho és.

Crispin Wright (1976, 156) descriu un terme F com a *tolerant* sempre que l'aplicació de F a un objecte x tolere canvis graduals en x : canvis petits però quantificables que afecten alguna cosa la possessió o no de la qual (en major o menor grau) és determinant per tal que x siga F . 'Calb' —com 'ric', 'pobre', etc.— és un predicat tolerant: la seua aplicació a un individu depèn de quants cabells tinga, i tolera ser aplicat a algú amb un cabell més (o menys) per exemple que un individu calb. Sembla preceptiu pensar que tots els termes vagues són tolerants en aquest sentit i que respecten principis de tolerància semblants a PT_R, altrament seria impossible distingir un predicat vague com 'ser ric' d'un predicat precís com 'tenir cent milions d'euros'. El primer es caracteritza

justament per la manca d'un límit clar a partir del qual deixa de (o comença a) aplicar-se a un objecte: no hi ha un nombre determinat de cabells a partir del qual hom esdevé calb (o perd tal condició), com tampoc no hi ha un nombre determinat de grans de sorra a partir del qual tenim una pila de grans o deixem de tenir-la. Hi ha objectes que clarament admeten un predicat vague com 'ric' i n'hi ha d'altres que clarament no l'admeten. També hi ha objectes per als quals és indeterminat si el predicat s'aplica o no (¿és ric qui té vuit cents mil euros?) però cal observar que, donat un terme vague F , no hi ha un límit clar entre els objectes que són F , els que no ho són i els que no està determinat si ho són o no. Si hi hagués un, podríem convertir F fàcilment en un predicat precís identificant-lo amb el predicat 'ser determinadament F ' (DF). Però DF és un predicat tan vague (i tolerant) com el propi F : si algú amb n euros és determinadament ric, també ho és algú amb $n-1$ euros. I el mateix es pot dir de DDF , $DDDF$, etc. (un fenomen que alguns anomenen "vaguetat d'ordre superior").

Arribats a aquest punt, però, tenim tots els ingredients que permeten articular la paradoxa de *sorites* (*munt* o *pila* en grec), una paradoxa o família de paradoxes ben conegudes en l'Antiguitat que afecten a qualsevol terme vague. Una d'aquestes paradoxes diu així: És clar que algú amb una fortuna de 100 milions d'euros és ric. Ara bé, aquest fet, unit a PT_R , permet inferir, apel·lant al principi d'inferència anomenat *Modus Ponens* (MP), que algú amb 99.999.999 € és ric. Apel·lant de nou a PT_R i a MP , inferim ara que algú amb 99.999.998 € és ric i seguint la mateixa rutina una i altra vegada podem establir —després d'aplicar 100 milions de vegades MP — que algú amb 0 euros és ric! Allò que abans havíem descrit com quelcom absurd: creure que algú sense diners és ric, esdevé ara la conclusió d'un argument que té com a premisses una obvietat —que un multimilionari és ric— i un principi, PT_R , que no podem abandonar si 'ric' és, com sembla, un predicat vague. Per altra banda, l'argument descansa en la iteració de MP , una regla d'inferència deductiva irreprotxable: "Si és el cas que p implica q i és el cas que p , llavors és el cas que q ". Sorprenentment, tenim una raó per creure quelcom manifestament fals: que tal falsedat és la conclusió d'un argument aparentment correcte que parteix de veritats trivials i principis *a priori* que formen part de la nostra comprensió del significat dels termes involucrats. És per això que les paradoxes copsen la nostra atenció, no perquè avancen tesis falses o absurdes, sinó perquè —com un prestidigitador que sembla materialitzar un objecte davant dels nostres ulls incrèduls— una paradoxa ens sorprèn donant-nos una raó poderosa per a creure una tesi absurda: un bon argument.¹

¹ Keefe i Smith (1997) ofereixen una excel·lent introducció als problemes i paradoxes relacionades amb la noció de vaguetat. Vegeu també Sainsbury (2009, capítol 3).

2. ¿PER QUÈ LES PARADOXES SÓN UN PROBLEMA?

Però, ¿és realment bo l'argument que dóna suport a una paradoxa? Si ho és, tenim un problema més greu del que puga semblar inicialment. Per això molts filòsofs des d'Aristòtil als nostres dies han mirat de prendre's seriosament les paradoxes. Parlant de la paradoxa del mentider —presentada en 3.3— Alfred Tarski diu:

Al meu parer, seria del tot erroni i perillós des del punt de vista del progrés científic menysprear la importància d'aquesta i d'altres antinòmies i tractar-les com a bromes o sofisteries. És un fet que som davant d'un absurd, que ens hem vist obligats a asseverar una oració falsa [...]. Si ens prenem el nostre treball seriosament, no ens podem resignar a aquest fet. Cal que en descobrim la causa, ço és, cal que analitzem les premisses sobre les quals descansa l'antinòmia; cal després que rebutgem almenys una d'aquestes premisses; i cal que investiguem les conseqüències que açò té per a tot el domini de la nostra recerca (TARSKI 1944, 348)

Qualsevol tesi paradoxal és la conclusió d'un raonament i els raonaments —com la percepció, la memòria o el testimoni— són un dels mecanismes o dispositius clàssics emprats en la formació i justificació de creences. L'objectiu de raonar és ampliar el nostre coneixement del món i la nostra comprensió de les conseqüències implícites que té allò que creiem o suposem. Quan raonem cerquem descobrir quina informació podem obtenir a partir d'allò que creiem saber o d'allò que suposem, i ho fem construint arguments que justifiquen la veritat d'una proposició (la conclusió) a partir d'altres (les premisses) que l'argument tracta com a vertaderes. Si els raonaments vàlids no preservaren la veritat de les premisses en la conclusió, raonar no resultaria una activitat cognitiva valuosa, no compliria la seua funció essencial: proporcionar coneixement. És cert que en un raonament inductiu la veritat de les premisses no assegura la veritat de la conclusió, però els arguments inductius són tan bons (o dolents) com alta o baixa siga la probabilitat que la conclusió siga vertadera quan les premisses ho són, i un argument inductiu amb premisses vertaderes i conclusió falsa és *sempre* inacceptable, sense importar com de bo ens puga semblar. La raó d'això és precisament que cap argument ha d'anar mai d'allò que és vertader a allò que no ho és. Com que les paradoxes incompleixen aquesta màxima i, no obstant això, semblen descansar en arguments perfectament correctes, plantegen un dilema: o revisem l'argument —rebutjant bé la validesa d'alguna inferència, bé la veritat d'alguna premissa, bé la falsedat de la conclusió—, o acceptem que hi ha arguments *vàlids* que proven la veritat o acceptabilitat d'una falsedat.

Apreciarem millor fins a quin punt és perillós optar per la segona banya del dilema si parem atenció a un principi, *Ex Falso Quodlibet (EFQ)*, que és respectat per tota lògica consistent. *EFQ* descansa en la idea que acceptar una falsedat és sempre absurd. Segons *EFQ*, d'una falsedat se segueix *qualsevol cosa*: “Si \perp és el cas, llavors q ”.² Si li atorguem a $2 = 1$ (per exemple) el mateix estatus que conferim a una veritat, *EFQ* diu que podem atorgar tal estatus a qualsevol proposició q , per absurda que siga. Per tant, si demostràrem que $2 = 1$, *EFQ* ens permetria inferir *totes* les proposicions. Arribats a aquest punt, la pràctica de raonar esdevindria ociosa perquè permetria justificar qualsevol cosa. Tota proposició seria trivialment vertadera, de fet seria inintel·ligible traçar una distinció entre veritat i falsedat. És per això que *EFQ* rep també el nom *d'Explosió*: una teoria que accepte *EFQ* i que permeta derivar una falsedat es col·lapsarà automàticament, esdevindrà trivial en el sentit d'acceptar-ho tot, de no destriar o separar les proposicions acceptables de la resta.

Però per què acceptar $2 = 1$ hauria d'obligar-nos a acceptar que els gosos són boles de sucre voladores? ¿No podríem prescindir d'*EFQ*? És cert que apel·lem tàcitament a *EFQ* quan diem coses tals com: “Si tu ets francès, jo sóc Napoleó” (convençuts que l'antecedent és fals), però *EFQ* sembla menys intuïtiu que principis d'inferència com ara *MP*. No obstant això, les aparences enganyen. La validesa d'*EFQ* pot justificar-se a partir d'altres principis fonamentals. Considerem-ne dos als quals anomenaré *Introducció de la Disjunció (ID)* i *Sil·logisme Disjuntiu (SD)*. *ID* diu simplement que “si és el cas que p , llavors és el cas que $p \vee q$ ”; i *SD* diu que “si $p \vee q$ és vertadera i no és el cas que p , llavors és el cas que q ”. Aquests principis guarden relació amb la nostra comprensió del significat de la disjunció i difícilment podríem renunciar a ells sense transformar per complet el significat de la paraula ‘o’ (entre d'altres) i la nostra forma de raonar. Tanmateix, aquests principis conjuntament permeten demostrar que si acceptem una falsedat, podem acceptar la veritat d'una proposició qualsevol q . Suposem que p és una falsedat acceptable. Com que p és falsa, la seua negació, $\neg p$,³ és vertadera i, per tant, igualment acceptable. Ara bé, a partir de p , podem inferir $p \vee q$ (per *ID*), on q és una proposició qualsevol; i a partir de $\neg p$ i $p \vee q$, podem inferir q (per *SD*). Per tant, si p és falsa, podem inferir q , com resa *EFQ*, i sols podem rebutjar aquest principi rebutjant també la validesa general d'*ID* o de *SD*, quelcom aparentment inacceptable. Per altra banda, moltes lògiques (entre elles la clàssica) tracten oracions de la forma ‘Si p , llavors q ’ ($p \rightarrow q$) com

² El símbol ‘ \perp ’ representa una falsedat lògica i equival a una contradicció.

³ Representaré la negació, la conjunció, la disjunció, el condicional i el bicondicional mitjançant els símbols ‘ \neg ’, ‘ \wedge ’, ‘ \vee ’, ‘ \rightarrow ’ i ‘ \leftrightarrow ’, respectivament, i usaré, com he fet fins ara, variables del tipus p , q , r , etc. per a representar proposicions.

a condicionals *materials* i deriven al seu si l'equivalència entre $p \rightarrow q$ i $\neg p \vee q$. Això vol dir que si el condicional '→' que apareix en "Si és el cas que $p \rightarrow q$ i és el cas que p , llavors és el cas que q " és material, *MP* és equivalent a la següent versió de *SD*: "Si és el cas que $\neg p \vee q$ i és el cas que p , llavors és el cas que q ". Així, renunciar a *SD* suposa en moltes lògiques renunciar a *MP*.

Un altre principi absolutament fonamental al qual caldria renunciar si acceptem la segona banya del dilema anterior és l'anomenat *Reductio ad Absurdum* (*RA*): "Si podem inferir una falsedat a partir del supòsit que la proposició p és vertadera, podem inferir la falsedat de p (o, equivalentment, la veritat de $\neg p$)". Aquest principi descansa en la intuïció que inferir —i acceptar consegüentment— una falsedat (\perp) a partir d'una veritat és absurd. Per tant, si de p se segueix \perp , p ha de ser rebutjada. Una vegada acceptem la correcció de demostrar falsedats, la motivació de *RA* s'esvaeix.⁴

Davant dels perills que ens sotgen, si acceptem la segona banya del dilema al qual ens aboquen les paradoxes —el trivialisme o la renúncia a *EFQ*, *RA* i un bon grapat d'altres principis fonamentals—, no és d'estranyar que la immensa majoria de filòsofs opten per revisar l'argument paradoxal. El problema rau en el fet que aquesta opció sol imposar en cada cas sacrificis gairebé tan dolorosos com els considerats fins ara i, sovint, tampoc no assegura l'eradicació d'alguna versió més subtil de la paradoxa en qüestió, o fins i tot la generació de noves paradoxes.

3. TRES IMATGES DE LA RELLEVÀNCIA TEÒRICA DE LES PARADOXES

3.1. La concepció doxàstica

Qui rebutja la possibilitat de falsedats demostrables abraça la primera banya del dilema anterior i té obertes, per tant, tres opcions quan revisa un argument paradoxal. Segons Sainsbury (2009, 1), una paradoxa és:

una conclusió aparentment inacceptable derivada mitjançant raonaments aparentment acceptables a partir de premisses aparentment acceptables. Les aparences han de ser enganyoses, ja que allò acceptable no pot portar a allò inacceptable a través de passos acceptables. Així, en general, cal que triem: o la conclusió

⁴ *RA* permet a més justificar *EFQ* en companyia de *MP* (i del principi clàssic, rebutjat pels intuïcionistes, de *Doble Negació* (*DN*): " p i $\neg\neg p$ són equivalents"). Partint com a premisses de p i de la seua falsedat: $p \rightarrow \perp$, suposem $\neg q$. Inferim després \perp de les premisses (*MP*) i reduïm el supòsit a l'absurd: $\neg\neg q$ (*RA*).

no és realment inacceptable, o altrament el punt de partida, o el raonament, contenen alguna errada que no és pas òbvia

Podríem dir que Sainsbury té una visió *doxàstica* de les paradoxes. D'acord amb aquesta concepció —sens dubte la dominant, compartida per Tarski i Quine—, en una paradoxa les coses no ens semblen mai com de fet són: tota paradoxa amaga en última instància una il·lusió que ens fa acceptar una opinió errònia bé sobre la correcció d'un argument, bé sobre la veritat de les seues premisses, bé sobre la falsedat de la conclusió. La nostra comesa és identificar i corregir l'error.

Quine (1966, 5-6), per exemple, distingeix entre paradoxes verídiques i falsàries. Les primeres són aquelles on la conclusió, que sembla falsa, es revela en última instància com a vertadera mentre que les segones tenen, com sembla, una conclusió falsa i allò que cal revisar són les premisses o les inferències de l'argument paradoxal. En principi, jutgem fals que no hi pugua haver en una aldea un barber que afaite tots els aldeans (i sols a ells) que no s'afaiten a si mateixos. Tanmateix, el següent argument mostra que ens enganyem: Suposem que A fos tal barber. O A s'afaita a si mateix o no ho fa. En el primer cas, A afaitarà algú que s'afaita a si mateix; en el segon, deixarà sense afaitar algú que no s'afaita a si mateix. Per tant, cap aldeà pot ser el barber de tots els aldeans que no s'afaiten a si mateixos. La paradoxa del barber és una paradoxa verídica: l'argument avançat demostra concloentment que un enunciat aparentment fals és en realitat vertader. Entre les paradoxes falsàries Quine situa, per exemple, les paradoxes de Zenó.

La paradoxa de la cursa —relacionada amb la d'Aquil·les i la tortuga— afirma que un corredor A que ha de cobrir en una cursa una distància qualsevol, un kilòmetre per exemple, mai no arribarà ni tan sols a moure's. Aquesta falsedat òbvia s'argumenta de la següent manera: (i) Ningú no pot completar en un temps finit un nombre infinit de tasques; però (ii) per a completar la cursa i arribar al punt "1 km" (la meta), A ha de córrer abans un nombre infinit de "curses": ha de córrer fins al punt intermedi 1/2 km, i abans fins al punt 1/4 km, i abans fins al punt 1/8 km, etc. Per tant —i atès que el problema es reproduïx per a qualsevol distància x , per ínfima que siga— A no pot fer ni una passa. On falla l'argument? Suposant que assolir el punt 1/2 km, i abans el punt 1/4 km, etc., siguin efectivament subtasques presents en la tasca de córrer un kilòmetre, la premissa (i) de l'argument assumeix quelcom il·lícit, a saber, que la suma d'una sèrie infinita d'interval·ls finits de temps —els requerits per a completar cadascuna de les subtasques de la cursa— no pot ser mai un interval finit de temps. Si tots els interval·ls foren de la mateixa durada, això seria cert, però podem imaginar que el temps invertit en cobrir cada distància

minva indefinidament en proporció a la llargària de la distància. Així, assumint una velocitat constant i suposant que cada tasca es completa en un interval finit de temps —cada vegada més breu a mesura que s'acurta la distància—, podem imaginar que A arriba en t al punt $1/2$ km; en $t/2$ al punt $1/4$; en $t/4$ al punt $1/8$, etc. Ara bé, la matemàtica moderna ens diu que la suma de la sèrie infinita: $t/2 + t/4 + t/8 + \dots$ és t (el límit on convergeix la sèrie): l'interval finit de temps que invertim en recórrer la meitat d'un kilòmetre. Per tant, no hi ha cap contradicció a pensar que A assolirà la meta en $2 \cdot t$.⁵

Per a Quine, les paradoxes de Zenó il·lustren un fet crucial i és que les nostres conviccions sobre la validesa *a priori* de certs principis generals —com ara que tota suma d'un nombre infinit de magnituds ha de resultar en una magnitud infinita— canvien amb el temps i ho fan habitualment com a resposta a paradoxes que ens forcen a revisar tals conviccions. Quine (1966, 7) admet que hi ha unes paradoxes —les anomena *antinòmies*— que, a diferència de les verídiques i les falsàries, no tenen encara una solució clara (les paradoxes de sorites i del mentider serien dos bons exemples) però també pensa (*ibidem*, 11) que les úniques diferències que hi ha entre una antinòmia i una paradoxa verídica o falsària són l'estat del nostre coneixement i l'evolució dels nostres esquemes conceptuals. Les antinòmies del passat són les paradoxes verídiques o falsàries del present o del futur i qualsevol paradoxa pot deixar de ser-ho amb el pas del temps i esdevenir (si és falsària) una falsedat manifesta basada en un pèssim raonament o (si és verídica) una veritat central establerta per un argument concloent. Les tesis de Quine pressuposen òbviament la concepció doxàstica: *cap* paradoxa, per recalitrant que siga, pot justificar una falsedat, tot i que “les paradoxes admeten gradacions, depenent de com de bé les aparences camuffen la realitat” (SAINSBURY 2009, 1). Sempre hi ha quelcom que se'ns escapa i el nostre repte és descobrir què i esmenar la nostra visió del món. Les paradoxes representen així el combustible que alimenta el motor del progrés conceptual i epistemològic. Idees que en el passat semblaven estranyes o absurdes, com per exemple que un conjunt pogués tenir tants elements com algun dels seus subconjunts propis,⁶ són ara veritats fonamentals. De fet, definim així el tret característic de tot conjunt infinit i demostrem, per exemple, que hi ha tants nombres naturals com naturals parells (un subconjunt propi del primer conjunt) especificant una funció, com ara $2 \cdot n$, que associa amb cada membre n

⁵ Per a una discussió de les paradoxes de Zenó i altres relacionades vegeu Sainsbury (2009, capítol 1); Clark (2007, 1, 181); Salmon (1970 i 1980, capítol 2).

⁶ Un conjunt A és un *subconjunt* d'un altre B ($A \subseteq B$) si i sols si és el cas, per a tot x , que: si x pertany a A ($x \in A$), x pertany a B . A és a més un subconjunt *propi* de B ($A \subset B$) si i sols si $A \subseteq B$ i $A \neq B$. (El conjunt buit, \emptyset , és trivialment un subconjunt de tot conjunt Z , ja que “si $x \in \emptyset$, $x \in Z$ ” és un condicional material amb antecedent fals per a qualssevol x i Z .)

del primer conjunt: $\{0, 1, 2, \dots\}$, un (i sols un) del segon: $\{2 \cdot 0, 2 \cdot 1, 2 \cdot 2, \dots\}$, sense deixar-ne cap sense parella. Trobar una relació biunívoca (una funció bijectiva) que emparelle exhaustivament tots els membres de dos conjunts permet demostrar que ambdós tenen el mateix nombre d'elements i aquesta forma de comparar conjunts permet formular i respondre noves preguntes sobre la seua grandària. L'argument que justifica la conclusió paradoxal inspira així altres arguments similars que ens ajuden a eixamplar el nostre coneixement.⁷

Com que tot conjunt infinit té un nombre inacabable d'elements, podríem pensar ara que dos conjunts infinits qualssevol X i Y sempre tindran la mateixa grandària i que trobar una bijecció entre tots dos és una qüestió de temps i enginy. De nou, les aparences enganyen. El *Teorema de Cantor* demostra mitjançant un argument senzill que $\text{Sub}(X)$, el conjunt de tots els subconjunts d'un conjunt X , té sempre més elements que X . Açò és obvi quan parlem de conjunts finits no buits com $A = \{a, b\}$ perquè $\text{Sub}(A)$ també és finit: $\{\emptyset, \{a\}, \{b\}, \{a, b\}\}$ i més gran que A , però no ho és tant quan A és infinit i, consegüentment, també ho és $\text{Sub}(A)$. El Teorema de Cantor redueix a l'absurd (per a qualsevol conjunt A) la idea que A i $\text{Sub}(A)$ tinguen la mateixa grandària.⁸

Suposem que hi ha una bijecció f entre A i $\text{Sub}(A)$ —una relació biunívoca que emparella cada membre d' A amb un i sols un dels subconjunts d' A sense deixar-ne cap solt—. Podem ara fer-nos la següent pregunta en relació a cada $x \in A$ i al subconjunt d' A , $f(x)$, amb el qual l'emparella f : “pertany x a $f(x)$?”. En alguns casos potser sí i en altres no. Centrem-nos ara en els x que *no* pertanyen a $f(x)$: els $x \in A$ tals que $x \notin f(x)$. A té almenys un element a tal que $a \notin f(a)$;⁹ per tant hi ha un conjunt no buit, B , que consta dels $x \in A$ tals que $x \notin f(x)$. Com que B és un subconjunt d' A , $B \in \text{Sub}(A)$ i hi ha d'haver un element $b \in A$ tal que $f(b) = B$. Ara bé, o $b \in B$, o $b \notin B$. Si $b \in B$, per definició $b \notin f(b)$, però $f(b) = B$! Si $b \notin B$, llavors $b \notin f(b)$ però en tal cas b satisfà la definició de B i, per tant, $b \in B$! El supòsit d'una bijecció f entre A i $\text{Sub}(A)$ ens porta a una contradicció i podem reduir-lo a l'absurd. La no existència d'una bijecció entre A i $\text{Sub}(A)$ implica que un dels dos conjunts és més gran que l'altre. Com que amb cada $x \in A$ podríem emparellar el subconjunt $\{x\} \in \text{Sub}(A)$, sabem que hi ha almenys *tants* elements en $\text{Sub}(A)$ com en A . El Teorema de Cantor mostra consegüentment que $\text{Sub}(A)$ té *més* elements que A : tot intent d'emparellar A i $\text{Sub}(A)$ deixarà orfes alguns elements del segon conjunt.

⁷ Així sabem també que hi ha tants nombres sencers i tants nombres racionals com naturals.

⁸ Vegeu, per exemple, Quine (1966, 16) o Devlin (1993, 79, Lemma 3.6.6).

⁹ Sabem que $\emptyset \in \text{Sub}(A)$ i hem suposat que tots els membres de $\text{Sub}(A)$ estan emparellats mitjançant f amb un (i sols un) membre d' A . Per tant, hi ha un $a \in A$, tal que $f(a) = \emptyset$. Ara bé, $a \notin \emptyset$, per tant $a \notin f(a)$.

En realitat, Cantor va concebre una tècnica —anomenada *diagonalització* i il·lustrada per l'argument anterior— per a construir demostracions indirectes. Un argument diagonal típic parteix del supòsit que certa col·lecció Ω existeix. A continuació, mostra un element que ha de pertànyer a Ω però que no hi pot pertànyer, i usa aquesta contradicció per reduir a l'absurd l'existència de Ω . Un tret crucial de la tècnica és que l'element problemàtic de Ω , que anomenarem *diagonal* i representarem així: $\delta(\Omega)$, és definit apel·lant a tots els elements de la col·lecció Ω , entre els quals es troba el propi $\delta(\Omega)$.¹⁰ Cantor va usar per primera vegada els arguments diagonals per a establir una tesi sorprenent al seu moment: que no tots els conjunts infinits són igual de grans (A és més petit que $\text{Sub}(A)$, i aquest més petit que $\text{Sub}(\text{Sub}(A))$, etc.). Aquests arguments són avui dia una eina fonamental de la teoria de conjunts.

Quine (1966, 18-20) dóna encara un últim exemple de paradoxa verídica crucial en la nostra comprensió actual de les matemàtiques: el primer teorema d'incompletesa de Gödel. La tesi que hi ha veritats matemàtiques indemostrables sembla un disbarat; tanmateix, Kurt Gödel va demostrar en 1931 que, si un sistema deductiu contenia entre els seus teoremes *totes* les veritats de l'aritmètica elemental, llavors també podia demostrar una falsedat. O, equivalentment, donat un sistema deductiu qualsevol S : o S no pot demostrar totes les veritats de l'aritmètica elemental, o S pot demostrar una falsedat. Com que la segona possibilitat sembla inacceptable, s'ha interpretat tradicionalment aquest teorema com una prova que l'aritmètica és incompleta: conté veritats indemostrables. Una de les fites del teorema és la prova que, donat un sistema deductiu qualsevol S , hi ha una proposició de l'aritmètica, que anomenaré G , amb les mateixes condicions de veritat que la proposició: *G no és demostrable en- S* (ometré la referència a S d'ara en davant). Això planteja un problema seriós perquè, si G és demostrable, podem demostrar una falsedat (ja que G és falsa quan és demostrable); i, si G no és demostrable (com diu), hi ha una veritat aritmètica indemostrable.¹¹ Una vegada més, els resultats i tècniques

¹⁰ Simmons (1993, capítol 2) conté una exposició detallada d'aquesta tècnica. En el teorema de Cantor, Ω seria el conjunt de totes les parelles $\langle x, y \rangle$ que suposadament formen la bijecció f entre A i $\text{Sub}(A)$, on $x \in A$, $y \in \text{Sub}(A)$, i $f(x) = y$; $\delta(\Omega)$ seria $\langle b, B \rangle$, on B és el conjunt dels $x \in A$ tals que $x \notin f(x)$. (Observeu que determinar B requereix examinar, per a tots els $\langle x, y \rangle$ de Ω , si és el cas que $x \notin y$, en particular, si $b \notin B$.)

¹¹ Aquest teorema incorpora un argument diagonal on Ω seria la col·lecció de *totes* les veritats aritmètiques (emparellades amb les seues demostracions): Ω constaria dels $\langle p, \Delta_p \rangle$ tals que p és una veritat aritmètica i Δ_p el conjunt de les seues demostracions; $\delta(\Omega)$ seria $\langle G, \Delta_G \rangle$. Si $\delta(\Omega) \in \Omega$, G és demostrable però falsa; i si $\delta(\Omega) \notin \Omega$, G és indemostrable però vertadera. Per tant, Ω no pot existir. (Observeu que G es defineix apel·lant a tots els membres de Ω , ja que “no ser demostrable” és “ser diferent de tota proposició demostrable”: G és vertadera si, i sols si, *per a tota veritat aritmètica* p , si hi ha un Δ_p tal que $\langle p, \Delta_p \rangle \in \Omega$, llavors $G \neq p$.)

emprades per Gödel en aquest teorema van impulsar el desenvolupament de diverses branques de les matemàtiques, com ara la teoria de la demostració i l'estudi de les funcions recursives i la noció de computabilitat.

3.2. La concepció dialetheista

És innegable que moltes antinòmies es dissolen amb el temps i que esbrinar quina peça de la nostra concepció del món cal abandonar per tal de domesticar una paradoxa dóna peu a importants descobriments i revisions conceptuals que fan evolucionar tant el llenguatge amb el qual pensem el món com la nostra imatge d'ell. Els exemples històrics abunden, no sols en l'àmbit de les matemàtiques.

En el seu famós poema, Parmènides afirma que és impossible pensar en allò *que no és* (o dir res al respecte) precisament perquè *no és res*; i ens exhorta a no ser com “gents sense judici, que creuen que ser i no ser són el mateix i no el mateix”.¹² Sota la convicció que és absurd dir d'allò que és *que no és*, o d'allò que no és *que és*, Parmènides dedueix *a priori* la naturalesa d'allò que és, conclouent que és u, indivisible, immòbil, immutable, ingènit, imperible, sense principi ni fi... Negar alguna d'aquestes afirmacions i mantenir que hi ha una pluralitat d'éssers limitats, que canvien, es mouen, naixen i moren ens abocaria a dir d'allò que és, que no és, com fem en afirmar que “les balenes *són* mamífers, *no són* peixos” o “Joan *era* al despatx però ja *no hi és*”.

Avui dia, les conclusions de Parmènides resulten tan xocants que dir que ens semblen falses seria un eufemisme, però això és, en part, perquè hem après a pensar com aquells que “desfan l'argument de Zenó i Parmènides afirmant que allò *que és* i allò *u* es diuen de moltes maneres”.¹³ Aristòtil observa que:

‘allò que és’ [...] es diu en molts sentits: en primer lloc, està allò que és accidentalment; en segon lloc, allò que és en el sentit de ‘és vertader’ i allò que és en el sentit de ‘és fals’; a més, estan les figures de la predicació [*les categories aristotèliques*] per exemple què és, de quina qualitat, de quina quantitat, on, quan és, i qualsevol altra cosa que signifiqui d'aquest mode, i, a més de tots aquests sentits, allò que és en potència i en acte (*Metafísica*, VI, capítol 2, 1026a, 33-1026b, 1)¹⁴

¹² Text 293 en Kirk, Raven i Schofield (1983).

¹³ *Sobre les refutacions sofístiques*, 182b, 25-28, en Aristòtil (1982).

¹⁴ Seguesc la traducció al castellà del text publicada en Gredos (ARISTÒTIL 1994). Faig el mateix en els textos indicats en les notes 13 i 15.

L'argument de Parmènides és fal·laç perquè tracta com a unívoc un terme, 'ser', que té distints significats i articula conceptes i idees distintes. Així, la llavor, que *no és* un arbre en acte, ho *és* en potència; i res no hi ha d'estrany en el fet que quelcom que hi *era* allà, ja *no* hi *sigui* (o que *sigui* un mamífer i *no sigui* un peix). Només “és impossible que allò mateix es done i no es done alhora en allò mateix i en el mateix sentit”:¹⁵ no pot passar que la mateixa cosa sigui i no sigui un peix, o sigui i no sigui en tal lloc i moment, etc. Ens hem habituat a trobar tan familiars aquestes idees que hem oblidat que descobrir distints sentits amagats en una mateixa paraula sol ser una gesta intel·lectual important, com ho fou en aquest cas.

Però aquest exemple històric de superació d'una paradoxa per mitjà d'un exercici d'anàlisi conceptual pot obscurir el fet que paradoxes com les de sorites (i altres) ens acompanyen des de temps immemorials. Sovint, l'aparell conceptual que una teoria introdueix per a solucionar una paradoxa, tot i ser valuós, no evita la formulació d'una versió més sofisticada d'aquesta. Com indicàvem adés, la paradoxa de sorites no es resol introduint el concepte *determinadament* per a qualificar conceptes vagues com *ric* i diferenciar així els individus *determinadament rics* (o *no rics*) d'aquells que no ho són. El concepte proporciona millors instruments per a pensar el món però no resol la paradoxa perquè el terme 'determinadament ric' és vague i la reproduceix. Altres vegades la solució a una paradoxa en genera de noves (emparentades en major o menor grau amb l'anterior), o deixa per resoldre problemes relacionats amb la paradoxa inicial. Assumint la solució discutida a la paradoxa de la cursa, Thomson (1970) va plantejar una altra paradoxa.¹⁶ Per altra banda, i tornant al raonament de Parmènides, trobem una paradoxa larvada que no resol l'estratègia aristotèlica. Centrem-nos en el sentit del verb 'ser' que el fa equivalent a 'existir' en certs contextos: Parmènides diu plausiblement que sobre allò que no és, no hi podem dir o pensar res perquè no és. Però, en tal cas, sembla que *tampoc* podem dir *això*. No podem dir o pensar del que no és que no és, com fa Parmènides, perquè res no es pot pensar o predicar del que, per no ser res en absolut, ni tan sols pot ser referit o indicat. I, tanmateix, ¿no és això el que acabem de fer?¹⁷ Kirk, Raven i Schofield (1983, 355) observen que “[a]questa línia d'argumentació ha atret poderosament [...] a molts filòsofs, des de Plató a Russell. La seua conclusió és paradoxal” —i afegeixen— “com totes les bones paradoxes,

¹⁵ *Metafísica*, IV, capítol 3, 1005b, 18-21.

¹⁶ Sainsbury (2009, 12-5) i alguns articles recopilats en Salmon (1970) discuteixen en detall la paradoxa del llum de Thompson.

¹⁷ Més informació sobre la paradoxa dels existencials negatius en Sorensen (2003, capítol 3), que també descriu formes més sofisticades i vigents de les paradoxes de Parmènides.

ens forcen a examinar, amb major profunditat, la nostra comprensió dels conceptes emprats —particularment, en aquest cas, les relacions entre significat, referència i existència—. Ara bé, quines garanties tenim, després de més de dos mil anys, que una revisió dels nostres conceptes resoldrà finalment aquesta o altres paradoxes que ens intriguen des de l'antiguitat?

Graham Priest (2002; 2006) pensa que les revisions i canvis conceptuals motivats al llarg dels segles pel desig de dissoldre les inconsistències que generen certes paradoxes recalcitrants han propiciat noves paradoxes anàlogues a les anteriors o, alternativament, alguna mutilació intolerable de la capacitat expressiva del nostre llenguatge. Per això, Priest explora seriosament la possibilitat que les coses siguin de fet com semblen: que les premisses i la conclusió de certs raonaments paradoxals siguin respectivament vertaderes i falsa, i que l'argument siga, no obstant això, correcte. Com que cap argument correcte pot anar d'allò acceptable a allò inacceptable, Priest planteja la possibilitat que la conclusió d'una antinòmia, a més de falsa, siga *també* vertadera. Per definició, que p siga falsa significa que $\neg p$ és vertadera, per tant, dir que p és falsa i vertadera implica que p i $\neg p$ són ambdues vertaderes o, com diu Priest, que p és una *dialetheia*. Tota falsedat és rebutjable però tant les veritats com les dialetheies són acceptables i poden ser —a diferència d'una falsedat *simpliciter*— la conclusió legítima d'un argument vàlid que parteix de premisses acceptables (vertaderes o dialetheies). Cal observar que si p és una dialetheia, $\neg p$ i $p \wedge \neg p$ també ho són.¹⁸ Per tant, encara que tota contradicció és falsa per a Priest, algunes també són vertaderes i el *Principi de No Contradicció (PNC)*, interpretat com l'exclusió d'aquesta possibilitat, no és vàlid en general.¹⁹

Com veiem, la proposta de Priest és radical. Rebutja dràsticament algunes de les nostres conviccions més arrelades, com que cap contradicció pot ser vertadera o que les categories *vertader* i *fals* són mútuament excloents. Per altra banda, si bé és cert que Priest reinterpreta com a vertadera la conclusió p d'un argument paradoxal, no és menys cert que, sent una *dialetheia*, p continua sent falsa, i recordem que *EFQ* permet inferir qualsevol cosa d'una falsedat. ¿Ens aboca el dialetheisme al *trivialisme*, a l'afirmació que tota proposició és vertadera, una dialetheia?

Durant els darrers quaranta anys, Priest ha assumit amb vigor el difícil repte de defensar el dialetheisme. Per sintetitzar millor un seguit de raons estretament relacionades entre si, dividiré en tres blocs l'estratègia argumental de

¹⁸ $\neg p$ és vertadera perquè p és falsa; i falsa perquè p és vertadera. $p \wedge \neg p$ és vertadera perquè tant p com $\neg p$ ho són; i falsa perquè p o $\neg p$ (ambdues de fet) són falses.

¹⁹ El dialetheisme accepta *PNC* entès com l'afirmació que tota contradicció és falsa: “per a tot p , $\neg(p \wedge \neg p)$ ”, i el rebutja entès com l'afirmació que *cap* contradicció és vertadera.

Priest i parlaré d'una defensa *lògica, epistemològica i dialèctica* del dialetheisme. La defensa lògica s'adreça a la pregunta anterior i busca mostrar que el dialetheisme és una teoria coherent i lògicament raonable que no ens aboca al trivialisme. Priest (2006, 104) insisteix sovint que el fet que hi haja *algunes* falsedats (àdhuc contradiccions) racionalment acceptables, les dialetheies, *no* implica que *totes* ho siguin. Per tant, *EFQ* no pot ser un principi lògicament vàlid i Priest, consegüentment, rebutja la lògica clàssica, adoptant-ne una de paraconsistent: LP (“*logic of paradox*”, PRIEST 1979; 2006, capítol 5). Les lògiques paraconsistents són lògiques no clàssiques caracteritzades pel rebuig d'*EFQ*. Aquesta restricció permet, per exemple, simular com raonem en presència d'informació contradictòria —quelcom habitual— sense caure en el trivialisme. Com que Priest pensa que hi ha dialetheies *al món* (no sols informació inconsistent), LP contempla la possibilitat que hi haja proposicions que siguin alhora vertaderes i falses en virtut de com és el món (PRIEST 2002, 295). Així, partint de dos valors de veritat: vertader (1) i fals (0), LP permet *tres* possibles avaluacions d'una proposició: 1, 0 i ambdós.²⁰ Hem vist, però (punt 2), que la renúncia a *EFQ* té conseqüències greus. En LP són invàlids principis clàssics com: *EFQ*, *SD*, *MP* i *RA*, als quals cal afegir-hi, entre d'altres: *Antecedent Fals (AF)*: “Si $\neg p$ és el cas, $p \rightarrow q$ ”, *Contraposició (CTR)*: “Si $p \rightarrow q$, llavors $\neg q \rightarrow \neg p$ ” i *Transitivitat del Condicional (TR)*: “Si $p \rightarrow q$ i $q \rightarrow r$, llavors $p \rightarrow r$ ” (Vegeu PRIEST 1979 (III.9), 228, 230). A primera vista, açò suposa una mutilació intolerable del nostre raonament, però Priest (2006) fa dues observacions importants al respecte: (i) *tots* els principis clàssics —incloent-hi els que acabem d'enumerar— són vàlids en LP *quan no operem amb dialetheies*; (ii) el condicional material *no* copsa intuïtivament la nostra noció de condicional. Molts dels anteriors principis: *MP*, *CTR* o *TR*, deixen de fallar si busquem condicionals alternatius i millors. Per altra banda, principis com *AF* sols són plausibles si el condicional és *material*. Priest explora diferents nocions intensionals de condicional que respecten invariablement *MP* (alguns també *CTR*) i que s'acosten més a la nostra noció intuïtiva (*ibídem*, capítol 6). Paral·lelament, Priest manté que, tot i no ser estrictament vàlids, els anteriors principis són *quasi-vàlids* (*ibídem*, capítol 16). La probabilitat d'entropessar amb una dialetheia en la vida ordinària és ínfima, per tant, en condicions normals és raonable per a tothom

²⁰ Algunes lògiques paraconsistents consideren també la possibilitat que *p* no siga ni vertadera ni falsa, rebutjant el *Principi de Terç Exclòs (PTE)*: “ $p \vee \neg p$ ”. El dialetheisme és compatible amb aquest rebuig (PRIEST 2006, 66) però Priest respecta *PTE*: *vertader* i *fals* són categories exhaustives. Segons Priest (*ibídem*, 61-4), l'asserció és anàloga a un joc individual on no hi pot haver “empats”: l'objectiu és dir la veritat. Si fallem, allò asseverat és fals: si *no* hi ha *res* al món que justifique la veritat (falsedat) de *p*, llavors tal absència de proves o fets constitueix en si un *fet* —que denomina “negatiu” — a favor de la falsedat (veritat) de *p* (*ibídem*, 65-6, 267-8).

—*també* per al dialetheista— raonar d'acord amb *EFQ*, *SD*, *MP*, *RA*, etc. Sols cal suspendre aquests principis en presència d'una dialetheia (*ibidem*, 118). Dissortadament, no hi ha criteris *formals* per dirimir quan una proposició és una dialetheia. Establir que ho és requereix sospesar quines raons tenim a favor de la seua veritat i falsedat. Tanmateix, açò il·lustra un fet crucial per a Priest (1998, 423) i és que no hi ha criteris exclusivament formals per a determinar si una proposició és o no racionalment acceptable. Tenir la forma ' $p \wedge \neg p$ ' (o qualsevol altra), no determina que una proposició siga rebutjable: moltes dialetheies tenen aquesta forma. Per això, la coherència d'una teoria no pot identificar-se amb la seua consistència. La segona noció és formal: una teoria és consistent quan no deriva oracions de la forma ' $p \wedge \neg p$ ', la primera no. Per a Priest (1998, 419-24; 2006, 96-107), la consistència, tot i ser una condició rellevant per a la coherència, no és ni una condició necessària, ni suficient. No és suficient perquè, quan demanem coherència, no demanem sols l'exclusió de proposicions contradictòries: demanem coherència *amb l'experiència, amb els nostres principis bàsics*, etc. La coherència d'una teoria està relacionada amb la veritat i la justificació, qualitats que no garanteix la mera consistència: hi ha teories falses i del tot increïbles que són malgrat tot consistents. La consistència tampoc no és una condició necessària perquè tenim raons per a acceptar dialetheies segons Priest i perquè, de fet, mantenim raonablement un cert nombre de creences inconsistentes. Priest (2006, 100) dóna sovint com a exemple la paradoxa del pròleg (vegeu CLARK 2007, 166-8): algú pot tenir bones raons per a creure conjuntament i per separat totes les afirmacions: $\alpha_0, \alpha_1, \dots, \alpha_n$ que ha fet en un llibre després d'una investigació minuciosa. Tanmateix, la mateixa persona tindria poderoses raons empíriques independents per a creure que tot llibre, també el seu, conté alguna afirmació falsa (i advertir-ho així als lectors en el pròleg). Aquesta persona creuria alhora —conscientment o no— tesis contradictòries: ' $\alpha_0 \wedge \alpha_1 \wedge \dots \wedge \alpha_n$ ' i ' $\neg \alpha_0 \vee \neg \alpha_1 \vee \dots \vee \neg \alpha_n$ ' però la seua actitud no seria per això irracional. La coherència d'una teoria o d'un sistema de creences és una propietat més difícil de guanyar i de destruir que la mera consistència: és una qualitat de caràcter holista que demana que avaluem l'acceptabilitat d'una creença individual x en el context de la resta de les nostres creences, no sols en el d'aquelles que tinguen la forma ' $\neg x$ ' (si n'hi ha). La percepció d'un conflicte entre dues normes morals que acceptem, o entre dues teories (la relativitat i la mecànica quàntica, per exemple) no ens porta immediatament a rebutjar-ne una de les dues prescindint d'altres raons independents.

La defensa epistemològica del dialetheisme està encaminada a justificar positivament que *hi ha* dialetheies i que certes proposicions ho són. La manca en l'actualitat d'una solució consistent satisfactòria a algunes paradoxes no pot constituir mai una *demonstració* de l'existència de dialetheies, Priest n'és ben

conscient. De fet, res no hi podria comptar com a prova irrefutable que una contradicció és vertadera perquè sempre podem interpretar tal prova com una reducció a l'absurd d'algun concepte o principi lògic involucrat en ella i revisar consegüentment els nostres esquemes conceptuals i principis lògics. Constatàvem en 3.2 que així ha estat sovint al llarg de la història. ¿Com podríem demostrar que P , un principi bàsic, no és revisable? Una prova tal seria segurament tan discutible com el propi principi (o encara més, si en fes ús circularment). Ara bé, per raons idèntiques, tampoc podem demostrar que *no* hi ha dialetheies, o que *PNC* no és revisable. En aquest camp, només podem aspirar a construir arguments raonables no concloents i, en aquest sentit, l'existència de paradoxes recalitrants mil·lenàries (vegeu PRIEST 2002) pot oferir indicis empírics a favor de la hipòtesi dialetheista. Considerats aïlladament, aquests indicis són insuficients per a justificar tal hipòtesi, però fan augmentar la probabilitat de la seua veritat significativament si els combinem, com fa Priest, amb la tesi que el dialetheisme és lògicament coherent i raonable, i amb la tesi que les teories rivals són menys satisfactòries. Un argument directe a favor de l'existència de dialetheies que permet a més identificar-ne algunes, es basa en el fet que certs principis del raonament —tant bàsics com ho puguen ser *MP*, *RA* o *SD*— són inconsistents aplicats sense restriccions. Aquests principis lògics, expressats mitjançant els esquemes següents, justifiquen *a priori* l'acceptació de tota proposició que els exemplifiqui, i això vol dir que justifiquen l'acceptació d'algunes dialetheies:

- (PA) $x \in \{y : \varphi(y)\} \leftrightarrow \varphi(x)$.
 (PS) x satisfà ' $\varphi(y)$ ' $\leftrightarrow \varphi(x)$.
 (PV) x és vertadera $\leftrightarrow p$.

El *Principi d'Abstracció (PA)* diu que, associat amb tot concepte o condició $\varphi(y)$, hi ha el conjunt $\{y : \varphi(y)\}$ de totes les coses y que satisfan tal concepte: *Sòcrates* és filòsof si, i sols si, *Sòcrates* pertany al conjunt $\{y : y \text{ és filòsof}\}$ que té com a membres els filòsofs. El *Principi de Satisfacció (PS)* diu que x satisfà un concepte o condició $\varphi(y)$ quan és el cas que $\varphi(x)$, ço és, quan x (*Sòcrates*) és efectivament φ (filòsof). Per últim, El *Principi de la Veritat (PV)* diu que una proposició com *la proposició que expressa l'oració 'neva'* és vertadera si, i sols si, és el cas que *neva*. Certament, és difícil pensar en principis més bàsics que aquests i, tanmateix, són inconsistents. Per a Priest aquests principis tenen validesa universal. Per tant, si l'*única* forma d'evitar una contradicció a partir d'una proposició p és restringir-los, és millor no fer-ho i concloure que p és una dialetheia.

En 3.3 i 4, veurem exemples de dialetheies derivades assumint la validesa universal de *PA*, *PS* i *PV* ara ens centrarem, però, en una especialment

sorprenent. Si *PV* és un principi lògic universal aplicable en qualsevol demostració, podem reinterpretar el Teorema de Gödel com una prova que l'aritmètica és *inconsistent*. Gödel va provar que hi ha una proposició aritmètica amb les mateixes condicions de veritat que la proposició *G*: *G no és demostrable*. L'exemplificació escaient de *PV* estableix que *G* és vertadera si i sols si *G no és demostrable*. Ara bé, o *G* és demostrable o no ho és. Si *G* no és demostrable, llavors (per *MP* i *PV*) podem demostrar que *G* és vertadera. Però, en aquest cas (i aplicant de nou *MP* a *PV*), podem demostrar *G* (ço és: *G no és demostrable*), que és per tant demostrable. Però si *G* és demostrable, és vertadera —com tota proposició demostrable— i, aplicant de nou *MP* a *PV*, deduïm que *G* no és demostrable. En resum, *G* és demostrable si i sols si no ho és.²¹ Si *PV* és un principi lògic universalment vàlid, *G* és una dialetheia i l'aritmètica és inconsistent. El teorema de Gödel planteja un dilema: o l'aritmètica és consistent però incompleta (hi ha veritats aritmètiques indemostrables), o és completa però inconsistent (totes les veritats aritmètiques, i alguna falsedat, són demostrables). Com que *PV* permet demostrar *G* i la seua negació, Priest conclou que la segona banya del dilema és la correcta.²²

En la seua defensa dialèctica del dialetheisme, Priest (2006, 5) ens recorda que tota línia argumental contra la seua teoria que la descarte (o descarte algun dels principis a què dóna suport) per inconsistent, incorre en una *petitio principii*, ja que la possibilitat d'acceptar algunes contradiccions és precisament la tesi a debat. Dit açò, Priest adopta tres estratègies principals per a mostrar que les teories rivals són insatisfactòries (parlarem de l'última en el punt 4). Podem resumir-les així:

1. Per tal d'evitar resultats com l'anterior en relació amb el Teorema de Gödel, les solucions consistents a les paradoxes restringeixen principis del raonament (*PA*, *PS*, *PV*) no menys bàsics que els que restringeix el dialetheisme.

2. Les restriccions que imposen les teories rivals: o són *ad hoc* (i, per tant, injustificades filosòficament); o introdueixen conceptes que permeten generar noves paradoxes (i per tant, no són plenament consistents); o garanteixen la consistència sols al preu de fer sacrificis —més enllà dels tres principis esmentats— en els nostres recursos conceptuals i expressius que mutilen intolerablement el nostre pensament.

²¹ La prova assumeix la validesa de *MP* perquè el bicondicional de *PV no és material* segons Priest (2006, 59). Recordem que *MP* no és vàlid per al condicional material en LP.

²² Priest (2006, capítol 3) discuteix en detall aquestes qüestions. Hi ha models inconsistents i tanmateix no trivials de l'aritmètica on per a algun nombre natural *n* i el seu successor *n'*: $n = n'$ i $n \neq n'$ (capítol 17).

3. Analitzades correctament, certes paradoxes presenten un problema *comú* que requereix una solució uniforme i sols el dialetheisme ofereix tal solució.

Priest fa una defensa admirable del dialetheisme. Sens dubte, és el responsable del fet que, avui dia, filòsofs que no defensen totes les seues tesis, es declaren, no obstant això, dialetheistes.²³ Però ¿és el dialetheisme una solució a les paradoxes?

3.3. La concepció mística

Suposem que Priest ens convenç del fet que el dialetheisme no sols és una teoria raonable sinó també correcta sobre certes paradoxes recalcitrants. ¿No seria açò un descobriment important? Certament ho seria, i caldria reaccionar reestructurant els nostres esquemes conceptuals i principis lògics: restringint *PNC* i altres lleis clàssiques per fer lloc a un nombre moderat i coherent de contradiccions; adoptant el concepte *dialetheia* per pensar el món; revisant les nocions de *condicional*, *coherència*, etc. Hauríem acabat així amb les paradoxes?

En realitat no, la paradoxa de Curry (1942) ho demostra. Siga κ (per estipulació) l'oració condicional: *si κ és vertadera, plouen cigrons*. Suposem que κ és vertadera, en tal cas podem inferir (per *PV*) que: si κ és vertadera, plouen cigrons (ja que ' κ ' és un nom d'aquest condicional). Però això vol dir que tenim el condicional κ i el seu antecedent (κ és vertadera) i podem inferir (per *MP*) una falsedat ben grollera: que plouen cigrons. Suposem ara que κ no és vertadera. En tal cas, κ és un condicional amb un antecedent fals i (per *AF*) podem inferir que κ és vertadera. Ara bé, si κ és vertadera podem inferir (per *PV*) que: *si κ és vertadera, plouen cigrons* i tornem a tenir un condicional amb el seu antecedent... Per tant: *plouen cigrons!* Observeu que podríem haver escrit en el conseqüent de κ qualsevol oració: 'no plouen cigrons', 'Déu existeix', 'tot és vertader'... En resum, si ' x ' és un nom de l'oració 'si x és vertadera, p ', podem demostrar que p és el cas siga p la proposició que siga, i això vol dir que amb pocs recursos lògics ens veiem abocats al trivialisme.

Priest (2006, 6.2) n'és ben conscient del problema i el bloqueja. La demostració anterior és incorrecta en LP perquè fa ús d'*AF*. El bicondicional de *PV* no és material, és una connectiva intensional que no satisfà ni *AF* ni altres principis que poden alimentar la paradoxa (entre ells, *Absorció* o *Contracció* (ABs): "Si és el cas que $p \rightarrow (p \rightarrow q)$, llavors $p \rightarrow q$ "). Allò rellevant, però, és que la resposta de Priest no passa ací per assumir que la conclusió paradoxal, p , és tal que p i $\neg p$ són vertaderes. Acceptar la validesa del patró de raonament

²³ Beall i Armour-Garb (2003), Beall (2004), Mares (2004).

anterior no suposaria acceptar *alguna* dialetheia, suposaria acceptar que *tota* proposició és trivialment una dialetheia.

Cal dir que si les paradoxes són —com pensa Quine— el combustible que la ciència consumeix quan progressa, demanar-li a Priest una teoria on tota antinòmia acabe consumint-se finalment seria com demanar-li que aturés la roda de la història, és a dir, un impossible. Tanmateix, si sacrificar la validesa universal de *PNC* no és una cura universal contra les paradoxes, una solució consistent a la paradoxa de Curry que alhora solucionés altres paradoxes considerades per Priest seria més temptadora que el dialetheisme. Però hi ha altres problemes.

Dirimir què signifiquen *negar* i *rebutjar* no és fàcil per a un dialetheista. Segons Priest, no hi ha criteris formals que determinen quan una proposició és rebutjable o absurda. Per als lògics consistents, en canvi, tota proposició de la forma ' $p \wedge \neg p$ ' (o tota proposició que implique una contradicció) és absurda i rebutjable: $(p \wedge \neg p) \leftrightarrow \perp$. Molts justifiquen aquesta idea (TENNANT 1998) apel·lant a la nostra concepció de la *negació*, que codificaria una relació d'*incompatibilitat* entre dues proposicions que s'exclouen mútuament, fent absurda la seua acceptació conjunta. Priest (1999; 2002, 273-4) no pot acceptar aquesta teoria. La negació és per a ell una operació que canvia la polaritat d'una proposició i que construeix parelles de pols contradictoris sense determinar necessàriament la incompatibilitat dels pols, que no s'exclouen en una dialetheia (tota proposició té *un* i *sols un* contradictori: el de $\neg p$ és p , atès que $p \leftrightarrow \neg \neg p$). Una proposició necessàriament absurda seria aquella que impliqués que tot és vertader: $\forall x \text{Vertader}(x) \leftrightarrow \perp$,²⁴ però tals proposicions no tenen cap tret sintàctic distintiu. Com la majoria, Priest manté (i) que la negació d'una proposició equival a la seua falsedat; i (ii) que veritat i falsedat tenen un caràcter normatiu en la realització d'actes de parla i la formació d'actituds doxàstiques. El lògic clàssic pensa que tenim el deure d'asseverar i creure només allò vertader, i de negar i rebutjar només allò fals. Per tant, si acceptem que $\neg p$, sembla que cal negar i rebutjar que p (SMILEY 1993, 20). Si p fos una dialetheia, tindríem aparentment els deures contradictoris d'afirmar i negar, de creure i rebutjar alhora que p i això sembla erroni. Contra Kant, Priest podria defensar que *deure no implica poder*, que tenim deures inconsistents que no podem complir, no sols en l'àmbit moral i jurídic, sinó també en el doxàstic. Però en tal cas tenim un dilema: ¿creiem la dialetheia per vertadera o la rebutgem per falsa? Per a Priest (2006, 4.7 i 4.8), entre la veritat de p i la seua falsedat, $\neg p$, no hi ha una tercera opció. Els estats de coses representats per p i $\neg p$ són exhaustius però no exclouents: quan el món és inconsistent, p i $\neg p$ el representen correctament. En canvi

²⁴ '∀' i '∃' designen els quantificadors universal i existencial, respectivament.

(*ibidem*, 98), la distinció entre acceptar i rebutjar que p no és exhaustiva (podem ser agnòstics sobre p) però sí exclouent (no podem acceptar i rebutjar que p alhora). Per què acceptar, doncs, una dialetheia en lloc de rebutjar-la? Priest (1998, 421) respon que en certa manera “*truth trumps falsity*”. La veritat: allò que cal creure, l’objectiu de tota recerca, té més pes epistemològic que la falsedat. La falsedat sols és la veritat de la negació, té un pes epistemològic derivat. Només tenim raons determinants per a rebutjar que p quan p és *no-verdadera* (quan sols és falsa). Aquesta resposta és poc convincent. La asimetria que traça Priest entre veritat i falsedat, conceptes mútuament dependents i interdefinibles, sembla artificial. Tampoc és obvi que privilegiem sistemàticament l’acceptació de la veritat sobre el rebuig de la falsedat. Ambdues actituds semblen cares d’una mateixa moneda: el nostre interès per acceptar la veritat de p va lligat al nostre interès per rebutjar p si no és vertadera, ço és, si és falsa. Tanmateix, Priest distingeix ser fals de ser no-vertader (*untrue*, en anglès). Tot allò que no siga vertader és fals però no tot allò que és fals és no-vertader (les dialetheies) i per a Priest allò que justifica genuïnament el rebuig és ser no-vertader (ser sols fals). Dissortadament, l’estratègia de privilegiar l’actitud envers la veritat d’una dialetheia per a explicar per què cal creure-les té conseqüències paradoxals.

Considerem la paradoxa del mentider, una de les més recalitrants i antigues, i un dels arguments esgrimits per Priest a favor del dialetheisme, assumint la validesa universal de *PV*. En la seua versió més simple, la paradoxa involucra una proposició que diu de si mateixa que és falsa. Siga λ la proposició: λ és falsa. Si λ és vertadera, llavors (per *PV*) λ és falsa; i si λ és falsa, llavors (per *PV*) λ és vertadera. Suposar la veritat o falsedat de λ porta a una contradicció si veritat i falsedat són categories exclouents i exhaustives. Alguns han defensat que no són categories exhaustives i que λ no és ni vertadera ni falsa. Però suposem ara que μ és la proposició: μ no és vertadera. Si μ és vertadera, per *PV* μ no és vertadera. Si μ és falsa o no és ni vertadera ni falsa, en particular μ no és vertadera i, per *PV*, μ és vertadera. Aparentment no hi ha escapatòria a la contradicció. Priest manté que λ i μ són dialetheies: veritat i falsedat són categories exhaustives però no exclouents. Tanmateix, la tesi de Priest respecte a μ , la versió més corretjosa del mentider, té una conseqüència desagradable: μ és vertadera i falsa, però dir que és vertadera equival (per *PV*) a dir que *no és vertadera* o, en altres paraules, a dir que μ és *no-vertadera*. Priest ens ha dit que cal acceptar allò vertader (incloses les dialetheies) i rebutjar allò no-vertader. Què fem llavors amb μ que és vertadera i no-vertadera? μ ha de ser rebutjada tant com acceptada.²⁵ Priest (2006, 71-2) no ignora aquests fets. En diferents llocs,

²⁵ Vegeu també Sainsbury (2009, 157-8). Smiley (1993) i Bromand (2002) formulen versions més sofisticades de paradoxes que afecten al dialetheista. Trobareu una rèplica en Priest (2006, capítol 20).

admet (*ibidem*, 114) que hi ha proposicions i principis racionalment acceptables i rebutjables, i que la seua acceptabilitat no cancel·la el fet que també són rebutjables racionalment. Aquest fet no és per a Priest sinó una conseqüència d'acceptar una teoria inconsistent. Però cal dir que certes inconsistències són paradoxals en el context de la seua teoria, algunes tesis centrals del dialetheisme són elles mateixes dialetheies vertaderes i *no-vertaderes* difícils d'encaixar.

Si ni tan sols acceptar contradiccions elimina les paradoxes, quina actitud podem adoptar davant d'elles? La visió doxàstica partia d'un dogma *a priori*: cap contradicció és acceptable. El dialetheisme, considerant la persistència històrica de certes paradoxes i constatant empíricament que cap solució consistent ha estat satisfactòria fins ara, explora la possibilitat d'acceptar algunes contradiccions. Ambdues perspectives tenen un objectiu comú: domesticar les paradoxes, i ambdues accepten que les paradoxes ens forcen a revisar i reformar principis i conceptes. En certa manera, sols discrepen pel que fa a si una revisió pot o no traspasar la línia que marca *PNC*. Però si hi ha paradoxes que no podem domesticar, ¿no podríem “rendir-nos” davant d'elles? ¿No podríem acceptar que el nostre enteniment té límits més enllà dels quals és ociós tot intent per comprendre, i que les paradoxes dibuixen en certa manera tals límits?

El problema és que la idea d'identificar paradoxes “irresolubles” o “límits” del nostre pensament és inestable, paradoxal de fet: (1) Solucionar una paradoxa requereix una investigació exhaustiva, el mer fet de no haver-hi trobat una solució fins ara no demostra que no en tinga. Sempre podem imaginar que els nostres esforços han estat insuficients. (2) D'altra banda, cap investigació podria establir concloentment que una paradoxa és irresoluble en el sentit que hi ha un raonament paradoxal vàlid que *no* preserva la veritat de les premisses en la conclusió. Qualsevol argument que establís la conclusió de tal investigació seria en si mateix inacceptable: una paradoxa sobre la qual ens preguntariem si té solució o no. (3) Això suposa que la qüestió de si una paradoxa és o no irresoluble podria ser indecidible: davant d'una hipotètica paradoxa irresoluble, no podríem determinar satisfactòriament que ho és. D'on inferim que no podem saber si hi ha paradoxes irresolubles. (4) Ara bé, la nostra ignorància no demostra que no n'hi haja; no podem descartar aquesta possibilitat. Ens trobem en un atzucac.

La idea que el nostre pensament té límits *identificables* que no pot traspasar és igualment inestable. Traçar els límits absoluts del pensament humà suposaria ocupar una posició que ens està vedada. Hauríem d'identificar una “línia” que separés allò que podem entendre, pensar i conèixer d'allò intel·ligible. Però això no és possible. Reconèixer un límit absolut entre allò que podem pensar i allò inaccessible al pensament, requeriria pensar els dos costats de la línia i identificar allò que de fet és impensable. Si poguéssim accedir a l'altre costat de la línia, traspasar-la, això sols mostraria que tal línia no és el límit que

cercàvem, atès que hi podem pensar en el que hi ha més enllà. Potser uns éssers amb capacitats cognitives que incorporen i superen les nostres podrien fer un mapa del pensament humà i identificar pensaments concrets que no podem articular, però nosaltres no som tals éssers.

Wittgenstein anticipa aquestes idees en el *Tractatus* quan parla de la posició que el subjecte metafísic ocupa en el “seu” món, comparant-la amb la que ocupa l’ull amb respecte al seu camp visual (5.631-5.6331):

El subjecte no pertany al món, sinó que és un límit del món. [5.632]

On es pot notar en el món un subjecte metafísic?

Tu dius que aquí es comporten les coses com amb l’ull i el camp visual. Però l’ull, tu realment *no* el veus. [5.633]

El subjecte metafísic és ací “el subjecte que pensa, que fa representacions” (5.631), la seua constitució determina què pot ser pensat o representat, tal i com la constitució de l’ull determina què pot ser vist. Wittgenstein ens diu que aquest subjecte no pot ser pensat o representat perquè no pot aparèixer en el món que representa i pensa, de la mateixa manera que l’ull no pot aparèixer en el seu camp visual (5.6331). El subjecte, com l’ull, només configura els límits d’alguna cosa: del món pensable o representable en un cas, del món visible en l’altre, però aquests límits no apareixen en el món que limiten. Més encara, els límits no poden ser reconstruïts a partir d’allò que apareix en el món o el camp visual, “*res en el camp visual* no permet concloure que aquest és vist per un ull” (5.633). En 5.634 es nega la possibilitat d’un coneixement *a priori* tant dels límits del nostre món com d’allò que ens ve donat en l’experiència. I Wittgenstein, en 5.631, ens diu que “el subjecte que pensa, que es fa representacions, no existeix”. No pot ser conegut ni pensat com a objecte d’experiència perquè està fora del món, és el seu límit. Tampoc pot ser conegut *a priori* perquè tal coneixement l’ubicaria de nou dins del món representable i generaria la il·lusió que podem conèixer *a priori* certs aspectes de l’experiència: aquells que el subjecte-límit fa possibles o impossibles (extrem que Wittgenstein nega). Com que allò que existeix, existeix en el món —que no és altre que el món del subjecte, el món que *ell* limita (5.64, 5.641)—, Wittgenstein conclou que el subjecte metafísic no existeix perquè no pot existir *en el món*. El cos (l’ull) del qual té experiència i veu reflectit en l’espill, efectivament existeix, però *no és ell*, no és el subjecte (l’ull) metafísic del qual parla Wittgenstein: el límit del món (del camp visual).

Arribats a aquest punt, Wittgenstein —o el lector que accepti el seu raonament o l’argument similar que he exposat adés— té el mateix problema que Parmènides. Si Wittgenstein té raó, el seu discurs és il·legítim. No sols *pensa* i *representa* el subjecte metafísic com a límit del món, sinó que pretén demostrar

a priori que el límit del món, el subjecte, no pot estar donat en l'experiència i en fer açò transgredeix els límits que suposadament no podíem transgredir. A diferència de Parmènides, Wittgenstein sí que és conscient d'aquesta i d'altres paradoxes presents en el *Tractatus* i deixa ben clar en 6.54 que qui entén les seues proposicions les reconeix finalment com a sense-sentit. Tanmateix, això no les fa del tot inservibles, al seu parer. Wittgenstein sembla pensar que tota reflexió sobre allò que fa possible el pensament, o sobre els seus límits, és inexorablement inconsistent i, no sent un dialetheista, no està disposat a acceptar que algunes contradiccions representen veritats sobre el món, o que tinguen contingut diferent (sent unes acceptables i altres no). Si bé és cert que una contradicció no és un sense-sentit (4.4611), no és menys cert que descriure un fet del món mitjançant una contradicció sí que ho és. I, si efectivament tota reflexió sobre els límits del pensament és incoherent, res no es pot dir amb sentit pel que fa al cas. Però la millor manera potser de mostrar açò és reflexionar seriosament sobre tals límits i fracassar, mostrar que fins i tot quan volem concloure la impossibilitat de conèixer els límits del pensament, els millors arguments són paradoxals. Si bé els arguments no podrien establir coherentment la seua conclusió, sí que podrien “mostrar” que un problema és irresoluble, o que hem topat amb un límit cognoscitiu, a través del seu fracàs i la nostra incapacitat de detectar cap error. De nou, açò no pot ser un fet demostrable del món, per tant, Wittgenstein sols podria esperar del *Tractatus* que ens ajudés a veure les qüestions allà examinades d'una determinada manera.²⁶ N'hi ha d'altres: podríem creure encara que totes les dificultats relacionades amb allò que es pot pensar i allò que no acabaran resolent-se finalment. El *Tractatus* no encoratja aquesta actitud però tampoc pot eliminar-la, com sí que va eliminar Schliemann tot dubte sobre l'existència de Troia en desenterrar-la.

Tota paradoxa sembla imposar-nos un deure, el de rebutjar allò inacceptable i restaurar una visió coherent del món, però la resistència d'algunes antinòmies a ser resoltes després de molts intents també inoculara una sospita, la de ser moderns Sísifs que empenyen cada dia la mateixa pedra fins al cim de la mateixa penya, només per veure-la caure a l'endemà davant nostre. La impossibilitat de decidir si una paradoxa és o no realment una aporia insalvable ens condemna a la ignorància, però no cancel·la ni els nostres deures ni els nostres temors. Conseqüentment, les certeses absolutes respecte a la resolubilitat o intractabilitat de certes paradoxes semblen adquirir un caràcter místic —atès que no tenim raons concloents (i potser no n'hi ha) a favor d'una o altra tesi— i fomenten dues visions metafísiques enfrontades de les paradoxes. El *Tractatus* il·lustra vívidament aquest fet i una de les perspectives possibles.

²⁶ Explore aquestes idees en Valor Abad (2008B).

4. LES PARADOXES D'AUTOREFERÈNCIA

Hi ha un grup de paradoxes que ha estat especialment rellevant per a la filosofia i la ciència al llarg del segle xx. En un text fundacional, Bertrand Russell (1908) les anomena paradoxes d'*autoreferència* i així les coneixem avui dia. Aquestes paradoxes amenacen qualsevol definició coherent dels conceptes fonamentals per a la ciència i el pensament en general, com ara *veritat*, *conjunt*, *classe*, *relació*, *definició*, *número cardinal*, *número ordinal*. Frank Ramsey (1931, 20) les va dividir en dos grups atenent a si involucraven conceptes propis de les ciències formals (*número*, *classe*) o conceptes semàntics que impregnen el nostre llenguatge (*veritat*, *relació*). La intuïció de Ramsey és que l'estratègia que cal seguir davant d'una paradoxa depèn de l'àmbit en què es produeix: els principis i conceptes afectats són diferents en cada cas, i els estàndards normatius d'una solució satisfactòria també. En l'àmbit d'un càlcul formal on la consistència és crucial —una idea que, gràcies a l'obra de Priest i d'altres, avui podem matisar—, evitar una paradoxa pot justificar introduir restriccions conceptuals que limiten la capacitat expressiva del nostre llenguatge, però que, a canvi, eliminen la possibilitat d'inferir contradiccions. En l'àmbit del llenguatge ordinari, en canvi, eliminar contradiccions no basta, i potser no és tant vital. Qualsevol restricció que impedisca expressar proposicions intuïtivament vertaderes per tal d'evitar l'expressió de proposicions paradoxals com λ , μ o κ semblarà arbitrària i insatisfactòria. Les idees de Ramsey en aquest sentit han dominat el segle xx. Branques de la matemàtica com la Teoria de Conjunts han incorporat restriccions importants pel que fa a allò que podem considerar un conjunt amb el propòsit d'evitar contradiccions i tals restriccions han passat a ser solucions canòniques a les paradoxes conjuntistes. En canvi, paradoxes d'autoreferència com la del mentider (examinada abans) encara no han trobat solucions satisfactòries, malgrat que Tarski (1936) o Kripke (1975) han construït teories consistents relacionades amb el concepte de *veritat* (en un cas en l'àmbit de la lògica clàssica i en l'altre en el de diverses lògiques trivalents). El problema en ambdós casos és que les teories introdueixen restriccions expressives que no tolerem. La solució de Tarski prohibeix l'autoreferència en oracions amb conceptes semàntics i no funciona per a llenguatges on podem expressar veritats òbvies com 'és veritat que aquesta oració conté huit paraules' i en la teoria de Kripke μ no és ni vertadera ni falsa, però no hi ha cap oració distinta de μ que pugui expressar allò que diu la teoria: *que μ no és vertadera* (i μ tampoc expressa aquesta veritat meta-teòrica, atès que no és ni vertadera ni falsa).

Les intuïcions de Ramsey sobre les paradoxes d'autoreferència difereien però de les del propi Russell. Russell pensava que aquestes paradoxes comparti-

en una sèrie de trets estructurals que justificaven la seua inclusió en *una* família (per això es va molestar a batejar-les). Però també pensava que l'estructura d'aquestes paradoxes revelava un problema comú que totes elles exemplificaven. Desactivar el problema oferiria, per tant, una solució uniforme a totes les paradoxes d'autoreferència. L'etiqueta 'autoreferència' és enganyosa en realitat, perquè no totes les paradoxes així etiquetades involucren problemes d'autoreferència. Russell (1908, 63) fou més precís quan va diagnosticar que totes elles involucren alguna mena de circularitat viciosa que podem descriure així: totes elles parlen d'una entitat a i una col·lecció Ω (en algunes paradoxes a pot ser la pròpia Ω) tals que a és un possible membre de Ω però un que sols pot ser definit apel·lant a tots els membres de Ω (incloent el propi a). Açò comporta un cercle viciós en les definicions de Ω i a , ja que cal conèixer tots els membres de Ω per a definir a però cal determinar si $a \in \Omega$ per a fixar l'extensió de Ω . La solució de Russell en el seu *Principi de Circularitat Viciosa (PCV)* és clara: tals col·leccions no existeixen.

En 1905, Russell va donar una descripció més acurada encara de l'estructura d'un subgrup de les paradoxes d'autoreferència que afecten a la Teoria de Conjunts. Basant-se en aquesta descripció i modificant-la mínimament, Priest (2002, 133-6) ha identificat un esquema, *Inclosure Schema (IS)*, en el qual encaixen totes les paradoxes batejades per Russell i moltes altres (que cal considerar com a membres de la mateixa família). Priest, igual que Russell, pensa que totes les paradoxes que són "inclosures" han de rebre una solució uniforme que ataque *el* problema comú que exemplifiquen i afegeix que sols el dialetheisme satisfà aquest criteri (*ibidem*, 166-7). Aquest argument configuraria la quarta estratègia dialèctica (mencionada en 3.2) a favor del dialetheisme. Però què és una inclosure i què una paradoxa d'autoreferència?

En una breu ficció titulada *Del rigor de la ciència*, Borges (1960) ens parla d'un Imperi on les arts cartogràfiques van assolir tal esplendor que els cartògrafs decidiren confeccionar un mapa descomunal de l'imperi que coincidís punt per punt amb ell, cada objecte hi tindria una representació fidedigna. El projecte és finalment abandonat per estèril i restes del mapa queden espargides per tot l'imperi... La qüestió és: ¿seria realment possible construir tal mapa? El mapa [a] conté representacions acurades de *totes* les coses que hi ha a l'imperi [Ω] i és ell mateix un dels objectes que hom troba a l'imperi [$a \in \Omega$]. Per figurar fidelment l'imperi, el mapa ha de contenir una representació de si mateix, però el mapa representat sols coincidirà punt per punt amb el real si el primer representa la complexitat del segon, on apareixen tots els objectes de l'imperi, incloent-hi el propi mapa... Sembla clar que ens sotja un regrés a l'infinit provocat per la relació circular que guarden el mapa, que sols pot definir-se apel·lant a tots els objectes de l'imperi, i el propi imperi, que conté, entre altres coses, el

mapa. Russell diria que el mapa no pot existir apel·lant a *PVC*. L'imperi sols pot existir entès com una totalitat Ω que *no* involucra un membre a (el mapa) únicament definible en termes de tots els membres de Ω .

Què hauria passat si, alertats pel problema, els cartògrafs hagueren decidit cartografiar només aquelles coses que “no s'autorepresenten”, ço és, que *no contenen representacions de si mateixes*? El projecte continuaria sent inviable per la mateixa raó però, a més d'un problema de circularitat, tindríem ara problemes de consistència. ¿Ha de contenir el mapa una representació seua o no? Si la conté, ha representat un objecte que s'autorepresenta; si no la conté, ha deixat de representar un objecte que no s'autorepresenta. En cap cas pot representar tots (i sols) els objectes de l'imperi que no s'autorepresenten. De nou, determinar la totalitat Ω d'objectes de l'imperi que satisfan la condició “no s'autorepresenta” involucra decidir si el mapa a pertany o no a Ω . Però definir a requereix saber quins són tots els objectes que satisfan la condició anterior. El problema addicional del nou mapa és que podem inferir que $a \in \Omega \Leftrightarrow a \notin \Omega$. La descripció estructural de Russell en 1908 arreplega el primer aspecte del problema: la circularitat. La descripció de Russell (1905) i Priest (2002) arreplega també el segon aspecte del problema: la inconsistència. Descrit formalment, l'esquema de Russell-Priest, *IS*, té aquest aspecte:

- (1) $\Omega = \{y : \varphi(y)\}$ existeix i $\psi(\Omega)$.
- (2) Si $X \subseteq \Omega$ i $\psi(X)$, llavors: [2.a] $\delta(X) \notin X$; i [2.b] $\delta(X) \in \Omega$

La primera part d'*IS*, (1), diu que existeix el conjunt Ω de totes les coses que satisfan la condició φ , i que Ω satisfà la condició ψ . La segona part, (2), diu que hi ha una funció δ definida per als subconjunts X de Ω que satisfan la condició ψ tal que δ assigna a X un element $\delta(X)$ de Ω [2.b] que no pertany a X [2.a]. La contradicció es produeix quan considerem un subconjunt particular de Ω que satisfà la propietat ψ , a saber, el propi Ω : en tal cas, δ li assigna a Ω un element $\delta(\Omega)$ tal que: $\delta(\Omega) \notin \Omega$ i $\delta(\Omega) \in \Omega$. Priest anomena *diagonalitzador* a la funció δ precisament perquè codifica un argument diagonal (vegeu 3.1). Per tant, *IS* identifica dos elements clau en tota paradoxa d'autoreferència: l'existència de la mena de relació circular descrita per Russell entre Ω i un dels seus membres $\delta(\Omega)$, i la presència d'un argument diagonal.

Els mapes de Borges no generen cap antinòmia. Cap descripció d'un mapa o d'un barber pot forçar-nos a creure en la seua existència. Quan descobrim la paradoxa, deixem de pensar (sense traumes) que puguem existir el mapa de Borges o el barber que afaita a tots aquells que no s'afaiten a si mateixos. Però ambdós casos exemplifiquen *IS* i poden considerar-se *inclosures*. Suposem que existeix la col·lecció Ω de totes les coses que no s'autorepresenten i

suposem que tal col·lecció satisfà la condició Ψ de *ser definible* (per exemple mitjançant l'expressió “la col·lecció Ω de totes les coses que no s'autorepresenten”). Suposem ara que δ és una operació que permet construir mapes (que no s'autorepresenten) de grups definibles X d'objectes de Ω ($X \subseteq \Omega$ i X satisfà Ψ). Donats tres objectes a, b i c de Ω , δ faria un mapa $\delta(X)$ que representa els membres del grup $X = \{a, b, c\}$. Com que δ construeix mapes que no s'autorepresenten, $\delta(X) \in \Omega$ però, per la mateixa raó, $\delta(X) \notin X$, ja que si $\delta(X)$ pertanyés a X (si per exemple $\delta(X) = c$), $\delta(X)$ s'autorepresentaria i no seria membre de Ω . ¿Què passa quan intentem construir un mapa $\delta(\Omega)$ de tots els objectes que no s'autorepresenten? En tal cas, el mapa hauria i no hauria d'autorepresentar-se: $\delta(\Omega) \notin \Omega$ i $\delta(\Omega) \in \Omega$.

Al marge de casos anecdòtics, hi ha *inclosures* que sí que són antinòmiques. La paradoxa del mentider n'és una. Identifiquem Ψ amb la propietat: *ser definible lingüísticament* i definim Ω com: *el conjunt V de coses que satisfan el predicat 'ser vertader'*. (La proposició: *p és vertadera*, es representa ací com $p \in V$; i *p no és vertadera* com $p \notin V$.) L'operació δ assigna a cada conjunt definible X de proposicions vertaderes una proposició α ($\delta(X) = \alpha$) que diu: $\alpha \notin X$ (α diu de si mateixa que no és cap de les proposicions que hi ha a X). Si α no és vertadera, llavors $\alpha \in X$, però això és impossible perquè totes les proposicions d' X són vertaderes. Per tant, α és vertadera: $\alpha \in V$, però (per *PV*) això significa que $\alpha \notin X$. Com que $\Omega = V$ i $\alpha = \delta(X)$, hem demostrat que per a tot X definible tal que $X \subseteq V$, $\delta(X) \notin X$; i $\delta(X) \in \Omega$. Quan apliquem δ a V , tenim la paradoxa del mentider: $\delta(V) \notin V$ i $\delta(V) \in V$, ja que $\delta(V) = \mu$ i μ és la proposició $\mu \notin V$. Açò és una antinòmia perquè negar l'existència de V és negar que hi haja un conjunt que siga l'extensió del predicat 'ser vertader'.

A més del mentider hi ha altres dues antinòmies famoses: la de Russell (relativa a la noció de conjunt) i la de Grelling, que també són *inclosures* generades per l'ús no restringit de *PA* en el primer cas, i de *PS* en el segon. La paradoxa de Russell considera el conjunt $R = \{x : x \notin x\}$ de tots els conjunts que no pertanyen a si mateixos i es genera quan preguntem si R pertany a R (ja que $R \in R \leftrightarrow R \notin R$). La de Grelling afecta a la relació semàntica de satisfacció —que x manté amb un predicat φ si, i sols si, x és φ — i defineix un predicat, 'heterològic', de la següent manera: *un predicat lingüístic ' φ ' és heterològic si ell mateix no és φ* . El predicat 'llarg' és heterològic perquè 'llarg' no és llarg, el predicat 'curt' no és heterològic perquè 'curt' és curt però el predicat 'heterològic' és heterològic si, i sols si, no és heterològic. Priest (2002, capítols 9-11) demostra que aquestes paradoxes, totes les descrites per Russell en 1908 (Burali-Forti, Mirimanoff, Berry, König, etc.) i moltes altres són *inclosures* i usa el mode en què cada una exemplifica *IS* per distingir diferents subfamílies dins de les *inclosures*. També són *inclosures* (VALOR ABAD 2008A, apèndix) la paradoxa de la

cita de Tarski (1936, 159-62) i la paradoxa de l'hiperjoc de Zwicker (1987). Actualment, Priest (2010) defensa que les paradoxes de sorites són *inclosures* i argumenta (Priest 2002, part 4) que podem trobar *inclosures* en les obres de filòsofs com Frege, Wittgenstein, Quine o Davidson quan reflexionen sobre (el que podríem descriure informalment, seguint Priest, com) “els límits del pensament”.

La importància d'aquestes paradoxes sembla cabdal per a entendre moltes de les antinòmies contemporànies de la filosofia i de les ciències. L'existència d'una descripció estructural uniforme de totes elles obri la possibilitat de trobar connexions il·luminadores entre àmbits i problemes aparentment diversos i de relacionar les solucions estables o provisionals que oferim a tals problemes en cada àmbit. Tanmateix, algunes de les tesis de Priest respecte a *IS* han estat fortament criticades. Alguns argumenten (WILLIAMSON 1996) que hi ha paradoxes d'autoreferència com la de Curry que no encaixen en *IS*. Priest (2002, 11.8) ja reconeix aquest fet i indica que només algunes versions d'aquesta paradoxa encaixen en *IS*. Altres pensen que la paradoxa de Yablo i la del mentider estan més relacionades entre si que la del mentider i altres *inclosures*, però la paradoxa de Yablo no és una *inclosure* perquè no involucra cap mena de circularitat: genera una contradicció a partir d'una sèrie infinita de proposicions tals que en ella cap proposició p parla de si mateixa o de cap altra que parle de p .

- (0) Per a tot x si $x > 0$, x no és vertader.
- (1) Per a tot x si $x > 1$, x no és vertader.
- (2) Per a tot x si $x > 2$, x no és vertader...

Si 0 és vertadera, 1, 2, 3, etc. són falses. Però si 2, 3, etc. són falses, 1 és vertadera. Per tant, 0 és falsa. Però si 0 és falsa, alguna proposició $n > 0$ serà vertadera i, reproduint el raonament anterior a partir de la veritat de n , tenim una nova contradicció. Priest (1997) respon que quan definim recursivament la sèrie infinita 0, 1, 2, etc., podem demostrar que la paradoxa de Yablo és una *inclosure*. Altres crítics argumenten que el fet de poder rebre una descripció estructural uniforme, no justifica que totes les paradoxes-*inclosures* hagen de rebre una solució uniforme (GRATTAN GUINNESS 1998). Priest indica que *no* tota *inclosure* és una paradoxa i que no hi ha criteris formals per a identificar dialetheies, per tant, exemplificar *IS* és una condició necessària però no suficient per a aïllar el problema que genera una paradoxa i afirmar que *IS* identifica el problema comú a totes elles, o la causa que les genera, és incorrecte (VALOR ABAD 2008A). En tot cas, la tesi que el dialetheisme és l'única solució uniforme a les *inclosures* perquè en cada cas accepta una contradicció és vàlua,

ja que, assumint un nivell d'abstracció similar, tota solució consistent a les paradoxes ofereix una solució uniforme: bloquejar la inferència d'una contradicció (SMITH 2000).

Agraiments. Vull agrair els comentaris d'Enric Casaban Moya i José Martínez Fernández a algunes idees arregades en aquest treball, que ha rebut el suport del projecte "Reference, Self-Reference, and Empirical Data" (FFI2011-25626), finançat pel Ministerio de Ciencia e Innovación.

BIBLIOGRAFÍA

- ARISTÒTIL 1982, *Tratados de lógica (Órganon) I*, traducció castellana i edició de M. CANDEL, Madrid, Gredos.
- ARISTÒTIL 1994, *Metafísica*, traducció castellana i edició de T. CALVO, Madrid, Gredos.
- BEALL, J. C. 2004, “True and False –As if”, en PRIEST, BEALL i ARMOUR-GARB 2004, 197-216.
- BEALL, J. C. i ARMOUR-GARB, B. 2003, “Should Deflationists Be Dialetheists?”, *Noûs*, 37: 303-24.
<http://dx.doi.org/10.1111/1468-0068.00440>
- BORGES, J. L. 1960, *El hacedor*, Buenos Aires, Emecé Editores.
- BROMAND, J. 2002, “Why Paraconsistent Logic Can Only Tell Half the Truth”, *Mind*, 111: 741-9.
<http://dx.doi.org/10.1093/mind/111.444.741>
- CLARK, M. 2007 (2a edició), *Paradoxes from A to Z*, Londres/NovaYork: Routledge.
- CURRY, H. 1942, “The Inconsistency of Certain Formal Logics”, *Journal of Symbolic Logic*, 10: 115-7.
<http://dx.doi.org/10.2307/2269292>
- DEVLIN, K. 1993 (2a edició), *The Joy of Sets*, Nova York: Springer Verlag.
<http://dx.doi.org/10.1007/978-1-4612-0903-4>
- GRATTAN-GUINNESS, I. 1998, “Structural Similarities or Structuralism”, *Mind*, 107: 823-34.
<http://dx.doi.org/10.1093/mind/107.428.823>
- KEEFE, R. i SMITH, P. 1997, “Introduction: Theories of Vagueness”, en *Vagueness: A Reader*, editat per R. KEEFE i P. SMITH, Cambridge (Massachusetts): MIT Press, 1-57.
- KIRK, G. S., RAVEN, J. E. i SCHOFIELD, M. 1983, *The Presocratic Philosophers. A Critical History with a Selection of Texts*, Cambridge: Cambridge University Press. Cite la traducció castellana de J. GARCÍA FERNÁNDEZ: *Los filósofos presocráticos. Historia crítica con selección de textos*, Madrid: Gredos, 1987.
- KRIPKE, S. 1975, “Outline of a Theory of Truth”, *Journal of Philosophy*, 72: 690-716.
<http://dx.doi.org/10.2307/2024634>
- MARES, E. 2004, “Semantic Dialetheism”, en PRIEST, BEALL i ARMOUR-GARB 2004, 264-75.
- PRIEST, G. 1979, “The Logic of Paradox”, *Journal of Philosophical Logic*, 8: 219-41.
<http://dx.doi.org/10.1007/BF00258428>
- 1997, “Yablo’s Paradox”, *Analysis*, 57: 236-242.
<http://dx.doi.org/10.1093/analys/57.4.236>

- 1998, “What is so Bad about Contradictions?”, *Journal of Philosophy*, 95, 410-26. (Reimprès en PRIEST, BEALL i ARMOUR-GARB 2004, 23-38.)
- 1999: “What Not? A Defence of a Dialethic Account of Negation”, en GABBAY i WANSING (eds.), *What is Negation?*, Dordrecht: Kluwer, 1999, 101-20.
- 2002 (2a edició), *Beyond the Limits of Thought*, Oxford: Oxford University Press.
- 2006 (2a edició), *In Contradiction. A Study of the Transconsistent*, Nova York: Oxford University Press.
- 2010, “Inclosure, Vagueness and Self-Reference”, *Notre Dame Journal of Formal Logic*, 51: 69-84.
<http://dx.doi.org/10.1215/00294527-2010-005>
- PRIEST, G., BEALL, J. C. i ARMOUR-GARB, B. (eds.) 2004, *The Law of Non-Contradiction: New Philosophical Essays*, Oxford: Oxford University Press.
- QUINE, W. V. O. 1966, “The Ways of Paradox”, en *The Ways of Paradox and Other Essays*, Cambridge (Massachusetts): Harvard UP, 1-18.
- RAMSEY, F. 1931, *The Foundations of Mathematics and Other Logical Essays*, editat per R. B. BRAITHWAITE, Londres: Routledge & Kegan Paul, 1-61.
- RUSSELL, B. 1905, “On Some Difficulties in the Theory of Transfinite Numbers and Order Types”, *Proceedings of the London Mathematical Society* (series 2), 4: 29-53.
- 1908, “Mathematical Logic as based on the Theory of Types”, *American Journal of Mathematics*, 30: 222-262. Reimprès en RUSSELL, *Logic and Knowledge*, editat per R. C. MARSH, Londres: Allen & Unwin, 1956, 59-102. (Referències a aquesta edició)
- SAINSBURY, R. M. 2009 (3a edició), *Paradoxes*, Cambridge: Cambridge University Press.
- SALMON, W. C. (ed.) 1970, *Zeno's Paradoxes*, Indianapolis: Bobbs-Merrill.
- 1980 (2a edició), *Space, Time and Motion: A Philosophical Introduction*, Minneapolis: University of Minnesota Press.
- SIMMONS, K. 1993, *Universality and the Liar. An Essay on Truth and the Diagonal Argument*, Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511551499>
- SMILEY, T. 1993, “Can Contradictions Be True? - I”, *Proceedings of the Aristotelian Society, Suppl. Vol.*, 67: 17-33.
- SMITH, N. J. J. 2000, “The Principle of Uniform Solution (of the Paradoxes of Self-Reference)”, *Mind*, 109: 117-22.
<http://dx.doi.org/10.1093/mind/109.433.117>
- SORENSEN, R. 2003, *A Brief History of Paradox*, Nova York: Oxford University Press.
- TARSKI, A. 1936, “Der Wahrheitsbegriff in den formalisierten Sprachen”, *Studia Philosophica*, 1: 261-405. Traduït a l'anglès en TARSKI, *Logic Semantics and Metamathematics*, traduït per J. H. WOODGER i editat per J. CORCORAN,

- Indianàpolis: Hackett, 1983 (2a edició), 152-278 (les referències remeten a aquesta traducció).
- TARSKI, A. 1944, "The Semantic Conception of Truth and the Foundations of Semantics", *Philosophy and Phenomenological Research*, 4: 341-75.
<http://dx.doi.org/10.2307/2102968>
- TENNANT, N. 1998, "Critical Notice of Beyond the Limits of Thought", *Philosophical Books*, 39: 20-38.
- THOMSON, J. F. 1970, "Tasks and Super-Tasks", *Analysis*, 15: 1-13.
<http://dx.doi.org/10.1093/analys/15.1.1>
- VALOR ABAD, J. 2008A, "The Inclosure Scheme and the Solution to the Paradoxes of Self-Reference", *Synthese*, 160: 183-202.
<http://dx.doi.org/10.1007/s11229-006-9109-x>
- 2008B, "¿Podemos representar el campo visual? (*Tractatus* 5.6331)", en *Sentido y sinsentido. Wittgenstein y la crítica del lenguaje*, editat per C. MOYA, Valencia: Pre-Textos, 2008, 227-42.
- YABLO, S. 1993, "Paradox without self-reference", *Analysis*, 53: 251-2.
<http://dx.doi.org/10.1093/analys/53.4.251>
- WILLIAMSON, T. 1996, "Review of *Beyond the Limits of Thought*, by G. Priest", *British Journal for the Philosophy of Science*, 47: 331-4.
<http://dx.doi.org/10.1093/bjps/47.2.331>
- WITTGENSTEIN, L. 1921, *Tractatus Logico-Philosophicus*. Cite la traducció al català a cura de J. ORDI FERNÁNDEZ, *Anuari de la Societat Catalana de Filosofia XVI (2004/2005)*, Barcelona: Societat Catalana de Filosofia/Institut d'Estudis Catalans, 2005, 117-217.
- WRIGHT, C. 1976, "Language Mastery and the Sorites Paradox", en *Truth and Meaning. Essays on Semantics*, editat per G. EVANS i J. MCDOWELL, Oxford: Oxford University Press, 223-47. Reimprés en *Vagueness: A Reader*, editat per R. KEEFE i P. SMITH, Cambridge (Massachusetts): MIT Press, 1997, 151-73; les referències remeten a aquesta edició.
- ZWICKER, W. 1987, "Playing Games with Games: The Hypergame Paradox", *The American Mathematical Monthly*, 94: 507-14.
<http://dx.doi.org/10.2307/2322840>