Aviso sobre esta traducción
Este libro es una traducción automática del texto original en inglés, realizada con herramientas de traducción automatizada. No ha sido revisado ni corregido por el autor ni por ningún otro ser humano. Por lo tanto, pueden presentarse errores, imprecisiones, frases poco fluidas o problemas de formato. La edición original en inglés sigue siendo la versión autorizada.

Paradojas y sus soluciones

Una recopilación temática

Avi Sion, Doctor en Filosofía.

© Copyright Avi Sion, 1990-2017.

Protegido por los convenios internacionales y panamericanos de derechos de autor.
Reservados todos los derechos.
Ninguna parte de este libro puede ser reproducida de ninguna manera.
o almacenado en un sistema de recuperación o transmitido,
sin el permiso expreso del autor-editor,
excepto en el caso de citas breves con el debido reconocimiento.

Publicado con orgullo por Avi Sion en 2017.
Ginebra, Suiza.

www.TheLogician.net
avi-sion@thelogician.net

Información de catalogación de la biblioteca:

Sion, Avi
Paradojas y sus soluciones
Una recopilación temática

Sin índice, sin bibliografía

Abstracto

Paradojas y sus soluciones Se trata de una «compilación temática» de Avi Sion. Reúne en un solo volumen los ensayos que ha escrito a lo largo de 27 años sobre este tema. Incluye exposiciones y resoluciones de numerosas paradojas antiguas y modernas (aunque no todas), entre ellas: la paradoja de Protágoras y Euatlus (Atenas, siglo V a. C.), la paradoja del Mentiroso y la paradoja de Sorites (ambas atribuidas a Eubúlides de Mileto, siglo IV a. C.), la paradoja de Russell (Reino Unido, 1901) y sus derivados, la paradoja de Barber y la paradoja del Catálogo Maestro (también de Russell), la paradoja de Grelling (Alemania, 1908), la paradoja de la confirmación de Hempel (EE. UU., década de 1940) y la paradoja de la predicción de Goodman (EE. UU., 1955). Este volumen también presenta y comenta algunos de los discursos antinómicos que se encuentran en algunos textos budistas (a saber, en Nagarjuna, India, siglo II d. C.; y en el Sutra del Diamante, fecha desconocida, pero probablemente de principios de siglo d. C.).

A es A (y no-A no es A);
nada es A y no-A a la vez;
Nada es ni A ni no-A.

(Las tres leyes del pensamiento de Aristóteles.)

Contenido

Prólogo 8
1. La vanidad del tetralema 13
   1. Los fenómenos son positivos 13
   2. No existen fenómenos negativos 14
   3. Un experimento mal interpretado 16
   4. Definición de negación 18
2. Aclaración de la contradicción 21
   1. Dialéctica 21
   2. Contradicción 23
   3. La coherencia es natural 29
3. Aclaración de la negación 32
   1. Negación en la aducción 32
   2. Fenómenos positivos y negativos 35
   3. La negación es secundaria 37
   4. La negación es una intención 41
4. Aclaración de la paradoja 44
   1. Inconsistencia interna 44
   2. La falacia del concepto robado 46
   3. Sistematización 49
   4. Propiedades 53
5. La paradoja del mentiroso (temprana) 59
   1. Doble paradoja 59
   2. La paradoja del mentiroso 61
   3. Más sobre la paradoja del mentiroso 63
   4. La utilidad de las paradojas 65
6. La paradoja del mentiroso (revisión) 70
   1. Primer acercamiento 70
   2. Segundo enfoque 72
   3. Tercer enfoque 75
   4. Cuarto enfoque 77
   5. Quinto enfoque 79
   6. Sexto enfoque 81

7. La paradoja de Russell (temprana) 84
1. Automembresía 84
2. La paradoja de Russell 86
3. Impermutabilidad 88
4. La paradoja del barbero 92
5. La paradoja del catálogo maestro 94
6. La paradoja de Grelling 98
8. La paradoja de Russell (revisión) 103
1. Elementos de la lógica de clases 103
2. Una aparente doble paradoja 109
3. Un poco de historia 121
4. Un poco de autocrítica 128
9. Más sobre la paradoja de Russell 131
1. Mi resolución de la paradoja de Russell 131
2. Por qué las resoluciones de Russell son inadecuadas 134
3. Por qué la resolución de Rescher es inadecuada 138
4. Otras encarnaciones de la paradoja 140
5. Sobre la paradoja de Barber 142
10. La paradoja de la confirmación de Hempel ... 144
1. Análisis tradicional 144
2. Análisis de la novela 146
3. El problema y su solución 152
11. La paradoja de la predicción de Goodman 160
1. El supuesto problema 160
2. La solución lógica 161
3. El concepto de 'grue' 164
4. ¿Hasta qué punto de la escala? 167
12. La paradoja de Sorites 171
1. ¿Qué es un montón? 171
2. El uso de términos vagos 172
3. Razonamiento con términos vagos 175
4. Inventar falsas paradojas 178

13. Protágoras contra Euatlo 182
1. Una antigua paradoja 182
2. Primera resolución 184
3. Segunda resolución 186
4. Resoluciones inadecuadas 190
14. Discurso antinómico budista 195
1. El tetralema 195
2. Ni real ni irreal 200
15. Más discurso antinómico budista 207
1. La tesis 'No tengo tesis' 207
2. Llamar a las cosas por su nombre 225
Referencias principales 230

Prefacio

Paradojas y sus soluciones Se trata de una «compilación temática». Reúne en un solo volumen los ensayos que he escrito en el pasado (a lo largo de unos 27 años) sobre este tema.

Comprende exposiciones y resoluciones de muchas, aunque no todas, las paradojas antiguas y modernas, entre las que se incluyen: la paradoja de Protágoras-Euathlus (Atenas, siglo V ^{a} . C.), la paradoja del Mentiroso y la paradoja de Sorites (ambas atribuidas a Eubúlides de Mileto, siglo IV ^{a} . C.), la paradoja de Russell (Reino Unido, 1901) y sus derivados, la paradoja de Barber y la paradoja del Catálogo Maestro (también de Russell), la paradoja de Grelling (Alemania, 1908), la paradoja de la confirmación de Hempel (EE. UU., década de 1940) y la paradoja de la predicción de Goodman (EE. UU., 1955).

También presento y comento aquí algunos de los discursos antinómicos que se encuentran en algunos textos budistas (a saber, en Nagarjuna, India, siglo II ^{d} . C.; y en el Sutra del Diamante, fecha desconocida, pero probablemente de principios de siglo d. C.).

A pesar de su título, cabe señalar que el presente libro no pretende ser un estudio exhaustivo; hay muchas paradojas que no menciona ni aborda.[1]

---

[1] En Wikipedia, https://en.wikipedia.org/wiki/List_of_paradoxes , se encuentra una interesante «lista de paradojas» que las agrupa bajo diversos epígrafes y contiene enlaces a entradas sobre cada una. También hay artículos sobre paradojas en la Enciclopedia de Filosofía de Stanford en línea. Un libro sobre este tema que merece la pena consultar es *Paradoxes: Their Roots, Range, and Resolution, de Nicholas Rescher* (Chicago, Ill.: Open Court, 2001).

Por «paradoja» se entiende aquí, por supuesto, una *aparente doble paradoja* . Una proposición paradójica tiene la forma hipotética contradictoria en sí misma «si P, entonces no P», o «si no P, entonces P», donde P es cualquier tipo de proposición. Tal proposición, tomada aisladamente, no es antinómica, es decir, no infringe las leyes del pensamiento, simplemente porque existe una salida formal a ella: «si P, entonces no P» implica la proposición categórica «no P»; y «si no P, entonces P» implica la proposición categórica «P». Una *sola* paradoja, entonces, constituye un discurso lógicamente legítimo. Una *doble* paradoja es muy diferente, es decir, una afirmación que sostiene que *tanto* "si P, entonces no P" como "si no P, entonces P" son simultáneamente verdaderas para una instancia dada de P. Dicha afirmación es diametralmente opuesta al principio de no contradicción, ya que concluye que tanto las proposiciones categóricas "no P" como "P" son simultáneamente verdaderas para una instancia dada de P.

Es evidente que resolver paradojas (dobles) es esencial para la lógica, la ontología y la epistemología, ya que aceptar cualquier antinomia de este tipo pondría en grave duda el conocimiento humano, e incluso las facultades cognitivas que lo hacen posible. Los escépticos se regodean en las paradojas, pues desean maliciosamente invalidar el conocimiento y la mente humanos. Por lo tanto, quienes defienden la razón humana tienen la obligación de afrontar cada desafío de este tipo y neutralizarlo de manera convincente. En un sentido más positivo, las paradojas son excelentes oportunidades para aprender algo nuevo sobre nuestra forma de pensar o nuestras creencias, y para descubrir y corregir nuestros errores. Si no nos topáramos con la paradoja, podríamos permanecer ajenos a nuestros errores; la paradoja nos abre la puerta a corregirlos.

La «resolución» de una paradoja consiste en demostrar que la aparente doble paradoja es, de hecho, por alguna razón específica, solo aparente; es decir, ilusoria, no real. Para encontrar la resolución, generalmente se requiere examinar con sumo cuidado el discurso subyacente o circundante y descubrir dónde se cometió el error. Es posible que se haya confiado en una noción demasiado vaga, caído en la ambigüedad, hecho una suposición injustificada, etc. La resolución de una paradoja no suele ser un proceso mecánico, sino que requiere considerable perspicacia y reflexión. Las principales paradojas distan mucho de ser fáciles de disolver; cada una requiere una consideración y un tratamiento especial. A menudo, después de ser descubiertas por algún lógico o filósofo, permanecen sin resolver durante mucho tiempo.

Respecto al orden de esta colección, quisiera hacer la siguiente aclaración. En mis compilaciones temáticas anteriores (a saber, *Las leyes del pensamiento* , *El yo* , *Ética* y *Teología*), generalmente he optado por una presentación cronológica. Sin embargo, en este caso, el orden se ha determinado tanto por consideraciones didácticas como cronológicas.

- He incluido como capítulo inicial el ensayo «La vanidad del tetralema», extraído de mi obra *A Fortiori Logic* (2013). Los dos capítulos siguientes, «Aclarando la contradicción» y «Aclarando la negación», provienen de mi libro *Ruminations* (2005). Estos tres primeros ensayos sirven, desde el principio, para centrar la atención en las leyes del pensamiento y reafirmarlas.
- A continuación, inserto, a modo de introducción al tema de las paradojas, el ensayo “Paradojas”, extraído de mi obra más temprana, *Lógica del futuro* (1990).
- A continuación, se presentan algunas reflexiones iniciales sobre las paradojas en general y la paradoja

del mentiroso en particular, en el ensayo «La paradoja del mentiroso (inicial)», extraído de *Lógica futura* y *reflexiones* . En este punto, inserto mi análisis más avanzado de esta paradoja en «La paradoja del mentiroso (revisión)», extraído de *Lógica más acentuada* .

- A continuación, en el ensayo «La paradoja de Russell (inicial)», también extraído de *Future Logic* and *Ruminations* , presento algunas reflexiones iniciales sobre la paradoja de Russell, así como sobre ciertas variantes, como la paradoja del barbero, la paradoja del catálogo de libros y la paradoja de Grelling. Posteriormente, inserto mi análisis más avanzado de esta paradoja en «La paradoja de Russell (revisada)», extraído de *A Fortiori Logic* .
- Posteriormente, he incluido los ensayos «La paradoja de la confirmación de Hempel» y «La paradoja de la predicción de Goodman», ambos extraídos de mis *Reflexiones lógicas y espirituales* (2008-2009). A continuación, los ensayos «La paradoja de Sorites es artificial» y «Protágoras contra Euathlus resuelto», ambos extraídos de un libro que aún no he terminado ni publicado (título provisional: «Temas de lógica, filosofía y espiritualidad», publicado parcialmente en línea en 2017).
- Finalmente, el presente volumen comprende dos ensayos, “Discurso antinómico budista” y “Más discurso antinómico budista”, extraídos respectivamente de mi libro *La ilógica budista* (2002) y de Reflexiones *lógicas* y *espirituales* .

En cuanto a la importancia relativa de las paradojas aquí presentadas y tratadas, cabe decir lo siguiente.

En mi opinión, la paradoja del mentiroso y la paradoja de Russell son especialmente importantes, dada su posible repercusión en la teoría de la lógica; y recomiendo

particularmente mis reflexiones más recientes al respecto (en los ensayos titulados «revisión»). Si bien ya había tratado estas dos paradojas en escritos anteriores (aquí titulados «iniciales»), en 2013, cuando me encontraba en la fase final de la redacción de mi libro *A Fortiori Logic* [2], logré profundizar mucho más en ellas. En aquel momento me encontraba en mi mejor momento, intelectualmente experimentado y maduro, y por lo tanto podía ver las cosas con mucha más claridad que nunca.

En el ensayo revisado sobre la paradoja del mentiroso, demuestro que esta paradoja es una hidra de muchas cabezas, que no puede explicarse únicamente mediante la autorreferencia, como se suele hacer, sino que implica una variedad de problemas que deben abordarse. En el ensayo revisado sobre la paradoja de Russell, demuestro que este enigma de la lógica de clases no se debe, como se suele suponer, a alguna dificultad con la idea de *no* pertenencia a uno mismo, sino, por el contrario, a la imposibilidad de la pertenencia *a uno mismo* ; esta es una resolución nueva y definitiva de la paradoja, y detallo sus múltiples implicaciones.

Las paradojas modernas de Hempel y Goodman también son significativas para la teoría lógica, ya que plantean dudas sobre la inducción; sin embargo, su resolución es relativamente sencilla. La antigua paradoja de Sorites también resulta lógicamente interesante, pues demuestra la importancia de la precisión intelectual y verbal en el discurso; pero, como paradoja, se resuelve fácilmente. La anterior paradoja de Protágoras-Euathlus no tiene gran importancia lógica, salvo quizás para enseñarnos que, al intentar resolver paradojas, no debemos enredarnos en

---

[2] De hecho, estos ensayos no tienen nada que ver con un argumento fortiori; pero casualmente los escribí cuando ese libro estaba casi terminado, y quería publicarlos lo antes posible, así que los incluí en un apéndice.

cuestiones lógicamente irrelevantes (en este caso, cuestiones legales).

En lo que respecta al discurso antinómico budista, abordarlo es importante en este momento de la historia, porque el budismo está de moda en algunos círculos y algunas de sus doctrinas se reciben allí con un espíritu demasiado dogmático.

Si en el futuro escribo ensayos adicionales sobre paradojas, espero incluirlos en ediciones ampliadas de este volumen.

1. La vanidad del tetralema

De A Fortiori Logic, Apéndice 7.3.

El ataque más radical a la razón consiste en intentar poner en duda las leyes del pensamiento, pues estas son, en efecto, los fundamentos de todo discurso racional. Primero, se niega la ley de identidad al afirmar que las cosas nunca son del todo lo que parecen ser, o que lo que son está más cerca del gris que del blanco y el negro. Esto es, por supuesto, una observación absurda, ya que en sí misma pretende ser una certeza y claridad absolutas. Luego, se niegan las leyes de no contradicción y del tercero excluido [3]al afirmar que las cosas pueden ser y no ser a la vez, o ni ser ni no ser. Este es el «tetralema», la lógica cuádruple que se favorece en las filosofías india y china, en el misticismo religioso, y a la que recurren cada vez más algunos «científicos». Para comprender la vanidad del tetralema, es necesario entender la naturaleza de la negación y su papel como uno de los fundamentos de la lógica humana.

1. Los fenómenos son positivos

Lo primero que debemos entender es que todo lo que experimentamos es *un fenómeno positivo.* Todo lo que percibimos a través de nuestros sentidos, recordamos, imaginamos o incluso comprendemos mediante la intuición, debe tener algún tipo de contenido para poder

[3] Sin darme cuenta, omití las palabras " y del tercero excluido " en las ediciones originales de *A Fortiori Logic* , aunque del resto de la frase se desprende claramente que mi intención era incluirlas.

ser percibido. Cada órgano sensorial es una ventana a un tipo distinto de fenómeno positivo. Vemos el cielo azul, oímos cantar a los pájaros, olemos el aire fresco, saboreamos una fruta, sentimos la textura y el calor de la tierra, etc. De manera similar, las imágenes y los sonidos en nuestra mente, ya provengan de la memoria o sean producidos por la imaginación, son fenómenos positivos; e incluso los objetos de la intuición deben tener algún contenido que podamos comprender. En segundo lugar, debemos darnos cuenta de que muchos fenómenos positivos pueden aparecer simultáneamente en el espacio en un momento dado. Esto es cierto para cada tipo de fenómeno. Así, el cielo azul puede ocupar solo una parte de nuestro campo de visión, rodeado de árboles verdes y edificios grises; podemos oír a la vez el canto de los pájaros y el sonido de los coches; y así sucesivamente. En tercer lugar, muchos fenómenos positivos pueden compartir el espacio que percibimos en un momento dado. Así, superpuestos a fenómenos visuales como el cielo, pueden aparecer otros tipos de fenómenos: el canto de los pájaros en los árboles, el olor del tráfico en las calles, las sensaciones corporales, etc. Incluso podemos alucinar, proyectando objetos de percepción mental en el espacio físico. Por ejemplo, la imagen de las gafas puede persistir durante un tiempo después de quitárselas. En cuarto lugar, cada fenómeno positivo, sea cual sea su tipo, varía con el tiempo, con mayor o menor rapidez. Así, el cielo azul puede tornarse rojo u oscuro, el canto de los pájaros o el ruido del tráfico pueden aumentar, disminuir o incluso cesar por un tiempo, y así sucesivamente.

2. No hay fenómenos negativos

Para expresar todas estas posibilidades perceptivas —diferencias en el espacio, en el tiempo y en otros aspectos— necesitamos el concepto de negación, o más

precisamente, el acto de negar. Sin la negación, no podemos comprender el mundo de manera racional; es el punto de partida del ordenamiento lógico de nuestra experiencia. Así, en un campo visual determinado, donde aparecen (por ejemplo) el cielo azul y los árboles, para poder decir «el cielo termina aquí, donde empiezan los árboles» necesitamos la idea de negación; es decir, que a un lado de un límite el cielo es visible y al otro no, mientras que al primer lado los árboles no son visibles y al otro sí. Del mismo modo, con respecto al tiempo, para poder describir un cambio, por ejemplo, de cielo azul a cielo rosa, necesitamos la idea de negación; es decir, que antes esta parte del cielo era azul y no rosa, y después era rosa y no azul. Nuevamente, necesitamos la idea de «negación» para expresar diferencias en otros aspectos; por ejemplo, para decir que «el canto de los pájaros parece provenir de los árboles, y no de los edificios». Así, la negación es una de las primeras herramientas de la lógica, que entra en juego ya en la clasificación de las experiencias.

Además, la negación sigue desempeñando un papel central cuando empezamos a trabajar con abstracciones. El conocimiento conceptual, que consiste en términos y proposiciones basados directa o indirectamente en fenómenos perceptivos, se fundamenta, en primer lugar, en nuestra capacidad para reconocer similitudes entre objetos de percepción: «esto parece parecerse a aquello en cierto modo». Así, proyectamos mentalmente la idea de que esto y aquello «tienen algo en común», una propiedad común abstracta (es decir, no fenoménica, no percibida de ninguna manera), a la que podríamos asignar un nombre. Sin embargo, para llevar este proceso conceptual más allá, debemos ser capaces de negar, es decir, de afirmar que «ciertas cosas distintas de esto y aquello no poseen la propiedad común abstracta que esto y aquello parecen tener», o que «esto y aquello no lo tienen todo en

común». Es decir, debemos ser capaces de afirmar no solo que una cosa se parece a otra de alguna manera, sino también que estas u otras cosas no se parecen entre sí de esa manera o de otra. Por lo tanto, la negación es esencial para dar sentido a la información también a nivel conceptual de la conciencia.

Ahora bien, ¿qué es la negación? Para responder a esta pregunta, primero debemos comprender que *no existen fenómenos negativos en el ámbito de la experiencia* . Todo lo que percibimos es un fenómeno positivo, pues si no lo fuera, obviamente no tendríamos nada que percibir. Solo podemos «percibir» un estado de cosas negativo definiendo mentalmente un estado de cosas positivo que debemos buscar, y luego buscándolo; si tras buscarlo diligentemente no lo encontramos, concluimos inductivamente que está «ausente», es decir, que «no está presente». Así, los fenómenos positivos preceden a los negativos, y no los siguen. La existencia precede lógicamente a la no existencia. Los fenómenos negativos son «fenómenos» solo metafóricamente, por analogía con los fenómenos positivos; en realidad, no lo son: no existen. La «negación» no es un concepto en el sentido de una abstracción de muchas experiencias particulares que comparten una propiedad determinada. La negación es una herramienta del observador pensante, como se describió anteriormente. Es un acto, una intención suya.

3. Un experimento mal interpretado

Para ilustrar la confusión que existe entre algunas personas —incluso entre algunos científicos— respecto a la negación, les presento el siguiente ejemplo extraído de * *El mayor espectáculo del mundo: La evidencia de la evolución* * , *de Richard Dawkins* [4]. En él, describe un

[4] Págs. 13-14.

experimento de Daniel J. Simons en el que se pide a un grupo de personas que vean un breve vídeo y observen cuántas veces se repite un determinado suceso; pero al final se les formula una pregunta completamente distinta: si notaron la presencia de un hombre disfrazado de gorila durante el vídeo, y la mayoría admite que no [5]. Según Dawkins, podemos inferir de este experimento que «el testimonio de un testigo ocular, la "observación directa", los "datos de la experiencia" son, o al menos pueden ser, completamente poco fiables».

Pero esta es una inferencia errónea a partir de los datos disponibles, porque confunde la experiencia positiva con la negativa. Las personas que vieron el video estaban demasiado ocupadas buscando lo que se les había pedido que observaran como para notar al gorila. Más tarde, cuando se les mostró el video por segunda vez, sí lo vieron. No hay razón para esperar que experimentemos realmente todo lo que se presenta a nuestros sentidos. Nuestras experiencias sensoriales son siempre, necesariamente, selectivas. La validez de la percepción sensorial como tal no se pone en duda por el alcance limitado de las percepciones sensoriales particulares. La prueba es que es *a través de una mayor percepción sensorial* que descubrimos lo que antes pasamos por alto. La no percepción de algo no constituye una percepción errónea, sino simplemente una percepción incompleta. «No vi X» no implica deductivamente «Vi la ausencia de X», aunque la repetición de lo primero tiende a implicar inductivamente lo segundo.

Un fenómeno negativo no es como un fenómeno positivo, algo que se puede percibir o intuir directamente. Una negación es necesariamente producto de la cognición indirecta, es decir, de un proceso inductivo

[5] El vídeo puede verse en: www.theinvisiblegorilla.com/videos.html .

(específicamente, aductivo). Mentalmente, planteamos la hipótesis de que tal o cual fenómeno positivo está ausente, y luego la comprobamos y confirmamos buscándolo repetidamente sin encontrarlo [6]. Si en algún momento encontráramos el fenómeno positivo, la hipótesis de la negación se rechazaría inmediatamente, ya que la fiabilidad de una negación es mucho menor que la de una experiencia positiva. Ni siquiera formularíamos la negación si ya hubiéramos percibido el fenómeno positivo en el pasado o en el presente. Y si la formuláramos, naturalmente nos retractaríamos si más tarde nos encontráramos con el fenómeno positivo. Por lo tanto, el contenido de los fenómenos negativos es necesariamente siempre hipotético, es decir, tentativo hasta cierto punto; nunca es firme y seguro como ocurre con los fenómenos positivos (experimentados).

Las afirmaciones negativas, al igual que las positivas, pueden ser correctas o incorrectas. Si se busca diligentemente un fenómeno positivo y no se encuentra, se puede afirmar lógicamente su negación. Esta afirmación es necesariamente inductiva: solo es válida mientras se busque activamente el fenómeno positivo y no se encuentre. En el momento en que se observa el fenómeno positivo, la negación deja de estar justificada. Si *no* se busca el fenómeno positivo, o no se busca con la debida diligencia, quizás debido a alguna distracción (como en el ejemplo citado anteriormente), entonces, por supuesto, la afirmación de negación es dudosa; ciertamente, es inductivamente débil, y es muy probable

---

[6] La falta de hallazgo es la ausencia del acto positivo de encontrar. Objetivamente, cabe destacar que la falta de hallazgo es en sí misma un fenómeno negativo, no positivo. Sin embargo, subjetivamente, puede experimentarse algo positivo en nuestro interior: tal vez una sensación de decepción o de alivio. Para más información sobre este tema , consulte el capítulo 9 de mis *Reflexiones* .

que se demuestre lo contrario mediante alguna observación posterior.

4. Definición de negación

¿Cómo se define, entonces, la negación? Podríamos decir que se define por las leyes de no contradicción y del tercero excluido. Es decir, con respecto a cualquier término 'X' y su negación 'no-X', la relación entre ellos es, por definición, la disyunción "o X o no-X", lo que aquí significa que estos términos (X y no-X) no pueden ser a la vez verdaderos ni falsos; es decir, son exclusivos y exhaustivos. ¿Qué entiendo aquí por "definición"? ¿Es un acto arbitrario? No, apunta a algo evidente; es intencional. Aquí, apunta al instrumento del discurso racional que necesitamos para ordenar la experiencia y producir derivados conceptuales coherentes a partir de ella. El instrumento necesario debe construirse de esta manera; cualquier otra construcción no cumpliría la función que necesitamos. Es decir, la única forma concebible de ordenar lógicamente nuestro conocimiento es mediante la negación, definida por las leyes de no contradicción y del tercero excluido. Sin esta herramienta, el análisis de la experiencia es imposible.

Supongamos ahora que alguien se presenta y, sin embargo, objeta la afirmación anterior. Bueno, dice, ¿cómo sabes *que* el dilema “o X o no X” es verdadero? Simplemente *definiste* las cosas arbitrariamente de esa manera, ¡pero eso no significa que sea un hecho! ¿No podríamos afirmar igualmente que el tetralema “o X o no X *o ambos o ninguno* ” es verdadero? La respuesta a esa objeción es muy simple. Supongamos que acepto esta crítica y estoy de acuerdo con el tetralema. Ahora, dividamos esta disyunción cuádruple, colocando a un lado la alternativa simple “X” y al otro lado la triple alternativa “no X o ambos o ninguno”. Ahora tengo nuevamente un

dilema, a saber, “o X o no X o ambos o ninguno”. Permítanme ahora definir un nuevo concepto de negación sobre esta base, de modo que obtengamos una disyunción de dos alternativas en lugar de cuatro. Llamemos a la segunda alternativa compleja 'no-X o ambos o ninguno' de esta disyunción 'NO-X' y llamémosla 'la supernegación de X'.

Así pues, ahora, el objetor y yo coincidimos en que la disyunción «o X o NO X» es exclusiva y exhaustiva. Coincidimos, presumiblemente, en que este nuevo dilema no puede ser refutado a su vez por un tetralema de la forma «O X o NO X o ambos o ninguno», pues si se intentara de nuevo tal oposición, sin duda se podría contrarrestar con otra división y redefinición. No podemos repetir razonablemente ese proceso hasta el infinito; hacerlo equivaldría a bloquear para siempre todo pensamiento racional. Habiendo bloqueado así todas las vías de pensamiento, el objetor no podría alegar tener un pensamiento mejor, ni ningún pensamiento en absoluto. Por lo tanto, no hay provecho en seguir objetando. Así pues, el tetralema es simplemente una provocación, pues pudimos esquivar el golpe. Habiendo llegado a un acuerdo de que la nueva disyunción «O X o NO X» es lógicamente irrefutable, debemos admitir que la disyunción original «O X o no X» era lógicamente sólida desde el principio. Porque puedo decirles que lo que yo entendía por no-X, o la 'negación de X', era desde el principio lo que ahora se entiende por NO-X, o la 'supernegación de X'.

Nunca me interesó una negación relativa y débil, sino que desde el principio busqué una negación absoluta y contundente. Pues esta negación total, y nada menos radical, es la herramienta que todos necesitamos para ordenar la experiencia y desarrollar el conocimiento conceptual de manera coherente y eficaz. En otras palabras, sea cual sea la versión más débil de la negación

que alguien intente inventar [7], aún podemos proponer una versión fuerte tal que tanto la ley de no contradicción como la del tercero excluido le sean aplicables sin duda. Si tal negación no existiera, habría que inventarla. Nadie puede destruirla negándola o diluyéndola. Quienes lo intentan son meros sofistas que no comprenden el origen, la naturaleza y la función de la negación en el discurso humano. Creen que se trata de una manipulación simbólica y no se dan cuenta de que su papel en el discurso humano es mucho más fundamental y complejo. La negación es el instrumento indispensable para cualquier intento de conocimiento que vaya más allá de la mera percepción.

[7] Hay quienes afirman que el principio de no contradicción es lógicamente necesario, pero el principio del tercero excluido no lo es. Evidentemente, esta afirmación puede refutarse del mismo modo. Si plantean las tres alternativas «O X o no X o ni X ni no X», podemos dividir la disyunción en dos, con «X» por un lado y «no X o ni X ni no X» por el otro, y proceder como hicimos con el tetralema. Lo mismo se aplica si alguien acepta el principio del tercero excluido pero rechaza el principio de no contradicción. Todos estos intentos son falaces absurdos.

2. Aclarar la contradicción

De las Reflexiones 1, 5 y 9.

1. Dialéctico

Las tres "Leyes del Pensamiento" pueden explicarse brevemente de la siguiente manera:

(i) *Tesis* : existen ciertas apariencias; las apariencias aparecen.

(ii) *Antítesis* : existen incompatibilidades entre algunas de estas apariencias; en tales casos, una o ambas deben ser falsas.

(iii) *Síntesis* : algunas de las apariencias restantes deben ser ciertas; ¡descubre cuáles!

Desde esta perspectiva, podemos considerar *la dialéctica* como una forma fundamental de pensamiento, a través de la cual el conocimiento progresa continuamente. No se trata de un mero detalle, un proceso mental ocasional, sino de una fuerza motriz, un motor del pensamiento.

Las leyes no son mera información, sino llamamientos a la acción cognitiva. Exigen medidas cognitivas proactivas y correctivas para garantizar (en la medida de lo posible en cada momento) la verificación continua, la coherencia y la exhaustividad.

(i) La ley de identidad nos insta a buscar los hechos y analizarlos lo mejor posible. Su propósito es inculcar en las personas un sano respeto por los hechos, mediante la observación y el juicio. Es, en esencia, un llamado a la honestidad y a la sumisión al veredicto de la verdad. A

menudo, las personas piensan, o actúan como si pensaran, que ignorar o negar hechos o argumentos desagradables hará que desaparezcan; la ley de identidad afirma: «No, no desaparecerán, debes tenerlos en cuenta».

Algunos creen que es imposible ignorar que «A es A». ¡Nada más lejos de la realidad! Todos lo hacemos con frecuencia: cuando nos negamos a ver o admitir la evidencia o una demostración lógica; cuando evitamos la realidad o la eludimos tras vislumbrarla; cuando nos mentimos a nosotros mismos o a los demás; y así sucesivamente. Si siempre acatáramos la ley de identidad, no habría necesidad de formularla. La lógica afirma lo obvio, porque a menudo se ignora.

(ii) Cuando la ley de no contradicción nos dice que "no se puede afirmar y negar una proposición al mismo tiempo", también nos está diciendo que si alguna vez en el curso del discurso nos encontramos con una situación en la que una proposición parece a la vez verdadera (por alguna razón) y falsa (por otras razones), ***tenemos que volver al principio de nuestro discurso y averiguar dónde nos equivocamos en el curso del mismo*** [8]***, y tenemos que efectuar una corrección apropiada para eliminar la dificultad*** .

No nos limitamos a decir: «¡Ah, hay una contradicción!», y quedarnos ahí, sin inmutarnos. Al contrario, nos vemos impulsados a buscar una solución al problema, es decir, a resolver la contradicción. Inferimos que debe haber algún error en nuestro razonamiento previo que nos llevó a esta incógnita, algún error de observación o de razonamiento que requiere corrección. Mientras se tolere esta situación y no podamos identificar la fuente del error, la credibilidad de todo el conocimiento relacionado se verá proporcionalmente mermada. La coherencia debe

[8] "Revisa tus instalaciones", solía decir Ayn Rand.

restablecerse cuanto antes, o corremos el riesgo de poner en duda todo el conocimiento posterior.

(iii) De igual modo, la ley del tercero excluido no solo nos informa que «ninguna proposición puede considerarse ni verdadera ni falsa». Esta ley insiste en que, si nos encontramos en tal situación, y resulta que tanto una proposición como su negación exacta parecen falsas, no podemos dejar el asunto en suspenso ni esperar encontrar una posición de compromiso; debemos, en última instancia, encontrar lo antes posible una buena razón para decantarnos por una u otra opción. No existe un punto medio lógicamente aceptable, ni una vía de escape.

Estas implicaciones prácticas inherentes a las leyes del pensamiento también pueden caracterizarse como pensamiento dialéctico. Desde esta perspectiva, la «tesis» es nuestro conocimiento (u opinión) en un momento dado; la «antítesis» es el descubrimiento de una falla lógica en esa tesis, lo que nos lleva a dudar de ella y a buscar su revisión; y, finalmente, la «síntesis» son las correcciones que introducimos en nuestras premisas para resolver la dificultad encontrada y obtener un nuevo estado de conocimiento menos problemático.

2. Contradicción

Mucha gente malinterpreta lo que los lógicos entendemos por « **contradicción** ». La contradicción de un término «A» es su negación, «no A», que se refiere a todo lo que existe en el universo distinto de A, es decir, a cualquier lugar donde A esté ausente. La relación de contradicción entre A y no A es mutua, reversible y perfectamente simétrica.

La presencia de algo (A) excluye su ausencia (es decir, no A) en esa misma cosa, y viceversa, si todas las coordenadas de espacio y tiempo son idénticas. Sin

embargo, esto no excluye la posibilidad lógica de que una misma cosa sea en parte A y en parte no A. Por lo tanto, la ley del pensamiento «o A o no A» también puede expresarse de forma más cuantitativa como «o bien "todo A" o bien "todo no A" o bien "parte A y parte no A"».

¡Algunos recurren a esta posibilidad de tres alternativas como argumento *contra* las leyes del pensamiento! Pero eso es un malentendido, o peor aún, una sofistería deliberada.

Si algo, por ejemplo 'B', implica pero no es implicado por no-A, entonces (es decir, B) es tan 'incompatible' con A como lo es no-A, pero no es contradictorio con A: es simplemente *contrario* a A. El no-A contradictorio de A se diferencia de los contrarios de A en que *la ausencia* de no-A implica A, mientras que en el caso de meros contrarios como B (o B1 o B2… etc.) esta relación lógica adicional de 'exhaustividad' no se aplica.

Cuando se colocan contradictorios en una disyunción, «o A o no A», la disyunción implica tanto exclusión mutua («o», que significa «no juntos») como exhaustividad («o», que significa «y no hay otra alternativa»). Su significado es: si «A», entonces no «no A»; y si no «A», entonces «no A».

Por otro lado, cualquier número de contrarios puede colocarse en una disyunción: «A o B o B1 o B2… etc.», de modo que la presencia de cualquier disyunción implica la ausencia de todas las demás; pero dicha disyunción no es exhaustiva, a menos que especifiquemos que la lista de contrarios que contiene es completa. Si esa lista *es* completa, entonces la negación de todos los disyuntos excepto uno implica la afirmación del restante. Así, «no-A» puede equipararse a la disyunción exhaustiva de todas las cosas en el mundo «contrarias a A».

Algo *distinto* de A, por ejemplo C, no es necesariamente contradictorio ni contrario a A. **El mero hecho de la**

diferencia no implica incompatibilidad . Cosas diferentes (como A y C) pueden ser compatibles, es decir, capaces de coexistir en la misma cosa, al mismo tiempo y en el mismo lugar. La «diferencia» simplemente significa que podemos *distinguir* entre las cosas en cuestión: es decir, no son una y la misma cuando se presentan ante nuestra conciencia. Las cosas «similares» pueden tener la misma apariencia, pero no ser una sola (por ejemplo, dos instancias del mismo tipo); o pueden ser una sola (es decir, partes de un mismo todo), pero no ser la misma.

Así, por ejemplo, la relación lógica entre los colores blanco y negro depende de la precisión con que nos enfoquemos en ellos. Son diferentes, puesto que se pueden distinguir. Dado que pueden coexistir en distintas partes de la misma superficie, son compatibles en términos generales. Sin embargo, en sí mismos , son contrarios; es decir: si yo percibo una superficie o parte de ella como totalmente blanca, y usted percibe *el mismo* lugar y momento como totalmente negro, nuestras afirmaciones son incompatibles [9]. Esta irreconciliabilidad no es una contradicción, ya que es posible que una superficie no sea ni blanca ni negra.

La expresión « **contradicción en los términos** » se refiere a un término compuesto por elementos incompatibles, como «A y no A» o «A y B (donde B es contrario a A)». Se puede decir que un término tan confuso es paradójico e internamente inconsistente, ya que implica que la contradicción es posible, de modo que niega las leyes del pensamiento y, por lo tanto (generalizando, si se quiere), «todo vale», incluyendo la negación de la conjunción «A y no A».

---

[9] Cabe señalar que nuestro desacuerdo no es terminológico. Anteriormente habíamos coincidido en cuanto a las experiencias "blancas" y "negras"; ahora, de repente, divergimos.

Por ejemplo, el término «realidad ilusoria» es una contradicción en sí misma. Por otro lado, cabe señalar que términos como «un ser humano inhumano» o «un judío antisemita» no son, en sentido estricto, contradictorios; se refieren a posibilidades naturales de conjunción, solo que la terminología empleada los hace parecer superficialmente contradictorios (es decir, hay personas que se comportan de forma inhumana o judíos que odian a su propio pueblo).

La proposición «A no es A» (o «algo que es A tampoco es A»), al ser contradictoria en sí misma, implica «A es A», su forma contradictoria. Esta afirmación debe reconocerse explícitamente, aunque resulte obvia, porque relaciona dos conceptos importantes: la «inconsistencia interna» y la «lógica de las paradojas».

La afirmación «A no es A» es lógicamente imposible, porque afirma y niega lo mismo. Por lo tanto, la afirmación opuesta es verdadera. Dicha afirmación, es decir, «A es A», es lógicamente necesaria, porque *incluso su contradictoria* «A no es A» la implica.

Quien afirma que "A no es A" está admitiendo que "A es A" – ***ipse dixit*** , ¡él mismo lo dijo! En cambio, quien afirma que "A es A" es coherente consigo mismo.

La autocontradicción consta de tres elementos:

(i) La proposición en cuestión, llamémosla P.

(ii) La admisión de que se trata de *una declaración asertiva* , es decir, una que afirma o niega algo.

(iii) La admisión de que todas las afirmaciones asertivas *implican pretensiones* de consciencia, de conocimiento, de verdad, etc.

Así pues, dada P (por ejemplo, «la realidad es incognoscible»), admitamos que P implica «esto es una afirmación»; pero toda afirmación implica cierto conocimiento de la realidad; *por lo tanto* , P implica no-P.

Existe un proceso que va de P a su negación, algo que la lógica exige que reconozcamos. Esta exigencia no puede rechazarse sin incurrir en la misma autocontradicción. No se trata de una demostración circular o ad infinitum, sino de una apelación a la honestidad, sin la cual ningún diálogo es posible.

Que todas las proposiciones asertivas afirmen es un aspecto de la Ley de Identidad. La Ley de No Contradicción puede discernirse en el argumento: Todas las afirmaciones afirman algo; P es una afirmación; por lo tanto, P afirma; de ahí que, si P niega afirmar, P implica no-P. La Ley del Tercero Excluido también está implícita aquí, en la conciencia de que no tenemos más remedio que rechazar firmemente P.

Contradicciones disfrazadas . Las contradicciones aparecen en el discurso de muchas maneras. No siempre son evidentes, sino que pueden estar ocultas en el hecho mismo de hacer una afirmación o en los criterios de juicio utilizados.

Una afirmación puede ser paradójica porque **implica inherentemente** su propia contradicción, aunque superficialmente no parezca ser inconsistente. Esta implicación no siempre es formal, sino que requiere comprender el significado de los términos utilizados. Esta forma de autocontradicción indirecta se conoce como la falacia del concepto robado [10].

Por ejemplo, la afirmación escéptica "No sé nada" puede rechazarse por ser contradictoria, porque en cuanto alguien la hace —alguien *que entiende y pretende el significado* de los términos "yo", "sé" y "nada"— eso es prueba absoluta de que la persona en cuestión "sabe" *algo* , por lo que la afirmación original (de ignorancia total) se revela inevitablemente contradictoria y, por lo tanto, necesariamente falsa.

---

10 Por Ayn Rand y (creo) Nathaniel Branden.

Así pues, en casos de este tipo, la implicación tácita es que uno de los términos utilizados (no saber nada) incluye implícitamente el acto en cuestión (saber que no sé nada), como ejemplo que contradice la afirmación explícita. (Huelga decir que reformular dicha afirmación como «no *sé* nada» no altera sus supuestos subyacentes).

Existen innumerables ejemplos de esta contradicción inherente. Decir «No tengo nada que decir» es decir algo. Afirmar «No tenemos memoria» es contradictorio, porque cada término presupone una palabra, un concepto y experiencias previas recordadas tanto por quien habla como por quien escucha. Un ejemplo común y divertido es: «¡No hablo ni una palabra de inglés!».

Otra forma importante de autoinconsistencia encubierta es el uso de **un doble rasero** . Esto consiste en aplicar criterios de juicio menos rigurosos al propio discurso que al de los oponentes intelectuales. Gran parte del discurso filosófico, y en particular del político y religioso, recurre a esta metodología inequitativa.

La contradicción que implica un doble rasero se hace evidente en cuanto analizamos el conocimiento y la metodología de quien lo utiliza en su conjunto. Desde esta perspectiva más amplia, quien emplea un doble rasero es claramente incoherente consigo mismo, aunque su discurso, visto fragmentado, pueda parecer superficialmente coherente.

Filosofías enteras pueden basarse en razonamientos tan falaces. Por ejemplo, el fenomenalismo establece como norma general la limitación del conocimiento a los datos sensoriales, sin permitir extrapolaciones a partir de ellos a supuestos objetos materiales externos; sin embargo, no critica sus propias aducciones utilizando la misma norma rígida.

Esta falacia puede cometerse de dos maneras: uno puede aplicar criterios laxos a su propio discurso, mientras que

aparentemente aplica normas universales al discurso de sus oponentes; o uno puede aparentar aplicar normas universales a sí mismo, mientras que inventa normas excesivamente estrictas para ellos. Uno puede *eximirse* de las reglas lógicas habituales, o puede *exigir lógicas inusuales a los demás* .

En ambos casos, quien defiende un doble rasero entra en conflicto con la exigencia de uniformidad propia de la lógica. La razón parte de la premisa de que todos los seres humanos se encuentran en el mismo plano epistemológico. La equidad es un aspecto del sentido común. La experiencia y la lógica deben utilizarse para convencerse a uno mismo y a los demás, no la manipulación sofística ni la autoridad.

Los criterios de juicio deben ser justos y universales; todo discurso debe recibir el mismo trato. Si se defienden las diferencias, deben justificarse de manera convincente. Si bien el principio de igualdad implica generalización, la carga de la prueba recae sobre cualquier particularización que se proponga.

Un ejemplo de doble rasero es el recurso al relativismo cultural. Se puede intentar racionalizar ideas o procesos de pensamiento contrarios a la razón ordinaria, alegando que pertenecen a un marco cultural diferente. Esta tolerancia parece, en apariencia, amistosa y abierta, pero se propone sin considerar plenamente sus implicaciones negativas, tanto humanas como epistemológicas.

3. La coherencia es natural.

Es importante reiterar aquí el principio de que la coherencia es natural , mientras que la incoherencia es excepcional.

Algunos lógicos modernos han propuesto la noción de «demostrar consistencia», pero esta idea es errónea. La

consistencia es el estado natural del conocimiento; no requiere prueba (deductiva) y somos incapaces de proporcionarla, ya que sería como «poner el carro delante de los bueyes». La única «prueba» posible de consistencia es que no se haya encontrado ninguna inconsistencia. La consistencia es un hecho inductivo que rara vez se refuta. Todo nuestro conocimiento puede y debe asumirse consistente, a menos que exista una razón para creer lo contrario.

En resumen: la armonía suele reinar sin que nos demos cuenta, mientras que los conflictos estallan ocasionalmente para nuestra sorpresa. Cabe preguntarse ahora si este principio es coherente con el principio aquí defendido de que las negaciones nunca son, en sí mismas, objetos de conocimiento, sino que solo existen por negación de las positivas correspondientes. Nuestro principio de que la coherencia se da por sentada parece implicar que, en ocasiones, tenemos intuiciones lógicas de *coherencia* , ¡algo negativo!

Para resolver este problema, debemos recalcar nuevamente la distinción entre la experiencia pura y las *interpretaciones* de la experiencia que, de forma implícita (por mera intención) o explícita, incorporamos habitualmente a nuestras vivencias. Generalmente, casi inmediatamente después de experimentar algo, comenzamos a interpretarlo, relacionándolo dinámicamente con el resto de nuestro conocimiento hasta el momento. Casi inevitablemente, toda experiencia genera en nosotros una serie de asociaciones, explicaciones, etc.

Las contradicciones que a veces encontramos en nuestro conocimiento no atañen a nuestras experiencias puras (que son necesariamente armoniosas, puesto que coexisten, podríamos añadir, de forma bastante «feliz»). *Nuestras contradicciones son necesariamente*

contradicciones entre una interpretación y una experiencia pura, o entre dos interpretaciones . Estrictamente hablando, las contradicciones no revelan dificultades en los datos brutos del conocimiento, sino simplemente en las hipótesis que hemos formulado sobre dichos datos.

Las contradicciones, por lo tanto, se deben a la razón, no a la experiencia. Esto no significa que la razón sea necesariamente defectuosa, sino solo que es falible. Las contradicciones no deben considerarse pruebas trágicas de nuestra ignorancia y estupidez, sino indicadores útiles de que hemos malinterpretado algo y que es necesario reinterpretarlo. Estos indicadores son precisamente una de las principales herramientas que utiliza la facultad de la razón para controlar la calidad de las creencias. La resolución de una contradicción no es más que una nueva interpretación.

Cómo sabemos que dos teorías, o una teoría y ciertos datos, se contradicen entre sí es una cuestión controvertida. Solemos desestimar esta pregunta con demasiada facilidad recurriendo a la «comprensión lógica». Dicha comprensión es en parte «experiencial», puesto que se basa en el análisis de las pruebas y doctrinas disponibles. Pero, evidentemente, no es del todo empírica e implica factores abstractos. Al fin y al cabo, la «contradicción» es una abstracción. Creo que la respuesta a esta pregunta se encuentra en gran medida en el análisis psicológico de la negación.

Existe una intuición que sugiere la existencia *de intenciones contradictorias* . Por lo tanto, la «comprensión lógica» de la inconsistencia no se refiere esencialmente a una negatividad (una falta de coherencia), sino a una positividad (la experiencia intuitiva del conflicto de intenciones). Si bien la palabra «inconsistencia» implica un prefijo negativo, evoca algo

empíricamente positivo: una tensión percibida entre dos tesis o entre una tesis y ciertos datos.

Por esta razón, afirmar que «la consistencia es asumible, hasta que se encuentre alguna inconsistencia» es coherente con nuestra afirmación de que «las negaciones no son puramente empíricas». (Cabe señalar, de paso, que aquí no «demostramos» la consistencia, sino que simplemente la *recuperamos* al aclarar las tesis implicadas).

El análisis anterior también aclara cómo se expresa en la práctica el principio de no contradicción. No se centra en las experiencias en sí mismas, sino en conceptos más abstractos del conocimiento. Para comprenderlo plenamente, debemos tener en cuenta las intenciones subyacentes. Un análisis similar puede proponerse para explicar el principio del tercero excluido.

En este último caso, insistiríamos en que (por la ley de identidad) «las cosas son algo, lo que son, sea lo que sea». No se puede decir que las cosas no sean *ni* esto *ni* la negación de esto, porque tales caracterizaciones son negativas (y, respectivamente, doblemente negativas) y, por lo tanto, no pueden constituir ni considerarse experiencia positiva. Estas situaciones remiten a incertidumbres *en el sujeto cognoscente* , que este debe completar eventualmente. No pueden proclamarse conocimiento definitivo (como han intentado hacer algunos sofistas modernos), sino que deben considerarse posturas temporales en la búsqueda del conocimiento.

3. Negación aclaratoria

De Reflexiones 9.

1. Negación en la aducción

Los conceptos y las teorías son construcciones hipotéticas. No pueden (en su mayoría) probarse (de forma definitiva, de una vez por todas), sino que solo se confirman repetidamente mediante la experiencia. Este es el aspecto positivo de la aducción: presentar evidencia que respalde las construcciones racionales. Este aspecto positivo es, por supuesto, indispensable, pues sin alguna evidencia concreta, una abstracción no es más que un producto de la imaginación, una mera especulación. Cuanta más evidencia presentemos, más fiable será nuestro concepto o teoría.

Pero, como bien comprendió Francis Bacon, la explicación de la aducción propuesta hasta ahora no le hace justicia. Tan importante como el aspecto positivo de aportar pruebas es su aspecto negativo: el rechazo de hipótesis que hacen predicciones que contradicen la experiencia. Como señaló, incluso si una hipótesis cuenta con numerosas confirmaciones, basta con que contenga *una sola* predicción errónea para que sea rechazada.

En resumen, esto significa que el proceso de aducción se centra en la selección de la hipótesis más probable entre dos o más explicaciones de un hecho (ya existentes o por concebir). Cada una de ellas puede tener numerosos «ejemplos positivos» (es decir, evidencia empírica que la respalda); y mientras todas sean competitivas, podemos

preferir aquellas con la mayor cantidad de ejemplos. Sin embargo, el avance decisivo en nuestro conocimiento conceptual/teórico se produce mediante la *eliminación sucesiva* de proposiciones que resultan tener «ejemplos negativos» (es decir, evidencia empírica en su contra).

Todo lo anterior es bien conocido y no requiere mayor explicación. Esta teoría de la lógica inductiva ha demostrado ser sumamente eficaz en la actualidad, constituyendo la base del método científico.

Pero, pensándolo bien, ¡el asunto no es tan simple y directo como parece a primera vista!

Consideremos, por ejemplo, la cuestión de si hay o no agua en Marte. Parecería que la proposición «Hay agua en Marte» es mucho más fácil de demostrar inductivamente que la proposición contradictoria «No hay agua en Marte». Ambas proposiciones son hipótesis.

La hipótesis positiva se confirmaría parcialmente si, mediante el uso de ciertos instrumentos a distancia, se detectaran indicios serios de la presencia de agua; se confirmaría de forma más sólida si se trajera a la Tierra una muestra de Marte y, tras su análisis, se encontrara agua. En ambos casos, la presencia de agua en Marte seguiría siendo incierta en cierta medida, ya que podrían surgir objeciones a nuestras suposiciones instrumentales o se podría descubrir que la muestra traída de vuelta se contaminó durante el viaje. No obstante, es posible alcanzar un grado de certeza bastante cercano a este tema.

La tesis negativa, por el contrario, es mucho más difícil de probar mediante la experiencia. Podemos asumirla fácilmente en la medida en que la tesis positiva no se haya confirmado ampliamente hasta el momento. Es decir, mientras no hayamos encontrado evidencia de la tesis positiva (es decir, agua en Marte), deberíamos optar por la tesis negativa. Pero esta última solo es confiable en la medida en que hayamos intentado, sin éxito, confirmar la

primera. Si buscáramos agua con ahínco por todos los medios posibles y no la encontráramos, podríamos asumir con la confianza proporcional que no hay agua.

Así pues, en nuestro ejemplo, la tesis negativa es en realidad *más difícil* de establecer que la positiva. *Depende de una generalización* , un razonamiento que va desde «Dondequiera y comoquiera que buscáramos agua en Marte, ***no la encontramos*** » hasta « ***No hay*** agua en Marte». Sin embargo, cabe señalar que sigue siendo concebible que algún día se encuentre una gota de agua en algún otro lugar de Marte, siglos después de que hayamos concluido que no la había.

Partiendo de este análisis, resulta evidente que la máxima de Bacon de que «Lo importante es el ejemplo negativo» es un tanto simplista. Presupone que un ejemplo negativo es tan accesible como uno positivo (o incluso más accesible), lo cual no siempre es así.

En la práctica, una negación puede ser inductivamente más remota que una positiva. Partiendo de esta conclusión, surge la pregunta: ¿es la instancia negativa *alguna vez* más accesible empíricamente que la positiva (o incluso igual de accesible)? Es decir, ¿ *cuándo* entra en juego realmente la formulación de la inducción de Bacon?

Si analizamos ejemplos históricos importantes de rechazo de teorías, nuestra duda puede persistir. Por ejemplo, la mecánica newtoniana se mantuvo vigente durante siglos, hasta que el descubrimiento de la constancia de la velocidad de la luz (que dio origen a la teoría de la relatividad) y, posteriormente, el descubrimiento de diversos fenómenos subatómicos (que dieron origen a la mecánica cuántica) la pusieron en duda. En este ejemplo, los casos considerados «negativos» eran, en esencia, «positivos»: ¡lo único «negativo» radicaba en su negación de la visión newtoniana del mundo!

Estas reflexiones me han llevado a sospechar que la «negación» a la que se refiere Bacon solo se entiende *en relación* con una abstracción específica. Su navaja no debe interpretarse como una defensa de la negación absoluta. Si analizamos el asunto con mayor claridad, nos damos cuenta de que los datos utilizados para negar una idea son esencialmente positivos. Por lo tanto, resulta evidente la necesidad de una reflexión más profunda sobre la naturaleza de la negación.

2. Fenómenos positivos y negativos

Siempre se ha considerado que existe una diferencia entre un término positivo y uno negativo. De hecho, por eso los lógicos les han dado nombres distintos. Sin embargo, a los lógicos también les ha resultado difícil expresar esa diferencia de forma sustancial. Aun así, existen diferencias fenomenológicas significativas entre los fenómenos positivos y negativos.

a. El mundo material y mental concreto se compone, evidentemente, únicamente de fenómenos particulares positivos, algunos de los cuales percibimos (ya sea a través de los sentidos corporales o en nuestra mente). Estos existen al menos como apariencias, aunque algunos resultan ser reales y otros ilusorios. Esta es una verdad fenomenológica, epistemológica y ontológica evidente.

Decir que un fenómeno es «particular» implica reconocer que siempre está limitado en el espacio y el tiempo. Tiene presencia, pero es finito y transitorio, es decir, se caracteriza manifiestamente por la diversidad y el cambio.

Normalmente no experimentamos nada concreto que se extienda uniformemente hasta el infinito y la eternidad (aunque tal totalidad de la existencia bien podría existir, e incluso los místicos afirman alcanzar la conciencia de ella en meditación profunda, caracterizándola como «el

presente eterno»). Sí solemos considerar algunas cosas como tan generalizadas. La «existencia» se considera el sustrato de todos los entes; «el universo» se refiere a la suma total de todos los entes; y pensamos en el «espacio-tiempo» como la definición de la extensión de todos los entes. Pero solo la «existencia» puede clasificarse como una experiencia (una cualidad presente en todos los entes); «el universo» y «el espacio-tiempo» deben admitirse como abstracciones.

Sin embargo, los límites de lo particular se pueden percibir sin necesidad de negar lo que está más allá de ellos, simplemente debido a la concentración variable de la conciencia, es decir, la dirección del foco de atención. Es decir, aunque «señalar» algún fenómeno positivo (por ejemplo, para nombrarlo) requiere cierta negación (queremos decir «esto, pero no aquello»), se pueden percibir los límites de ese fenómeno independientemente de la negación.

b. Los fenómenos negativos (y de igual modo los abstractos, sean positivos o negativos), por otro lado, dependen para su existencia de un Sujeto/Agente: una "persona" cognoscente (o sinónimo: un yo, alma o espíritu) con conciencia y voluntad que busca algún fenómeno positivo recordado o imaginado y no logra percibirlo (o en el caso de los abstractos, compara y contrasta particulares).

Así, los fenómenos particulares negativos (y, en general, los abstractos) poseen una existencia especial, más «relativa». No son tan independientes del Sujeto como los fenómenos particulares positivos. Esto no significa que sean, en un sentido kantiano, «a priori» o «trascendentales», ni puramente «subjetivos», sino que son potenciales ontológicos que solo se realizan en el contexto del conocimiento (racional).

Para tal realización se requiere otro tipo de experiencia: la autoexperiencia del Sujeto, su conocimiento intuitivo de sus cogniciones y voliciones. Este tipo de experiencia, al ser inmediata, puede ser positiva o negativa sin dificultad lógica. El Sujeto razona inductivamente de la siguiente manera:

Estoy buscando X;

No encuentro X;

Por lo tanto, X " *no está"* allí.

La conclusión negativa puede ser «verdadera» o «falsa» , al igual que una percepción o conclusión positiva. Es verdadera en la medida en que las premisas sean verdaderas, es decir, que la supuesta búsqueda de X fue diligente (inteligente, imaginativa, bien organizada, atenta y minuciosa), y que el supuesto fracaso en encontrar X no es deshonesto (una mentira diseñada para engañarse a uno mismo o a los demás).

De ahí que sea justo afirmar que, a diferencia de algunos términos positivos, los términos negativos nunca se basan *únicamente* en la percepción; *necesariamente* implican un proceso de pensamiento: la proyección mental previa o al menos la intención del término positivo que niegan.

Esta verdad epistemológica refleja una verdad ontológica: la verdad de que las «ausencias» de fenómenos carecen de aspectos fenoménicos. Un «no» no es una especie de «sí».

Nótese la diferencia lógica entre « **no percibir X** » y « **percibir la ausencia de X** ». No tenemos experiencia directa de lo segundo, sino que solo podemos inferirlo indirectamente mediante *inferencia inductiva* (o extrapolación) a partir de lo primero. En el caso de una afirmación, este razonamiento no es necesario, ya que a menudo se puede «percibir X» directamente.

Supongamos que dibujamos un cuadrado de oposición para las proposiciones (etiquetadas por analogía con las

posiciones estándar): «Percibo X» (A), «No percibo X» (I), «No percibo X» (E), «No percibo X» (O). Aquí, la forma A es cognoscible por experiencia, mientras que la forma I es cognoscible quizás solo por implicación deductiva. Por otro lado, la forma E no es cognoscible por experiencia, sino solo por generalización inductiva a partir de la forma O (que se basa en la experiencia).

3. La negación es secundaria

La negación es un pilar tanto de la lógica deductiva como de la inductiva, y requiere un análisis cuidadoso. Debemos comprender que los términos negativos son fundamentalmente distintos de los positivos si queremos empezar a entender la naturaleza de la lógica. La siguiente observación me parece crucial para dicho análisis:

> ***Podemos experimentar algo positivo sin haber experimentado (o pensado) primero en su negación, pero no podemos experimentar algo negativo sin antes pensar en (y por lo tanto haber experimentado previamente de alguna manera) lo positivo correspondiente.***

a. La cognición, en su forma más simple, es percepción. Nuestras percepciones son siempre *de particularidades positivas* . El contenido de nuestras cogniciones más básicas son imágenes, sonidos, olores, sabores y tacto fenomenales, así como otras sensaciones corporales que aparentemente surgen a través de la interacción de nuestros órganos sensoriales con la materia, o equivalentes mentales de estos fenómenos que aparentemente surgen a través de la memoria de experiencias sensoriales, o en recombinaciones imaginarias de tales supuestos recuerdos.

Una experiencia positiva puede vivirse de forma directa y pasiva. Podemos simplemente relajarnos, por así decirlo,

y observar receptivamente lo que se presente en nuestro campo de visión o audición, etc. Esto es lo que hacemos en la meditación. No necesitamos pensar activamente (recordar, visualizar o conceptualizar) en otra cosa para tener una experiencia tan positiva. Por supuesto, en la práctica, dicha observación puede verse complicada por pensamientos (preverbales o verbales), pero en algunos casos es posible tener una experiencia pura. Esto debe admitirse lógicamente si los conceptos se basan en percepciones.

b. En el caso de *los particulares negativos* , la situación es radicalmente diferente. Un particular negativo *no tiene* un contenido fenoménico específico, sino que se define *enteramente* por la «ausencia» de los contenidos fenoménicos que constituyen algún particular positivo. Si observo mi entorno material o mental, siempre veré fenómenos presentes. La ausencia de algún fenómeno solo se hace evidente si primero pensamos en ese fenómeno positivo y nos preguntamos si está presente.

Es cierto que nuestra constatación de su ausencia refleja una verdad o un hecho empírico; pero también es un hecho que simplemente no percibiríamos lo negativo sin antes haber pensado en lo positivo. Por lo tanto, el conocimiento negativo es necesariamente (por necesidad lógica) más indirecto y activo. En el mejor de los casos, se fundamenta perfectamente en la experiencia, pero dicha experiencia negativa requiere un proceso racional (ya sea verbal o de otro tipo).

Para experimentar algo negativo, primero debo imaginar (recordar o inventar) una experiencia positiva; luego debo observar (o escuchar o lo que sea) si esta imagen coincide o no con mi experiencia actual; y solo entonces (si resulta que no coincide) puedo concluir que he “experimentado” algo negativo.

Pensar en X puede considerarse como posicionarse en un punto de vista desde el cual se puede (por así decirlo) experimentar lo que no es X. Si uno no centra primero su atención en X, es imposible experimentar su negación. Uno puede experimentar todo tipo de cosas extrañas y maravillosas, pero no específicamente lo que no es X.

A partir de esta reflexión, podemos afirmar que, si bien las afirmaciones pueden experimentarse, las negaciones son inherentemente actos racionales (que implican imaginación, experiencia e intención). Una negación implica necesariamente pensamiento: el pensamiento de la positiva correspondiente (el elemento imaginativo), la comprobación de su presencia o ausencia (el elemento experiencial) y la conclusión racional de "negación" (el elemento intencional).

c. El proceso de negación puede implicar palabras, aunque no necesariamente.

Supongamos que tengo una experiencia momentánea de imágenes, sonidos, etc., y la denomino "X". El *contenido de la conciencia* en el que baso el término X es un conjunto específico de experiencias fenomenológicas positivas, es decir, percepciones físicas y/o mentales. Siempre que puedo hablar de esta X, mentalmente imagino *un* objeto de cierto color y forma que se mueve de ciertas maneras, emite ciertos sonidos, etc.

Muy diferente es la negación de un término tan simple como «no X». Este último no se define por ninguna percepción específica; no *se refiere a cualidades perceptibles* . No puede identificarse con los fenómenos positivos que, por casualidad, están presentes en ausencia de aquellos que constituyen X. Por lo tanto, en sentido estricto, «no X» solo se define por la «negación» de X.

Cabe señalar que no sería exacto afirmar (salvo a posteriori) que "no-X" se refiere a todas las experiencias distintas de X (como Y, Z, A, B, etc.), porque cuando

busco X aquí y ahora y no la encuentro, me refiero únicamente a la experiencia presente dentro de mi ámbito actual y no a todas las posibles experiencias de este tipo. No etiquetaríamos una situación desprovista de X como "no X" *sin pensar en X* ; en cambio, la etiquetaríamos de forma positiva (como "Y", "Z" o cualquier otro término).

Así, podemos nombrar (o pensar sin palabras) algo concreto "X", *después de* experimentar los fenómenos que lo constituyen; pero en el caso de "no-X", necesariamente evocamos el nombre (o un pensamiento sin palabras) de ello *antes* de experimentarlo.

“No-X” es, por tanto, un concepto más que una percepción, incluso cuando “X” se refiere a una mera percepción (y más aún cuando “X” implica cierta abstracción, como suele ocurrir). El concepto “no-X” se construye hipotéticamente primero y luego se confirma mediante el intento fallido de revivir X.

En resumen, la negación —incluso en el nivel más perceptivo— implica un proceso aductivo. Nunca es una mera experiencia. Un término negativo nunca implica la simple percepción de algo negativo, sino que consiste en una hipótesis con alguna confirmación perceptiva. La negación siempre tiene un carácter conceptual, además de perceptivo.

Una teoría no puede refutarse antes de ser formulada; del mismo modo, no se puede encontrar la ausencia de X a menos que primero pensemos en X.

4. La negación es una intención

Ahora bien, no existe ninguna experiencia fenomenológica específica detrás de la palabra «no». La negación no tiene color ni forma, ni sonido, ni olor, ni sabor, ni sensación, ¡ya sea real o ilusoria! ¿Qué es entonces? Sugiero lo siguiente:

La negación como tal se refiere a un "acto mental" – o, dicho más precisamente, es un acto de voluntad (o, aún más precisamente, de veleidad) por parte de un Sujeto de conciencia. Específicamente, la negación es *una intención* . Nótese que nuestra voluntad de negar es en sí misma *un acto positivo* , aunque nuestra intención sea negar otra cosa.

La negación expresa una experiencia: el "fracaso" de encontrar algo que se ha buscado. Se persigue voluntariamente algún resultado cognitivo (la percepción de algún fenómeno positivo), pero este sigue sin alcanzarse (esta experiencia es cualitativamente una especie de sufrimiento, pero sigue siendo una intención positiva, cabe destacar). Por ello, mentalmente (o más precisamente, intencionalmente) marcamos la cosa como "ausente", es decir, construimos una idea de "negación" de lo buscado.

Así pues, la negación *no es un fenómeno* (una percepción física o mental), *sino algo intuido* (un acontecimiento de la voluntad del sujeto cognoscente). Cabe destacar que «intuición» se refiere aquí al autoconocimiento del sujeto de conciencia y agente de la voluntad. Se trata de una experiencia de naturaleza *no fenoménica* . Dicha autoexperiencia es inmediata: no hay distancia que salvar en el espacio ni en el tiempo.

Cuando un sujeto niega la presencia de un fenómeno material o mental, tras haberlo buscado en la experiencia sin encontrarlo, la «negación» consiste en un acto intencional específico. Esta intención es lo que llamamos «negación» o «rechazo de una hipótesis». Ocurre en el sujeto, aunque se refiere al objeto.

Esta intención, sin embargo, no es un acto arbitrario. Si lo fuera, sería puramente subjetivo. Este acto (en el mejor de los casos) sigue dependiendo lo suficiente de la percepción como para ser considerado «objetivo». El

sujeto aún debe observar y comprobar si X está presente; si esa experiencia positiva no se corresponde con su prueba empírica, concluye la ausencia de X.

En efecto, una negación inicial puede resultar errónea tras un análisis más detenido; es decir, a veces pensamos que algo «no existe» y, tras una investigación más exhaustiva, descubrimos que, por el contrario, sí existe. Por lo tanto, esta teoría de la negación no debe interpretarse como una afirmación de que negar algo lo convierte en realidad. La negación se rige por los principios de la aducción: se basa en la apariencia, que es creíble mientras se confirme, pero que posteriormente puede ser refutada.

Podemos hablar a posteriori de una ausencia objetiva, pero no podemos definir completamente la «ausencia» más allá de su «no-presencia», y el prefijo «no-» aquí no se refiere a un fenómeno, sino a una intención. La «ausencia» se experimenta, en efecto, pero resulta *imperceptible* sin que el Sujeto se pregunte previamente «¿está X presente?».

La ausencia, entonces, no es producida por el Sujeto, sino que se hace perceptible por su vana búsqueda de presencia. Porque, reiteramos, la no-X no se experimenta como un contenido específico de la conciencia, sino como un fracaso continuo para experimentar los fenómenos positivos particulares que definen X para nosotros.

Aunque solo somos conscientes directamente de las realidades aparentes, podemos inferir inductivamente las realidades no aparentes a partir de la experiencia de que las apariencias son cambiantes y a veces inexistentes. En base a esto, consideramos las categorías de «existencia» y «apariencia» como desiguales, y la primera como más amplia que la segunda. De manera similar, inferimos inductivamente la «ausencia objetiva» a partir de la «búsqueda no encontrada», aun cuando no tengamos acceso directo a la primera, sino solo acceso indirecto por

extrapolación a partir de la segunda. Dicha inferencia es válida, con un grado de probabilidad proporcional a la diligencia debida que ejerzamos.

Por estas razones, considero que el acto de negación es una clave importante para comprender la naturaleza y el estatus de la lógica. La negación es tan fundamental para la razón, un hecho epistémico tan crucial, que no puede reducirse a otra cosa.

Podemos describirlo, *a grandes rasgos* , como la intención de «tachar» (bajo la influencia de alguna razón) el elemento propuesto de nuestra lista mental de entes. Sin embargo, esto seguramente parecerá una definición circular o una repetición de sinónimos. Es evidente que *no podemos hablar de negación sin practicarla* . Por lo tanto, conviene que admitamos el acto de negación como un concepto fundamental para la lógica.

Cabe señalar, de paso, que la presente teoría de la negación proporciona a la biología una distinción interesante con respecto a los animales racionales.

Los seres sensibles que carecen de esta facultad de negación solo pueden responder al presente, mientras que una vez que esta facultad aparece en un organismo (como sucedió en la especie humana), puede trascender mentalmente el aquí y el ahora. Un animal meramente sensorial simplemente reacciona a los acontecimientos actuales, mientras que un ser humano puede temer los peligros y prepararse para ellos.

Una vez que aparece la facultad de negación, la mente puede comenzar a *abstraer, concebir alternativas y formular hipótesis* . Se requiere memoria e imaginación para proyectar una idea positiva, pero también la intención de negar para rechazar proyecciones inadecuadas. Sin esta capacidad crítica, nuestras fantasías nos conducirían rápidamente a situaciones destructivas.

4. Aclarando la paradoja

De Future Logic 31.

Un campo fundamental de la lógica es el de la paradoja, pues nos proporciona una poderosa herramienta para establecer algunas de las certezas más fundamentales de esta ciencia. Nos permite reivindicar para la epistemología y la ontología el estatus de ciencias verdaderas, en lugar de meras digresiones especulativas. Esta elegante doctrina puede considerarse parte del estudio de los axiomas.

1. inconsistencia interna

Consideremos la forma hipotética "Si P, entonces Q", que es una parte esencial del lenguaje de la lógica. Se definió como "P y no Q es una conjunción imposible".

Es axiomático que la conjunción de cualquier proposición P y su negación noP es imposible; por lo tanto, una proposición P y su negación noP no pueden ser ambas verdaderas. Una consecuencia obvia de esto, que se obtiene al considerar noP como la proposición en cuestión en lugar de P, es que la conjunción de cualquier proposición noP y su negación no-noP es imposible; por lo tanto, una proposición P y su negación noP no pueden ser ambas falsas.

Así pues, la Ley de Identidad podría formularse como: «Para cualquier proposición, "Si P, entonces P" es verdadera, y "Si no P, entonces no P" es verdadera». Las Leyes de Contradicción y del Tercero Excluido podrían enunciarse como: «Para cualquier proposición, "Si P,

entonces no-no P" es verdadera (P y no P son incompatibles), y "Si no-no P, entonces P" es verdadera (no P y P son exhaustivos)».

Ahora bien, consideremos las proposiciones paradójicas «Si P, entonces no P» o «Si no P, entonces P». A primera vista, tales proposiciones parecen ser antinomias obviamente imposibles y necesariamente falsas.

Pero analicemos sus significados con más detenimiento. La primera afirma «P y (no no)P es imposible», lo que simplemente significa «P es imposible». La segunda afirma «noP y no P es imposible», lo que simplemente significa «noP es imposible». Presentadas en este formato definitorio, estas afirmaciones ya no parecen antinómicas. Simplemente nos informan de que la proposición P, o noP, según sea el caso, contiene una falla intrínseca, una contradicción interna, una propiedad de autonegación.

De esto se desprende que puede haber proposiciones lógicamente autodestructivas, que a su vez sustentan lógicamente sus propias negaciones. Propongamos entonces las siguientes definiciones. Una proposición es ***autocontradictoria*** si se niega a sí misma, es decir, implica su propia negación. Por consiguiente, una proposición es ***autoevidente*** si su negación es autocontradictoria, es decir, si se deduce de su propia negación.

Así, la proposición «Si P, entonces noP» nos indica que P es contradictoria en sí misma (y, por lo tanto, lógicamente imposible), y que noP es evidente por sí misma (y, por lo tanto, lógicamente necesaria). De igual modo, la proposición «Si noP, entonces P» nos indica que noP es contradictoria en sí misma, y que P es evidente por sí misma.

La existencia de paradojas no indica en absoluto una falla formal. La *paradoja* , la proposición hipotética en sí misma, no es antinómica. Puede ser verdadera o falsa,

como cualquier otra proposición. Si se le concede su verdad, es su tesis antecedente la que es antinómica y falsa, ya que se niega a sí misma; la tesis consecuente es entonces verdadera.

Si la proposición paradójica «Si P, entonces no P» es verdadera, entonces su contradictoria «Si P, no entonces no P», que significa «P no es imposible», es falsa; y si esta última es verdadera, la primera es falsa. Del mismo modo, «Si no P, entonces P» puede ser contradicha por «Si no P, no entonces P», que significa «no P no es imposible».

Las dos paradojas «Si P, entonces noP» y «Si noP, entonces P» son contradictorias entre sí, ya que implican la necesidad de proposiciones incompatibles, noP y P respectivamente. Por lo tanto, aunque dichas proposiciones tomadas individualmente no son antinómicas, la doble paradoja, una situación en la que ambas proposiciones paradójicas son verdaderas a la vez, es inaceptable para la lógica.

A diferencia de las proposiciones hipotéticas positivas, las negativas carecen de la capacidad de expresar paradojas. Las proposiciones «Si P, entonces no P» y «Si no P, entonces no P» *no tienen* sentido, ni son lógicamente concebibles, ni son verdaderas. Cabe destacar que tales proposiciones son formalmente falsas. Dado que una forma como «Si P, entonces no Q» se define con referencia a una conjunción positiva como «{P y no Q} es posible», no podemos sustituir P por Q aquí sin incurrir en antinomia (para decir «{P y no P} es posible»), ni no P por P y Q (para decir «{no P y no-no P} es posible»).

De ello se deduce que la proposición «si P, entonces no P» no implica la forma en minúscula «si P, no-entonces P», y la proposición «si no P, entonces P» no implica la forma en minúscula «si no P, no-entonces no P». Es decir, en el contexto de la paradoja, las proposiciones hipotéticas se

comportan de forma anómala, y no como formas basadas en contingencias.

Esto no debería sorprendernos, puesto que lo contradictorio es lógicamente imposible y lo evidente es lógicamente necesario. Dado que las proposiciones paradójicas implican tesis y antítesis incontinentes, están sujetas a las leyes propias de dicha base.

Las implicaciones y la coherencia de todo esto se analizarán próximamente.

2. La falacia del concepto robado

Las proposiciones paradójicas ocurren en la práctica; además, nos brindan resultados muy significativos. He aquí algunos ejemplos:

- La negación, o incluso la duda, de las leyes de la lógica oculta una apelación a esos mismos axiomas, lo que implica que se debe creer en la negación en lugar de en la afirmación;
- La negación de la capacidad del hombre para conocer objetivamente cualquier realidad constituye, en sí misma, una pretensión de conocimiento de un hecho de la realidad;
- La negación de la validez de la percepción humana, o de su capacidad conceptual, o de su razonamiento, todas esas afirmaciones escépticas presuponen la utilización y la confianza en las mismas facultades puestas en duda;
- La negación, por principio, de toda generalización, necesidad o absoluto, es en sí misma una pretensión de una verdad general, necesaria y absoluta.
- La negación de la existencia de los "universales" no elude el problema de los universales, puesto que apela a algo en sí mismo, a saber, que los "universales" "no" y "existen".

Encontrará más detalles sobre estas y otras paradojas a lo largo del texto. Así pues, la detección de paradojas es una técnica lógica importante y de uso frecuente. La escritora Ayn Rand hizo gran hincapié en este método para refutar las filosofías escépticas, demostrando que *estas apelan implícitamente a conceptos que pretenden negar explícitamente* ; a esto lo denominó «la falacia del concepto robado».

Una forma de comprender el funcionamiento de una paradoja es considerarla en el contexto de un dilema. Una proposición evidente P podría enunciarse como «Tanto si se afirma como si se niega, es verdadera»; una verdad absoluta es algo que resulta ser cierto independientemente de nuestras suposiciones iniciales.

Esto puede escribirse como un argumento constructivo cuyo extremo izquierdo es la proposición axiomática de la identidad de P consigo mismo, y cuyo extremo derecho es la paradoja de la autocontradicción de nonP; la premisa menor es el axioma de la contradicción total entre los antecedentes P y nonP; y la conclusión, la verdad absoluta del consecuente P.

Si P, entonces P — y — si no P, entonces P

pero ya sea P o no P

Por lo tanto, P.

Una versión destructiva puede formularse igualmente bien, utilizando la forma contrapuesta de identidad, 'Si noP, entonces noP', como cuerno izquierdo, con el mismo resultado.

Si noP, entonces noP — y — si noP, entonces P

pero o no-noP o noP

por lo tanto, no-noP, es decir, P.

La conclusión 'P' aquí significa que P es lógicamente necesaria, no simplemente que P es verdadera, nótese bien; esto se deduce de la necesidad formal de la premisa

menor, la disyunción de P y noP, suponiendo que el cuerno derecho esté bien establecido.

Otra forma de entender la paradoja es considerarla en términos de contextos de conocimiento. Si interpretamos la paradoja "si noP, entonces P" como "todos los contextos con noP son contextos con P", y la identidad "si P, entonces P" como "todos los contextos con P son contextos con P", podemos inferir que "todos los contextos son con P", lo que significa que P es lógicamente necesario.

De forma similar, podemos abordar la paradoja «si P, entonces no P» para llegar a la conclusión «no P», o mejor aún: P es imposible. El proceso de resolver una paradoja, extrayendo sus conclusiones categóricas implícitas, puede denominarse *dialéctica* .

Cabe señalar que la expresión abreviada del dilema simple, en una sola proposición, resulta ahora más comprensible. La proposición compuesta «Si P, entonces {Q y noQ}» simplemente significa «noP»; «Si noP, entonces {Q y noQ}» significa «P»; «Si (o si) P o noP, entonces Q» significa «Q»; y «Si (o si) P o noP, entonces noQ» significa «noQ». Estas proposiciones también podrían clasificarse como paradójicas, aunque la contradicción generada se refiera a otra tesis.

Sin embargo, recuerde que las dos formas anteriores no deben confundirse con las relaciones hipotéticas negativas menores 'Si P o no P, (si no, entonces no) Q' o 'Si P o no P, (si no, entonces no) no Q', respectivamente, que no son paradójicas, a menos que existan condiciones bajo las cuales se eleven al nivel de hipótesis positivas.

3. Sistematización

Normalmente, partimos de la base de que nuestra información ya está libre de tesis evidentes o

contradictorias, mientras que en situaciones anormales, como en el caso de una paradoja, las tesis necesarias o imposibles son eventualidades formalmente aceptables.

Una hipótesis de la forma primaria «Si P, entonces Q» se definió como «P y no Q son imposibles juntas». Pero existen varias maneras en que esta situación puede surgir. O bien (i) ambas tesis, P y no Q, son individualmente contingentes, y solo su conjunción es imposible; esta es la situación normal. O bien (ii) la conjunción es imposible porque una u otra de las tesis es individualmente imposible, mientras que la restante es individualmente posible, es decir, contingente o necesaria; o porque ambas son individualmente imposibles; estas situaciones generan una paradoja.

Asimismo, una hipótesis de la forma primaria contradictoria «Si P, no Q» se definió como «P y no Q son posibles juntas». Pero esta situación puede darse de varias maneras. O bien (i) ambas tesis, P y no Q, y también su conjunción, son contingentes; esta es la situación normal. O bien (ii) una u otra es individualmente no solo posible sino necesaria, mientras que la otra es individualmente contingente, de modo que su conjunción sigue siendo contingente; o bien ambas son individualmente necesarias, de modo que su conjunción también es no solo posible sino necesaria; estas situaciones generan una paradoja.

Estas alternativas se aclaran en las siguientes tablas, tanto para las formas primarias como para sus derivados que incluyen una o ambas antítesis. El término «posible» significa, por supuesto, «contingente o necesario», y constituye el punto en común entre ambos. Aquí utilizaremos los símbolos « **N** » para necesario, « **C** » para contingente (que significa posible pero innecesario) y « **M** » para imposible. Las combinaciones están numeradas

para facilitar la consulta. Las simetrías presentes en estas tablas garantizan su exhaustividad.

Tabla 1Modalidades de tesis y conjunciones.

No.	Theses				Conjunctions			
	P	nonP	Q	nonQ	P and Q	P and nonQ	nonP and Q	nonP and nonQ
Normal (P,Q both contingent)								
1.	C	C	C	C	C	C	C	C
2.	C	C	C	C	M	C	C	C
3.	C	C	C	C	C	M	C	C
4.	C	C	C	C	C	C	M	C
5.	C	C	C	C	C	C	C	M
6.	C	C	C	C	C	M	M	C
7.	C	C	C	C	M	C	C	M
Abnormal (one or both of P, Q not contingent)								
8.	M	N	C	C	M	M	C	C
9.	N	M	C	C	C	C	M	M
10.	C	C	M	N	M	C	M	C
11.	C	C	N	M	C	M	C	M
12.	N	M	N	M	N	M	M	M
13.	N	M	M	N	M	N	M	M
14.	M	N	N	M	M	M	N	M
15.	M	N	M	N	M	M	M	N

La siguiente tabla se deriva de la anterior. ' **Sí** ' indica que una implicación y su contrapuesta están implícitas en la forma en cuestión, mientras que ' **no** ' indica que están excluidas de ella. ' □' aquí significa implica, y ' □' significa es implicado por.

Tabla 2Hipótesis definidas correspondientes.

No.	Name	Implications (→) and Contraposites (←)			
		P→nonQ	P→Q	nonP→nonQ	nonP→Q
		nonP←Q	nonP←nonQ	P←Q	P←nonQ
	Normal (P,Q both contingent)				
1.	Neutral	no	no	no	no
2.	Contrary	yes	no	no	no
3.	Subalternating	no	yes	no	no
4.	Subalternated	no	no	yes	no
5.	Subcontrary	no	no	no	yes
6.	Implicant	no	yes	yes	no
7.	Contradictory	yes	no	no	yes
	Abnormal (one or both of P, Q not contingent)				
8.	P impossible, Q contingent	yes	yes	no	no
9.	P necessary, Q contingent	no	no	yes	yes
10.	P contingent, Q impossible	yes	no	yes	no
11.	P contingent, Q necessary	no	yes	no	yes
12.	P, Q both necessary	no	yes	yes	yes
13.	P necessary, Q impossible	yes	no	yes	yes
14.	P impossible, Q necessary	yes	yes	no	yes
15.	P, Q both impossible	yes	yes	yes	no

La lógica hipotética normal presupone que las tesis de los hipotéticos son siempre contingentes, y por lo tanto se limita a los casos n. ° **1** a **7** de las tablas anteriores. Sin embargo, para desarrollar una lógica hipotética completa, también deben considerarse los casos anormales n. ° **8** a **15** , en los que una o ambas tesis no son contingentes (es decir, son evidentes o contradictorias).

La definición de la forma positiva primaria 'Si P, entonces Q', aunque permanece sin cambios como 'P más no Q no es posible', ahora se considera que comprende con mayor precisión las siguientes situaciones: números **3** , **6** , **8** , **11** , **12** , **14** o **15** , es decir, todos los casos en los que 'P y no

Q' es imposible ('M'), o 'P implica Q' está marcado como 'sí'.

La definición de la forma negativa primaria 'Si P, no-entonces Q', aunque permanece sin cambios como 'P más no-Q no es imposible', ahora se ve que comprende con mayor precisión las siguientes situaciones: Nos. **1** , **2** , **4** , **5** , **7** , **9** , **10** o **13** , es decir, todos los casos en los que 'P y no-Q' es contingente (C), o 'P implica Q' está marcado 'no'.

Las otras seis formas hipotéticas, que implican las antítesis de P y/o Q, también pueden recibir definiciones mejoradas, haciendo referencia a las tablas anteriores.

Observe las simetrías en estas tablas. En el caso n.° **1** , todas las conjunciones son «C» y todas las implicaciones son «no». En los casos n. **° 2** a **5** , una conjunción es «M» y una implicación es «sí». En los casos n.° **6** a **11** , dos conjunciones son «M» y dos implicaciones son «sí». En los casos n. **° 12** a **15** , tres conjunciones son «M» y tres implicaciones son «sí». Observe el estado correspondiente de cada tesis en cada caso.

El proceso de contraposición es universalmente aplicable a todas las hipótesis, positivas o negativas, normales o anormales, ya que procede directamente de las definiciones. Por esta razón, en las tablas anteriores, cada implicación está firmemente vinculada a su contrapuesta. Asimismo, la negación de cualquier implicación genera la negación de su contrapuesta, de modo que las tablas anteriores también se refieren indirectamente a hipótesis negativas.

Al desarrollar nuestra teoría de proposiciones hipotéticas, debemos tener cuidado de formular claramente el alcance y los límites de aplicación de cualquier proceso que consideremos, y especificar las excepciones, si las hubiere, a sus reglas. La validez o invalidez de los procesos lógicos a menudo depende de si nos centramos

en formas normales o anormales, aunque en algunos casos estas dos clases de proposiciones se comportan de la misma manera. Si no se tienen en cuenta estas distinciones, podemos incurrir fácilmente en inconsistencias formales.

4. Propiedades

Las proposiciones paradójicas obedecen las leyes de la lógica que resultan aplicables a todas las hipótesis, es decir, a las hipótesis de base no especificada. Sin embargo, las paradójicas, al ser hipótesis basadas en incontinencias, poseen propiedades que las hipótesis normales no tienen, o bien carecen de propiedades que sí poseen. En tales situaciones, donde se presentan diferencias en las propiedades lógicas, la lógica hipotética general sigue el caso más débil.

Las similitudes y diferencias en el comportamiento formal ya se han tratado con el detalle adecuado en los capítulos correspondientes, pero algunas se revisan aquí para subrayar el papel que desempeña la paradoja.

a. Oposición.

En la doctrina de la oposición, afirmábamos que «Si P, entonces Q» y «Si P, entonces no Q» debían ser contrarias, porque si P fuera verdadera, tanto Q como no Q serían verdaderas, lo cual es absurdo. Sin embargo, si hubiéramos planteado estas proposiciones en un dilema destructivo, como el que se muestra a continuación, habríamos obtenido un argumento legítimo:

Si P, entonces Q — y — si P, entonces no Q
pero ya sea noQ o Q
por lo tanto noP

De igual modo, 'Si P, entonces Q' y 'Si noP, entonces Q' podrían encajar en un dilema constructivo simple válido, dando como resultado Q, en lugar de argumentar como lo hicimos que deben ser contradictorios porque sus contraposiciones dan como resultado el absurdo de que noQ implique noP y P.

De ello se deduce que estas contrariedades solo son válidas condicionalmente, para hipótesis basadas en contingencias. Existen circunstancias excepcionales en las que no se cumplen, concretamente en relación con hipótesis anormales (incluidas las paradójicas).

Esto también queda claro de forma independiente a partir de la observación de marcas de "sí" que aparecen en paralelo, en los casos n. **° 8** , **14** , **15** (que permiten tanto "P implica noQ" como "P implica Q", donde P es imposible), y en los casos n. **° 11** , **12** , **14** (que permiten tanto "P implica Q" como "noP implica Q", donde Q es necesario).

Se aplican restricciones similares automáticamente a la subcontrariedad entre «Si P, entonces no Q» y «Si P, entonces no Q», y de igual modo a la subalternancia por la mayúscula «Si P, entonces Q» de la minúscula «Si P, entonces no Q» (que corresponde a la obversión). Estas oposiciones solo son válidas para proposiciones hipotéticas normales; al tratar con proposiciones hipotéticas anormales (y, por lo tanto, en lógica general), debemos, por coherencia, considerar dichas proposiciones como neutrales entre sí.

b. Educación.

De igual modo ocurre con las deducciones derivadas. El proceso primario de contraposición es incondicional, aplicable a todas las hipótesis, pero los demás procesos pueden ser criticados del mismo modo que antes, mediante la formulación de dilemas simples válidos,

utilizando la proposición original y la negación de la propuesta objetivo, o la(s) contrapuesta(s) de una u otra, o de ambas, como cuernos.

Alternativamente, estas proposiciones pueden combinarse en un silogismo, lo que da como resultado una conclusión paradójica. Por lo tanto:

En el caso de obversión o conversión obvertida (en el primero, negar el contrapuesto del objetivo):

Si Q, entonces noP (negación del objetivo)

Si P, entonces Q (fuente)

Entonces, si P, entonces noP (paradoja = noP)

En el caso de conversión por negación o inversión obvertida (en este último caso, negar el contrapuesto del objetivo):

Si P, entonces Q (fuente)

Si no es Q, entonces P (negación del objetivo)

Entonces, si no Q, entonces Q (paradoja = Q)

Por lo tanto, los procesos eductivos distintos de la contraposición solo son útiles para hipótesis basadas en contingencias y no pueden imitarse en la lógica anormal de las paradojas. Esto se aclara en las tablas anteriores, como se muestra a continuación.

Consideremos la forma paradigmática 'Si P, entonces Q'. Si limitamos nuestra atención a los casos 1-7, entonces aparece solo en dos situaciones: subalternación (**3**) o implicación (**6**). En estas dos situaciones, 'P implica no Q' es uniformemente 'no', por lo que la inversa, 'Si P, no-entonces no Q', es verdadera; y la contrapuesta 'Q implica no P' también es 'no', por lo que la inversa obvertida, 'Si Q, no-entonces no P', es verdadera; 'no P implica Q' es uniformemente 'no', por lo que la inversa obvertida 'Si no P, no-entonces Q' es verdadera; y la contrapuesta 'no Q implica P' también es 'no', por lo que la inversa por negación 'Si no Q, no-entonces P' es

verdadera. Con respecto a la inversión y la conversión, no son aplicables, porque 'no P implica no Q' y 'Q implica P' son 'no' en un caso, pero 'sí' en el otro.

Sin embargo, si ahora ampliamos nuestra atención para incluir los casos n. **° 8** a **15** , vemos que «Si P, entonces Q» ocurre adicionalmente si P es contradictorio y Q es contingente (**8**), o si P es contingente y Q es evidente (**11**), o si P y Q son evidentes (**12**), o si P es contradictorio y Q es evidente (**14**), o si P y Q son contradictorios (**15**). Las uniformidades antes mencionadas, que hicieron factibles las deducciones indicadas, ya no se cumplen. Existe una mezcla de «no» y «sí» en las alternativas disponibles que inhibe dichas deducciones.

c. Deducción.

En lo que respecta al silogismo, los *modos no subalternos* , validados mediante reducción al absurdo, conservan su validez universal, puesto que dicha reducción indirecta es esencialmente una contraposición y no se presupone ningún otro proceso eductivo. Sin embargo, los *modos subalternos,* presentes en las tres figuras, solo son válidos para hipotéticos normales. Dado que estos modos presuponen subalternancias para su validación, es decir, dependen de reducciones directas mediante obversión o inversión obvertida, no son válidos para hipotéticos anormales.

En lo que respecta a la apódosis, los modos con premisa menor modal nos proporcionan el punto de entrada a la lógica anormal. En cuanto al dilema, es el instrumento *por excelencia* para desenterrar paradojas en el curso del razonamiento cotidiano. Si planteamos *cualquier* dilema simple, constructivo (como se muestra a continuación) o destructivo (mutatis mutandis), en forma silogística, obtenemos una conclusión paradójica:

Si P, entonces R — y — si Q, entonces R

pero P y/o Q

Por lo tanto, R

Esto implica las sorites:

Si no es R, entonces no es P (contraposición del cuerno izquierdo)

si noP, entonces Q (menor)

Si Q, entonces R (cuerno derecho)

Por lo tanto, si no es R, entonces es R (conclusión paradójica = R).

Así pues, las proposiciones paradójicas son parte integral de la lógica hipotética general, no un apéndice extraño. Resaltan la continuidad esencial entre el silogismo y el dilema simple, siendo este último reducible al primero.

De ello se deduce incidentalmente que, puesto que (como se ha visto anteriormente) la apódosis puede considerarse un caso especial y límite del dilema simple, y el dilema simple un caso especial y límite del dilema complejo, todos los procesos inferenciales relacionados con las hipótesis están estrechamente relacionados.

La paradoja que genera un dilema simple, por supuesto, depende de la veracidad de sus premisas. No debemos inferir apresuradamente, a partir de la paradoja inherente a todo dilema simple, que todas las verdades son, en última instancia, evidentes por sí mismas, ni que todas las falsedades son, en última instancia, contradictorias. El conocimiento no es una empresa puramente racional, sino que depende en gran medida de hallazgos empíricos.

Como ya se ha señalado, un dilema simple arroja como conclusión una necesidad o imposibilidad categórica solo si todas sus premisas son indudablemente incontinentes. Si existieran condiciones tácitas o dudas sobre la incondicionalidad de las hipótesis (los cuernos) y/o la disyunción (la premisa menor), la conclusión se vería

proporcionalmente debilitada en cuanto a su modalidad lógica.

Así pues, en referencia al ejemplo anterior, *aceptando* la premisa mayor: en el caso específico en que nuestra premisa menor sea una disyunción formalmente dada —si, por ejemplo, P y Q son contradictorias entre sí (P = noQ, Q = noP)— entonces la conclusión R es necesaria. Pero, por lo general, las alternativas P y Q enumeradas son solo exhaustivas en su contexto, de modo que la conclusión R es solo objetivamente cierta.

Así pues, aunque toda necesidad lógica es evidente por sí misma y toda imposibilidad lógica es contradictoria en sí misma, formalmente hablando, según nuestras definiciones, sería prudente afirmar que estas afirmaciones no son recíprocas en la práctica, y distinguir entre paradoja apodíctica y paradoja fáctica. La primera es evidente por sí misma; la segunda se deriva de datos más empíricos y, por lo tanto, aunque contextualmente fiable, tiene menos peso y carácter definitivo.

Por último, cabe señalar que la inconsistencia de dos dilemas simples "igualmente convincentes" ahora puede entenderse mejor, ya que implican paradojas contrarias.

5. La paradoja del mentiroso (temprana)

De Future Logic 32 y Ruminations 1 y 5.

1. Doble paradoja

Hemos visto que las proposiciones lógicas de la forma «si P, entonces noP» (que equivale a «noP») o «si noP, entonces P» (que equivale a «P») son perfectamente válidas. Significan que el antecedente es contradictorio y lógicamente imposible, y que el consecuente es evidente y lógicamente necesario. Como proposiciones en sí mismas, no son en absoluto antinómicas; es uno de sus componentes el que resulta absurdo.

Si bien cualquiera de esas proposiciones, por separado, es formalmente aceptable y susceptible de ser cierta, jamás pueden ser ambas verdaderas: son contrarias irreconciliables y su conjunción es formalmente imposible. Pues si alguna vez fueran ambas verdaderas, entonces tanto P como noP serían necesariamente verdaderas.

Por lo tanto, debemos distinguir entre *la paradoja simple* , que tiene (de forma más precisa de lo sugerido anteriormente) la forma 'si P, entonces noP; pero si noP, no-entonces P; de donde noP', o la forma 'si noP, entonces P; pero si P, no-entonces noP; de donde P', y *la paradoja doble* , que tiene la forma 'si P, entonces noP, *y* si noP, entonces P'.

Repito, una paradoja simple se encuentra dentro de los límites de la lógica, mientras que una paradoja doble los sobrepasa. La primera bien puede ser cierta; la segunda siempre denota un error de razonamiento. Sin embargo, cabe objetar que la paradoja doble se presenta con bastante frecuencia en la práctica. No obstante, esto no la convierte en correcta, del mismo modo que la ocurrencia de otros tipos de errores en la práctica no los convierte en verdaderos.

La doble paradoja es posible, como veremos, gracias a un *uso indebido y oculto de conceptos* . Se trata de una sofistería por excelencia, ya que nos genera la ilusión superficial de una afirmación significativa que produce resultados contrarios a la razón. Sin embargo, tras un análisis más profundo, podemos detectar que se trata de una falacia, como una ambigüedad o una equivocación, lo que significa que, de hecho, la aparente contradicción nunca se produjo.

La lógica exige que *una o ambas* de las proposiciones hipotéticas que constituyen la doble paradoja, o paradoja sobre paradoja, *sean falsas* . Mientras que la paradoja simple se *resuelve* concluyendo categóricamente el consecuente, sin negar la relación antecedente-consecuente, la doble paradoja se *disuelve* demostrando que una o ambas de las paradojas simples involucradas son falsas, inexistentes. Nótese la diferencia en la solución del problema: la resolución «explica» la paradoja simple, mientras que la disolución «elimina» la doble paradoja.

La doble paradoja *sirve para demostrar* que estamos cometiendo algún tipo de error; el hecho de haber llegado a una contradicción es prueba suficiente de que hemos hecho una suposición errónea. Nuestra capacidad para intuir correctamente las conexiones lógicas no se pone en duda, ya que el juicio inicial fue precipitado, sin

reflexionar sobre los términos implicados. Una vez aclarados los conceptos, es la propia facultad racional la que emite el juicio en contra de su impresión previa de conexión.

Es importante comprender que toda doble paradoja (al igual que toda paradoja simple) *nos enseña algo* . Estos sucesos no deben considerarse amenazas a la razón que pongan en duda la lógica en su conjunto, sino simplemente lecciones. Son fuentes de información que nos revelan ciertas reglas lógicas de formación de conceptos que, de otro modo, habríamos pasado por alto. Nos muestran los límites de la corrección lingüística.

2. La paradoja del mentiroso

Un antiguo ejemplo de doble paradoja es la conocida «Paradoja del Mentiroso», descubierta por Eubúlides, un griego del siglo IV a. C. perteneciente a la escuela megarense. Dice así: «¿Acaso un hombre que afirma estar mintiendo dice la verdad?». Las implicaciones parecen ser que, si miente, dice la verdad, y si no miente, también dice la verdad.

Aquí, el error conceptual que subyace a la dificultad es que la proposición se *define por referencia a sí misma* . La paradoja del mentiroso es cómo descubrimos que tales conceptos no están permitidos.

La palabra «ahora» (que define la proposición misma como su sujeto) se usa en referencia a algo que aún no existe, cuya aparente existencia solo es posible gracias a ella. Así, de hecho, en este caso la palabra carece de referentes específicos. La palabra «ahora» suele tener significado, ya que en otras situaciones posee referentes precisos; pero en este caso se usa antes de que tengamos algo a lo que señalar como sujeto del discurso. Parece y suena como una palabra, pero no es más que eso.

Una versión más moderna y clara de esta paradoja es "esta proposición es falsa", porque resalta la función indicativa de la palabra "ahora" en la palabra "esto".

La palabra «esto» acompaña nuestras indicaciones y presupone que ya existe algo a lo que señalar. No puede crear un referente de la nada. Esta es la valiosa lección que nos enseña la paradoja del mentiroso. Podemos usar la palabra «esto» para señalar otra palabra «esto», pero no a sí misma. Así, puedo decirte «este "esto", que está en la proposición "esta proposición es falsa"», sin dificultad, porque mi «esto» tiene un referente, aunque sea un símbolo vacío; pero el «esto» original carece de significado.

Además, esta versión parece implicar que «si la proposición es verdadera, es falsa, y si es falsa, es verdadera». Sin embargo, al examinarla con más detenimiento, observamos que la expresión «la proposición» o «ella» tiene un significado diferente en antecedentes y consecuentes.

Si, a modo de ejemplo, entendemos esas implicaciones como: si esta proposición es falsa, entonces esta otra es verdadera; y si esta otra es verdadera, entonces esta otra es falsa —tomando el "esto" en el sentido de *autorreferencia* por cada tesis— entonces vemos que las tesis no tienen de hecho un mismo sujeto, y solo se presume que están en contradicción.

Formalmente no lo son, del mismo modo que, para cualquier P1 y P2, «P1 es verdadera» y «P2 es falsa» no se contradicen. Las implicaciones no son lógicamente necesarias, y por lo tanto, las dos paradojas se disuelven. No hay autocontradicción, ni en «esta proposición es falsa» ni, por supuesto, en «esta proposición es verdadera»; simplemente carecen de sentido, porque los indicativos que utilizan no tienen referencia.

Alternativamente, intentemos interpretar estas implicaciones de la siguiente manera: si «esta proposición es falsa» es verdadera, entonces aquella proposición es falsa; y si aquella proposición es falsa, entonces aquella proposición es verdadera. Tomamos el primer «esto» como autorreferencia y los «aquellos» subsiguientes como referencias a la proposición original y no a las tesis posteriores en sí mismas. En otras palabras, queremos decir: si «esta proposición es falsa» es verdadera, entonces «esta proposición es falsa» es falsa, y si «esta proposición es falsa» es falsa, entonces «esta proposición es falsa» es verdadera.

Aquí, los temas de las tesis son los mismos, pero las implicaciones ya no parecen necesarias, como se evidencia al sustituir el símbolo P por «esta proposición es falsa». El matiz de la paradoja ha desaparecido: solo existía mientras «esta proposición es falsa» parecía implicar, o implicar, «esta proposición es verdadera»; en cuanto el tema se unifica, ambas paradojas se desmoronan.

No podemos eludir el problema formulando la paradoja del mentiroso como una generalización. La proposición «Siempre miento» se puede refutar fácilmente con «A veces mientes (como en el caso de «Siempre miento»), pero a veces dices la verdad»; esto solo da lugar a una doble paradoja en forma indicativa. Del mismo modo, la proposición «Todas las proposiciones son falsas» se puede refutar sin dificultad con «Algunas sí, otras no».

Sin embargo, cabe señalar que ambas proposiciones generales son contradictorias en sí mismas; generan paradojas. Por consiguiente, ambas son falsas: no se puede afirmar que «siempre se miente», ni que «no existen proposiciones verdaderas». Esta es una inferencia lógica ordinaria y perfectamente legítima, puesto que existen alternativas lógicas.

Respecto a esas alternativas, la proposición «Nunca miento» no es inconsistente en sí misma, salvo para quien intencionadamente afirma «Siempre miento». De igual modo, la proposición «Todas las proposiciones son verdaderas» no es inconsistente en sí misma, pero sí lo es con el conocimiento lógico de que algunas proposiciones son inconsistentes; por lo tanto, es falsa. Así pues, en este caso solo puede sostenerse la proposición contingente «Algunas proposiciones son verdaderas y otras falsas».

3. Más sobre la paradoja del mentiroso

Una vez que comprendemos que el significado de las palabras reside en su intención, tanto individual como colectivamente, la solución a la paradoja del mentiroso se vuelve evidente. La autorreferencia carece de sentido, porque una intención no puede pretenderse a sí misma, ya que aún no existe; una intención solo puede pretender algo que ya existe, por ejemplo, otra intención dirigida a una tercera cosa.

En vista de esto, la proposición «esta proposición es falsa» carece de sentido, al igual que la proposición «esta proposición es verdadera». Ambas pueden declararse libremente como igualmente verdaderas o falsas, o ni verdaderas ni falsas; en este caso, no hay diferencia, porque las palabras «esta proposición» no se refieren a nada en absoluto [11].

Si bien las palabras utilizadas en estas oraciones tienen significado por separado y su estructura gramatical es correcta, la falta de contenido colectivo de las palabras implica que su valor lógico colectivo es nulo. La autorreferencia es sintácticamente coherente, pero semánticamente incoherente. Es como un argumento

---

11 Véase *Lógica del futuro* , capítulo 32.2.

circular, que se queda en el aire y no conduce a ninguna parte específica.

Respecto a la proposición exclusiva « *Solo* esta proposición es verdadera», implica tanto « *Esta* proposición es verdadera» como « *Todas las demás* proposiciones son falsas», es decir, es equivalente a la proposición exceptiva « *Todas* las proposiciones, *excepto esta* , son falsas». Esta última es frecuentemente afirmada por algunos filósofos, por ejemplo, por aquellos que dicen «todo es ilusión (excepto este hecho)».

Mi argumento aquí es que tales afirmaciones no solo implican la falacia de la autorreferencia (es decir, "esta proposición"). Dichas afirmaciones también implican una referencia a "todos los demás", lo cual es criticable, porque:

- Afirmar conocer "todas las demás proposiciones" equivale a pretender ser *omnisciente* , una pretensión de saberlo todo, tanto lo que hay que saber como lo que jamás se sabrá. Generalmente, tales afirmaciones se hacen sin ofrecer una justificación creíble, a pesar de contradecir todos los hallazgos previos de la experiencia y la razón.
- Sin duda, *otros* filósofos consideran y admiten como verdaderas otras proposiciones. Suelen ser bastante locuaces, incluso verborrágicos; no se limitan a afirmar una sola cosa y se niegan a decir nada más .
- Y por supuesto, formalmente, si "esto" no tiene sentido (como se demostró anteriormente), entonces "todos los demás", que significa "cualquier cosa *que no sea esto* ", ¡tampoco tiene sentido!

Por cierto, la paradoja del mentiroso se atribuye a los antiguos griegos, ya sea a Eubúlides de Mileto (siglo IV a . C.) o al anterior Epiménides de Creta (siglo VI ª . C.). Desconozco si su solución era evidente para estos primeros lógicos, pero un texto anónimo (¿europeo?) del

siglo XIV ^{d} . C. explicaba, según se cuenta, que la afirmación del mentiroso no es ni verdadera ni falsa, sino simplemente carente de sentido. Por lo tanto, esta explicación es históricamente mucho anterior a la lógica moderna (Russell y otros, aunque estos lógicos posteriores sin duda aclararon la cuestión).[12]

4. La utilidad de las paradojas

Una *proposición paradójica (simple)* tiene la forma «si P, entonces no P» o «si no P, entonces P», donde P es cualquier proposición. Es importante comprender que ***tales proposiciones son lógicamente legítimas dentro del discurso: una paradoja (simple) no es una contradicción*** . Por otro lado, una *doble paradoja* , es decir, una afirmación de que tanto «si P, entonces no P» *como* «si no P, entonces P» son verdaderas en un caso dado de P, sí es una contradicción.

El principio de no contradicción establece que la conjunción «P y noP» es lógicamente imposible; es decir, dos proposiciones contradictorias no pueden ser verdaderas a la vez. Del mismo modo, el principio del tercero excluido establece que «noP y no-noP» es lógicamente inaceptable. La razón de estas leyes es que tales situaciones de antinomia nos colocan en un dilema cognitivo: nos quedamos sin salida a la dificultad lógica, sin solución al problema inherente.

Por otro lado, una paradoja simple no representa tal amenaza para el pensamiento racional. Nos deja una salida lógica: la negación del antecedente (por ser contradictorio) y la afirmación del consecuente (por ser evidente). La proposición «si P, entonces no P» implica lógicamente «no P», y la proposición «si no P, entonces P» implica lógicamente «P». Por lo tanto, salvo la doble

---

12 Véase *Lógica futura* , capítulo 63, secciones 3 y 6.

paradoja, *una proposición que implica su propia negación es necesariamente falsa, y una proposición que se deriva de su propia negación es necesariamente verdadera* .

Por cierto, la conjunción de estas dos proposiciones hipotéticas, es decir, la doble paradoja, infringe el principio de no contradicción, puesto que da como resultado la conclusión compuesta de que «P y noP son ambas verdaderas». La doble paradoja también infringe el principio del tercero excluido, puesto que implica igualmente que «P y noP son ambas falsas».

Estas diversas inferencias pueden probarse y dilucidarse de diversas maneras:

- Dado que una proposición hipotética como "si x, entonces y" significa "x y no y es imposible", se deduce que "si P, entonces no P" significa "P y no no P son imposibles" (es decir, P es imposible), y "si no P, entonces P" significa "no P y no P son imposibles" (es decir, no P es imposible). Tenga en cuenta esta explicación.

  Sabemos que la negación de P es lo mismo que noP, y la negación de noP es igual a P, gracias a las leyes de no contradicción y del tercero excluido. Además, por la ley de identidad, repetir el nombre de un objeto no lo duplica: sigue siendo el mismo; por lo tanto, la conjunción "P y P" es equivalente a "P" y la conjunción "noP y noP" es equivalente a "noP".

  Nótese que el significado de «si P, entonces no P» es «(P y no no P) es *imposible* ». Por lo tanto, aunque esto implica que «no P es verdadero», no se *deduce* que «si no P es verdadero, P implica no P». De forma similar, *mutatis mutandis* , para «si no P, entonces P». Aquí nos interesa la implicación estricta (necesidad lógica), no la llamada implicación material.

  La razón por la que esta postura estricta es necesaria es que, en la práctica, la verdad y la falsedad son

contextuales: gran parte de lo que hoy creemos verdadero podría resultar falso mañana, y viceversa. Por otro lado, la necesidad o imposibilidad lógica se refiere a una relación mucho más fuerte que, en principio, una vez establecida, no debería variar con los cambios en el contexto del conocimiento: se aplica a *todos* los contextos imaginables.

- Dado que una proposición hipotética como "si x, entonces y" puede reformularse como "si x, entonces (x e y)", se deduce que "si P, entonces no P" es igual a "si P, entonces (P y no P)", y "si no P, entonces P" es igual a "si no P, entonces (no P y P)". Desde esta perspectiva, una proposición autocontradictoria implica una contradicción; dado que la contradicción es lógicamente inadmisible, se deduce que dicha proposición debe ser falsa y su contradictoria debe ser verdadera. Esto puede expresarse mediante la apódosis, en la que las leyes del pensamiento proporcionan la premisa menor categórica, lo que nos permite, excepcionalmente, extraer una conclusión categórica de una premisa hipotética.

  Si P, entonces (P y no P)
  pero: no(P y noP)
  Por lo tanto, no P

  Si noP, entonces (noP y P)
  pero: no(noP y P)
  Por lo tanto, no es noP

- También podemos tratar estas inferencias a modo de dilema, combinando la proposición "si P, entonces no P" con "si no P, entonces no P" (esta última derivada de la ley de identidad); o, de igual modo, "si no P,

entonces P" con "si P, entonces P". Esto nos da, de forma constructiva:

Si P, entonces noP; y si noP, entonces noP.
pero: o P o noP
Por lo tanto, noP

Si no P, entonces P; y si P, entonces P.
pero: o noP o P
Por lo tanto, P

Las paradojas a veces tienen consecuencias remotas. Por ejemplo, supongamos que Q implica P, y P implica noP (lo cual, como vimos, puede reescribirse como P implica tanto P como noP). Al combinar estas proposiciones en un silogismo, obtenemos la conclusión «si Q, entonces P y noP». Esta última también es una proposición paradójica, cuya conclusión es «noQ», aunque la contradicción en el consecuente no afecta directamente al antecedente. De manera similar, la no exclusión del término medio puede aparecer en la forma «si Q, entonces ni P ni noP». Este tipo de proposiciones también se encuentran en la práctica.

Es interesante que estas formas, "Q implica (P y noP), por lo tanto Q es falso" y "Q implica (no P y no noP), por lo tanto Q es falso", sean los argumentos implícitos en nuestra aplicación de las leyes del pensamiento correspondientes.

Cuando nos topamos con una antinomia en el conocimiento, buscamos dialécticamente eliminarla **encontrando y corrigiendo errores *previos* de observación o razonamiento** . Así, el argumento paradójico no solo deriva de las leyes del pensamiento,

sino que, en un sentido más amplio, es la forma en que las aplicamos habitualmente en la práctica.

Es decir, el proceso dialéctico que utilizamos tras el descubrimiento de una contradicción o un tercero excluido (o, en su caso, una violación de la ley de identidad) significa que creemos que:

> **Toda aparente antinomia es, en realidad, una ilusión.**

Se trata de una ilusión *debida a una paradoja* ; es decir, implica que *alguna de las premisas* que llevaron a esta conclusión aparentemente contradictoria o excluyente son erróneas y requieren corrección. La antinomia nunca es categórica, sino hipotética; es un indicio de, y depende de, alguna suposición o supuesto previo erróneo. La aparente antinomia contribuye al conocimiento al revelar alguna deficiencia en su totalidad y al animarnos a revisar nuestro pensamiento anterior.

Contradicción y paradoja están estrechamente relacionadas, pero no son lo mismo. Paradoja (es decir, paradoja simple, no doble) no equivale a antinomia. Podemos considerarlas como dificultades cognitivas de distinto grado. Desde esta perspectiva, mientras que la antinomia categórica sería un callejón sin salida que bloquearía cualquier pensamiento posterior, la paradoja es un grado de contradicción más leve (más hipotético), susceptible de resolución.

Como se desprende de todo lo anterior (y de otras observaciones posteriores), la paradoja desempeña un papel crucial en la lógica. La lógica de las proposiciones paradójicas no se limita a casos excepcionales como la paradoja del mentiroso. Es una herramienta esencial en la búsqueda del conocimiento, que nos ayuda a establecer

los fundamentos del pensamiento y, en general, a mantenerlo libre de impurezas lógicas.

La comprensión de las formas paradójicas no es un descubrimiento de la lógica moderna [13], aunque sí relativamente reciente (que data quizás del siglo XIV d. C., de la lógica escolástica).

[13] Por ejemplo, Charles Pierce (EE. UU., 1839-1914) observó que algunas proposiciones implican a todas las demás. Desconozco si se percató de que esta es una propiedad de las proposiciones contradictorias o lógicamente imposibles, y que las proposiciones evidentes o necesarias tienen la propiedad opuesta de estar implicadas por todas las demás. Sospecho que pensaba en términos de implicación material, más que de implicación estricta.

6. La paradoja del mentiroso (revisión)

De A Fortiori Logic, apéndice 7.4.

Ya abordé la paradoja del mentiroso en mi *libro Lógica del futuro* [14], pero ahora me doy cuenta de que es necesario profundizar en ella. Esta paradoja es particularmente difícil de abordar porque recurre simultáneamente a varios «trucos» discursivos.

1. Primer enfoque

La afirmación «Esta proposición es falsa» parece plausible a primera vista, hasta que nos damos cuenta de que si la asumimos verdadera, entonces debemos admitir que es falsa, mientras que si la asumimos falsa, entonces debemos admitir que es verdadera; todo lo cual parece inconcebible. Obviamente, existe una contradicción en tal discurso, ya que nada puede ser a la vez verdadero y falso. Pero la pregunta es: ¿qué la causa y cómo se puede resolver? No estamos «deduciendo» la contradicción a partir de una «ley del pensamiento», sino que la estamos «observando» mediante nuestra facultad racional. Tampoco podemos «deducir» la resolución de la contradicción a partir de una «ley del pensamiento»; debemos analizar el problema en cuestión con detenimiento y proponer de forma creativa una solución satisfactoria, es decir, una que realmente alivie nuestra

[14] Véase allí, capítulo 32.2. (Véase también *Reflexiones* 5.1.)

inquietud intelectual. Como veremos, esto no es en absoluto un asunto sencillo ni directo.

La proposición «Esta proposición es falsa» es una doble paradoja, porque: *si es verdadera, entonces es falsa; y si es falsa, entonces es verdadera* . Nótese la circularidad de verdadero a falso y de falso a verdadero. Las implicaciones que extraemos de la proposición dada parecen inevitables a primera vista. Pero debemos comenzar preguntándonos *cómo sabemos que estas implicaciones* (las dos proposiciones condicionales) son verdaderas. ¿Cómo sabemos que «es verdadero» implica «es falso», y que «es falso» implica «es verdadero»? Aparentemente, no estamos «deduciendo» estas implicaciones de alguna proposición implícita. Más bien, estamos utilizando una intuición racional ad hoc de algún tipo, es decir, en cierto sentido, «percibimos» directamente (comprendemos intelectualmente) las implicaciones de la proposición dada. Pero dicha intuición racional, aunque en principio fiable, es claramente *inductiva* , en lugar de deductiva, en su estatus epistemológico. Es decir, es digna de confianza hasta que, por alguna razón, se demuestre que es incorrecta. Esto significa que puede haber uno o más errores en nuestro razonamiento; no es una verdad absoluta. Y, en efecto, debe haber algún error, puesto que ha dado lugar a una doble paradoja. Por lo tanto, debemos buscarlo.

Quizás el uso del pronombre «ello» sea problemático, pues es un término bastante vago. Por lo tanto, preguntémonos: *¿a qué se refiere exactamente el pronombre «ello»* en este contexto?

A primera vista, el “ello” en “si es verdadero, entonces es falso; y si es falso, entonces es verdadero” se refiere a *la afirmación completa* , “Esta proposición es falsa”. En ese caso, debemos reformular la doble paradoja de la siguiente manera: *si “esta proposición es falsa” es*

verdadero, entonces “esta proposición es falsa” es falso; y si “esta proposición es falsa” es falso, entonces “esta proposición es falsa” es verdadero . Aquí, el sujeto de las dos afirmaciones condicionales (si-entonces) se distingue más claramente como “esta proposición es falsa”, y por lo tanto permanece constante en todo momento. Pero esta aclaración revela cambios anormales de predicado, de “verdadero” a “falso” y de “falso” a “verdadero”, que no se pueden explicar fácilmente. Normalmente, diríamos: si “esta proposición es falsa” es verdadero, entonces “esta proposición es falsa” es *verdadero* ; y si “esta proposición es falsa” es falso, entonces “esta proposición es falsa” es *falso* . La razón por la que aquí invertimos los predicados es que consideramos que la proposición original, "esta proposición es falsa", indica dicha inversión.

Sin embargo, mientras que una proposición de la forma «'esta proposición es falsa' es verdadera» se interpreta fácilmente en la forma más simple «esta proposición es falsa», una proposición de la forma «'esta proposición es falsa' es falsa» no puede simplificarse de la misma manera. ¿Cómo expresaríamos la doble negación implicada? ¿Como «esta proposición es verdadera»? Claramente, el significado de esta última no es idéntico al de la primera, ya que el sujeto «esta proposición» se refiere a proposiciones diferentes en cada caso. Por lo tanto, la formulación completa de la paradoja del mentiroso, es decir, «si 'esta proposición es falsa' es verdadera, entonces 'esta proposición es falsa' es falsa; y si 'esta proposición es falsa' es falsa, entonces 'esta proposición es falsa' es verdadera», no permite reproducir la fórmula inicial expresada en términos del pronombre «ello».

2. Segundo enfoque

Probemos, pues, otra cosa. Si el pronombre «ello» se refiere al *término* «esta proposición», entonces la doble paradoja debería reformularse de la siguiente manera: *si «esta proposición» es verdadera, entonces «esta proposición» es falsa; y si «esta proposición» es falsa, entonces «esta proposición» es verdadera* . Pero al hacerlo, vemos que en cada una de estas dos proposiciones condicionales, aunque el sujeto («esta proposición») permanece constante, el predicado («verdadero» o «falso», según sea el caso) no es el mismo en el consecuente que en el antecedente. No hay explicación lógica para estas inversiones del predicado. Normalmente, la verdad de una proposición P no implica su falsedad ni viceversa.

Podríamos sentir la tentación de usar la proposición dada "Esta proposición es falsa" como premisa para justificar la inferencia de dichos antecedentes a dichos consecuentes. Podríamos intentar formular dos apódosis, como sigue:

> Si esta proposición es verdadera, entonces es falsa (hipótesis),
>
> y esta proposición es falsa (dada);
>
> Por lo tanto, esta proposición es verdadera (conclusión hipotética).

> Si esta proposición es falsa, entonces es verdadera (hipótesis),
>
> y esta proposición es falsa (dada);
>
> Por lo tanto, esta proposición es verdadera (conclusión hipotética).

Obviamente, en el primer caso tenemos una inferencia inválida, ya que intentamos negar el antecedente para

negar el consecuente, o afirmar el consecuente para afirmar el antecedente. En el segundo caso, la supuesta conclusión se deriva de las premisas; pero aún podemos preguntarnos de dónde proviene la premisa mayor (la proposición hipotética), por lo que seguimos sin entender nada. Así pues, este enfoque también es inútil, es decir, no demuestra nada.

Como alternativa, podríamos intentar formular los dos silogismos siguientes:

> Esta proposición es falsa (dada),
>
> y esta proposición es verdadera (suposición);
>
> Por lo tanto, esta proposición es falsa (conclusión hipotética).

> Esta proposición es falsa (dato).
>
> y esta proposición es falsa (suposición);
>
> Por lo tanto, esta proposición es verdadera (conclusión hipotética).

Es evidente que estos argumentos no son silogísticos propiamente dichos; sin embargo, pueden reformularse ligeramente para formar silogismos. Las dos primeras premisas darían como resultado la conclusión «existe una proposición que es verdadera y falsa» (3/RRI), lo cual es contradictorio (por lo que una de las premisas debe ser falsa); las dos últimas premisas, al ser una misma proposición, no darían como resultado otra conclusión silogística que «existe una proposición que es falsa y falsa» (3/RRI), lo cual es evidente (y trivial). Pero estas no son las conclusiones que buscamos, las cuales deben referirse a «esta proposición» y no simplemente a «alguna proposición».

Un enfoque más adecuado consiste en analizar los dos últimos argumentos de la siguiente manera. En el primer caso, las premisas «esta proposición es falsa» (dato) y «esta proposición es verdadera» (suposición) parecen implicar conjuntamente «esta proposición es a la vez verdadera y falsa»; y la conclusión paradójica posterior sugiere, en efecto, que «esta proposición es falsa», puesto que la contradicción es imposible. En el segundo caso, las premisas «esta proposición es falsa» (dato) y «esta proposición es falsa» (suposición) coinciden en que «esta proposición es falsa», y por lo tanto, esta es su conclusión lógica. Dado que ambos argumentos concluyen con «esta proposición es falsa», esta última debe ser la conclusión general.

Sin embargo, este último resultado no es tan concluyente como parece, ya que, tras un análisis más detenido, resulta

obvio que «esta proposición es falsa» y «esta proposición es verdadera» no se refieren al mismo sujeto, puesto que el predicado cambia. La primera «esta proposición» se refiere a la proposición «esta proposición es falsa» y la segunda a la proposición «esta proposición es verdadera». Por lo tanto, estas dos proposiciones tienen, de hecho, sujetos y predicados diferentes (falso y verdadero, respectivamente). Los sujetos parecen iguales superficialmente, porque se expresan verbalmente con las mismas palabras; pero su intención subyacente no es la misma, puesto que se refieren a proposiciones significativamente diferentes (proposiciones con predicados manifiestamente diferentes, incluso contradictorios). Esto significa que cuando el predicado cambia, el sujeto también cambia efectivamente. Cuando el predicado es «verdadero», el sujeto significa una cosa; y cuando el predicado es «falso», el sujeto significa otra. Aunque las palabras «esta proposición» son constantes, su intención subyacente varía. Es decir, el término "esta proposición" no tiene un significado uniforme en todo momento y, por lo tanto, no puede utilizarse como base para las inferencias propuestas anteriormente.

3. Tercer enfoque

Ahora veamos otro enfoque. Si examinamos con más detenimiento nuestro razonamiento inicial en términos del pronombre «it», podemos ver lo que realmente sucede. Dado que «esta proposición es falsa» es verdadera, podemos decir de forma más breve: «esta proposición es falsa». Asimismo, dado que «esta proposición es falsa» es falsa, podemos deducir por negación que «esta proposición *no es* falsa» es verdadera, lo que significa que «esta proposición es verdadera» es verdadera, o dicho de

forma más breve: «esta proposición es verdadera» [15]. De esta manera, *parece* que argumentamos, con respecto al sujeto «esta proposición es falsa», desde «it is true» a «it is false», y desde «it is false» a «it is true». Pero, de hecho, el uso del pronombre «it» o del término «esta proposición» como sujeto abreviado es un truco, ya que el sujeto subyacente cambia en el curso de la segunda transición (la que termina en «esta proposición es verdadera»). Cuando se utilizan abreviaturas de forma generalizada, parece que nos referimos a una misma proposición como verdadera y falsa a la vez. Pero dado que esto se basa en una ambigüedad implícita, las paradojas desaparecen.

Es interesante observar que, al analizar el razonamiento de forma más explícita, la proposición «esta proposición es verdadera» también se vuelve paradójica. Podemos argumentar: si «esta proposición es verdadera» es verdadera, entonces obviamente «esta proposición es verdadera». Y: si «esta proposición es verdadera» es falsa, entonces su contradictoria «esta proposición *no es* verdadera» debe ser verdadera, lo que significa que «esta proposición es falsa» es verdadera, es decir, más sucintamente: «esta proposición es falsa». Aquí, superficialmente, no parece haber paradoja, porque parece que argumentamos, con respecto al sujeto «esta proposición es verdadera», desde «es verdadera» a «es verdadera», y desde «es falsa» a «es falsa». Pero si observamos la conclusión final, a saber, «esta proposición es falsa», vemos que *corresponde a* la paradoja del

---

[15] Algunos lógicos han intentado resolver la paradoja del mentiroso negando que verdadero y falso sean términos contradictorios, es decir, que no verdadero = falso y no falso = verdadero. Tal afirmación es un completo disparate; el intento de ignorar las leyes de no contradicción del tercero excluido para resolver una paradoja es una contradicción en sí misma.

mentiroso. [16]Y aquí, de nuevo, la explicación de la doble paradoja es que el sujeto aparente «ello» o «esta proposición» cambia de significado al extraer las implicaciones.

Nótese que, en ambos razonamientos, el primer paso es una autoimplicación ordinaria, una mera tautología, mientras que el segundo es la autocontradicción operativa, la paradoja. Si la proposición dada (ya sea «esta proposición es falsa» o «esta proposición es verdadera») es verdadera, simplemente la repetimos tal cual (sin necesidad de añadir la predicación «es verdadera»). Pero si la proposición dada es falsa, no podemos omitir la predicación adicional (es decir, «es falsa») sin alterar la proposición original. Por lo tanto, podríamos decir que las dos proposiciones, «esta proposición es falsa» o «esta proposición es verdadera», no presentan ningún problema cuando se consideran verdaderas; y solo cuando se consideran hipotéticamente falsas surge el problema. Así pues, podríamos decir que la solución a la paradoja del mentiroso (y su análogo positivo) consiste simplemente en aceptar ambas afirmaciones como verdaderas, ¡y no imaginarlas falsas!

Además, si quisiéramos, podríamos argumentar que «esta proposición es falsa» y «esta proposición es verdadera» se diferencian en que la primera parece explícitamente ponerse en duda, mientras que la segunda no. En base a esto, podríamos rechazar inmediatamente la primera y aceptar en cierta medida la segunda, aun admitiendo que

---

[16] Que «esta proposición es verdadera» sea implícitamente (aunque solo potencialmente) tan paradójico como «esta proposición es falsa» es, hasta donde sé, un descubrimiento reciente. Nótese bien cómo *ambas* paradojas surgen a través de deducciones bastante comunes: es decir, si se afirma «P es Q», entonces P es Q; y si se niega «P es Q», entonces se afirma «P no es Q», entonces P no es Q (donde P representa «esta proposición» y Q «falsa» o «verdadera», según el caso).

esta última carece igualmente de información útil. Es decir, dado que la primera parece «más paradójica» que la segunda, esta última es preferible *en caso de extrema necesidad* . Pero, cabe señalar, esto ignora las dificultades igualmente insuperables que presenta. Es mejor rechazar categóricamente ambas formas como construcciones viciosas.

4. Cuarto enfoque

Para comprender la naturaleza ilusoria de la paradoja del mentiroso, es importante darse cuenta de que las dos formas, «esta proposición es falsa» y «esta proposición es verdadera», *no son* contradictorias entre sí; y que, de hecho, ¡ninguna de ellas *tiene* una contradictoria! Esta es *una anomalía lógica* , una falla fatal en el discurso de la paradoja del mentiroso; pues, en principio, toda proposición bien formada y con sentido requiere lógicamente una contradictoria. Si una forma proposicional carece de una forma contradictoria, no puede juzgarse como verdadera o falsa, ya que tal juicio depende de que exista una elección. Ni siquiera tenemos que limitar nuestras proposiciones a los predicados «verdadero» o «falso»: cualquier predicado X y su negación no-X mostrarían la misma propiedad dado el mismo sujeto. Es decir, «esta proposición es X» y «esta proposición no-X» *no son* contradictorias entre sí, y por lo tanto, *ambas* carecen igualmente de contradicción.

Podríamos, por supuesto, señalar que «esta proposición es X» puede negarse con « *esa* proposición (es decir, la anterior) no es X», o incluso introducir *un símbolo* para la proposición original en la nueva proposición. En tal caso, aunque los sujetos serían verbalmente diferentes, sus intenciones serían sin duda las mismas. Pero la forma «esa proposición no es X» se asemeja más a la forma «'esta proposición es X' no es X», en la que toda la

proposición original asume el rol de sujeto y su predicado el de predicado. Sin embargo, aunque estas dos formas son en cierto modo equivalentes en significado entre sí y a la proposición original, sus patrones de comportamiento lógico no son idénticos a los de la proposición original, como ya hemos visto. El hecho es que «esta proposición no es X» no es la contradictoria de «esta proposición es X».

Es evidente que cualquier proposición que involucre el sujeto especial "esta proposición" exhibe una propiedad muy inusual y puede descartarse solo por ese motivo. La razón por la que dicha proposición carece de una contradicción es que su sujeto se refiere a la proposición en la *que se encuentra* , y esa proposición es claramente diferente cuando el predicado es el término "falso" y cuando es el término "verdadero" (o, más generalmente, cualquier par de predicados "X" y "no-X"). Cuando el predicado cambia, *también lo hace el sujeto* ; por lo tanto, el sujeto no puede fijarse, es variable, no es constante como debería ser. El término "esta proposición" tiene una referencia diferente en cada caso, que depende del predicado; en consecuencia, cada sujeto solo puede asociarse con un predicado y nunca con el otro (es decir, su negación).

De esto se desprende que, al principio, cuando al considerar la afirmación «Esta proposición es falsa», pensábamos que, si la tomábamos literalmente, debíamos considerarla falsa, y por lo tanto preferíamos «Esta proposición no es falsa», es decir, «Esta proposición es verdadera», etc., no nos dábamos cuenta de que, debido a la ambigüedad inherente al término «esta proposición» o «ello», estábamos cambiando su significado constantemente. Este cambio de significado pasa desapercibido, porque el término utilizado, por su propia naturaleza, no es fijo. Los pronombres «este» y «ello» pueden aplicarse a cualquier cosa y a su opuesto sin que

dicho cambio de significado se señale verbalmente en ellos. No están permanentemente ligados a ningún objeto, sino que son meras designaciones contextuales. En la terminología técnica de la lingüística, se caracterizan como «deícticos» o «indexicales».

Así pues, parece que la paradoja del mentiroso surge, independientemente de cómo entendamos sus términos, como resultado de una ambigüedad en el sujeto. Aunque superficialmente parezca que nos referimos al mismo sujeto en el antecedente y el consecuente de nuestro razonamiento condicional, en realidad existe un cambio de significado implícito que, una vez que lo percibimos, desmiente la aparente contradicción inicial. Las implicaciones imposibles sugeridas quedan así en duda, se vuelven inverosímiles. Las contradicciones aparentemente producidas se disipan o disuelven, en virtud de nuestra incapacidad para hacerlas persistentes.

5. Quinto enfoque

Otra forma complementaria de abordar la paradoja del mentiroso es señalar *la dificultad lógica de la autorreferencia* . Este es un enfoque que muchos lógicos han adoptado, incluyéndome a mí en mi primera incursión en este tema en *Lógica del Futuro* . El argumento que se propone aquí es que el término «esta proposición» se refiere a un objeto (es decir, «Esta proposición es falsa» o «Esta proposición es verdadera») que incluye el término mismo. Un dedo no puede señalarse a sí mismo, y «esto» es el equivalente conceptual de un dedo. En efecto, la expresión «esto» carece de contenido cuando se dirige a sí misma o a una oración que la incluye. Está vacía, sin sustancia. Es como si no se dijera nada cuando recurrimos a tal autorreferencia.

Así, «Esta proposición es X» (donde X representa falso, verdadero o cualquier otra cosa) carece de sentido; y una

oración sin sentido no puede ser verdadera ni falsa. Dicha oración puede describirse razonablemente como ni verdadera ni falsa, sin contravenir el principio del tercero excluido, ya que ninguna de estas evaluaciones lógicas se aplica a oraciones sin sentido. «Esta proposición es falsa» parece tener sentido porque sus cuatro componentes (es decir, «esta», «proposición», «es» y «falsa») tienen significado por separado. Pero en esta combinación particular, donde uno de los elementos (a saber, «esta») no se refiere a nada preexistente, la oración resulta ser sin sentido.

Las aparentes contradicciones que produce la autorreferencia nos ayudan a comprender su falta de sentido. Y es a través de la comprensión intelectual de esta falta de sentido que explicamos y anulamos las aparentes contradicciones. Sobre esta base, podemos afirmar que, si bien la oración «Esta proposición es verdadera» no genera a primera vista ninguna paradoja (como suele pensarse: «si es verdadera, es verdadera; y si es falsa, es falsa»), no obstante, dado que implica autorreferencia tanto como «Esta proposición es falsa», carece igualmente de sentido y no puede caracterizarse como verdadera o falsa. De hecho, como he demostrado anteriormente, «Esta proposición es verdadera» también genera una doble paradoja.

Alguien podría objetar: ¿Qué pasa con las proposiciones: «esta afirmación es autorreferencial» y «esta afirmación no es autorreferencial»? Sin duda, podemos decir que son significativas y que la primera es verdadera, mientras que la segunda es falsa. La respuesta a esa objeción es que las dos proposiciones «esta proposición se refiere a sí misma» y «esta proposición no se refiere a sí misma» *no son* contradictorias entre sí, porque (al igual que en la paradoja del mentiroso) sus sujetos difieren radicalmente, cada uno refiriéndose a la proposición en la que se encuentra y no a la otra. Así, aunque la versión positiva

pueda parecer más coherente que la negativa, y por lo tanto preferible *en casos extremos* , en realidad ambas son fundamentalmente erróneas, porque (al igual que en la paradoja del mentiroso) ninguna de ellas *tiene* una contradicción, y sin la posibilidad lógica de negar un discurso es imposible juzgar si es correcto o incorrecto.[17]

6. Sexto enfoque

Poco después de las reflexiones anteriores, me topé con otro ejemplo interesante de autorreferencia paradójica: «¡Desobedéceme!». [18]Esto implica una doble atadura: *si lo obedezco, lo desobedezco, y si lo desobedezco, lo obedezco* . Para resolver esta paradoja, primero debemos

---

[17] Otra objeción (que me planteó un lector) se refiere a proposiciones como «esta afirmación tiene cinco palabras» y «esta afirmación tiene seis palabras»: aunque contienen el demostrativo «esta», la primera parece verdadera y la segunda falsa. En respuesta, podríamos señalar que, si bien las proposiciones «esta afirmación tiene cinco palabras» y «esta afirmación no tiene cinco palabras» parecen significar cosas opuestas, no pueden ser contradictorias, ya que ambas parecen verdaderas. Compárese también: «esta afirmación tiene cinco palabras» y «esta afirmación sí tiene cinco palabras»: la primera es verdadera, mientras que la segunda es falsa, aunque ambas *significan* esencialmente lo mismo. Claramente, el comportamiento de estas proposiciones dista mucho de ser normal, debido a su inusual dependencia de la redacción utilizada. En un primer nivel, comprendemos el mensaje de la proposición, contamos el número de palabras y luego comprobamos si este número coincide con el número dado: si coincide, la proposición se considera «verdadera», y si no, se considera «falsa». Pero al mismo tiempo, debemos estar atentos al cambio de referencia del demostrativo "este", lo cual complica las cosas como ya se explicó, y además, en este contexto particular, debemos tener cuidado con el impacto de la redacción. Los Kneales citan "Lo que estoy diciendo ahora es una oración en inglés" como ejemplo de "autorreferencia inofensiva" (p. 228).

[18] Encontré este ejemplo en *La philosophie au jour le jour* de Robert Maggiori ; el autor no dice si es una invención suya o de otra persona (p. 438).

formular la afirmación con mayor precisión, por ejemplo: «¡Debes desobedecer esta orden!». Podemos entonces desenredar el nudo al darnos cuenta de que la orden *tiene una forma externa imperativa, pero internamente carece de contenido* . No define una acción específica y concreta que deba hacerse o no hacerse. Si quisiéramos obedecerla o desobedecerla, ¡no sabríamos qué se supone que debemos hacer o no hacer! Por lo tanto, es una orden que no puede ser obedecida ni desobedecida. Reflexionar sobre este caso me llevó a lo que ahora creo que es la clave, que resuelve de manera convincente la paradoja del mentiroso, aun cuando las reflexiones anteriores sigan siendo relevantes.

Entonces comprendí que este es precisamente el problema de la paradoja del mentiroso. Afirma que «esta proposición es falsa», pero *no nos dice nada sobre el mundo que pueda juzgarse como verdadero o falso* . Una «proposición» es una afirmación que hace alguna aseveración sobre el mundo. Si la afirmación no hace tal aseveración, si no «propone» nada, no puede evaluarse lógicamente como verdadera o falsa. Si no se refiere a nada —ya sea físico, mental o espiritual, perceptivo, intuitivo o conceptual—, carece de significado. Una afirmación sin sentido no califica como «proposición». Los atributos de «verdadero» o «falso» no son predicados comunes, como «blanco» o «negro», que pueden adjuntarse a cualquier sujeto y luego juzgarse como verdaderos o falsos. Los atributos de «verdadero» o «falso» requieren que se haga una afirmación precisa antes de que puedan utilizarse.

La veracidad de esta explicación se evidencia al considerar las «formas proposicionales» utilizadas en la teoría lógica. Por ejemplo, «Todos los X son Y». Esta forma proposicional no puede considerarse verdadera ni falsa, pues carece manifiestamente de contenido. Solo cuando se le otorga un contenido específico, como «Todos

los hombres son mortales», podemos empezar a cuestionar su veracidad. Una forma proposicional es demasiado vaga para considerarse una proposición. No nos revela nada sobre el mundo, salvo la implicación de que existen (o incluso la posibilidad de que existan) proposiciones concretas con esta forma. Del mismo modo que no podemos desobedecer ni obedecer un imperativo sin contenido, tampoco podemos juzgar como verdadera o falsa una expresión puramente formal.

Lo mismo ocurre con la paradoja del mentiroso: al igual que una proposición formal, carece de contenido concreto y, por lo tanto, no puede considerarse verdadera ni falsa. Esta paradoja carece de contenido en parte debido a que posee un sujeto autorreferencial («esta proposición»). Sin embargo, *incluso si su sujeto no fuera autorreferencial* , seguiría teniendo un contenido insuficiente. Esto se debe a que su predicado «falso» (y, del mismo modo, su opuesto, «verdadero») no es un predicado ordinario, sino más bien un predicado formal. Solo puede utilizarse si ya se ha propuesto otro predicado más concreto para el sujeto en cuestión. Por ejemplo, «esta proposición es interesante» podría considerarse verdadera o falsa (si no fuera autorreferencial) porque ya tiene un predicado («interesante»). Así pues, el problema de la paradoja del mentiroso no reside únicamente en la autorreferencia que implica, sino también en la falta de un predicado más concreto que el predicado lógico «falso» (o «verdadero»).

Todo esto ilustra cómo las «leyes del pensamiento» no son axiomas en el sentido de premisas fundamentales en la búsqueda del conocimiento, de las cuales derivamos mecánicamente otras premisas. Más bien, la expresión «leyes del pensamiento» se refiere a ideas recurrentes que nos brindan cierta orientación intelectual, pero que por sí solas no pueden determinar el resultado. El individuo que busca el conocimiento, y en particular el lógico, se ve *impulsado* por lo obvio o lo absurdo de una situación a

buscar soluciones creativas a los problemas. Aun así, debe pensar en posibles soluciones y ponerlas a prueba.

7. La paradoja de Russell (temprana)

De Future Logic 32, 42 y 45, y Ruminations 5.

1. Automembresía

Respecto al tema de la autopertenencia, es necesario profundizar en el tema. Intuitivamente, al menos para mí, la idea de que algo pueda ser a la vez contenedor y contenido resulta difícil de aceptar.

Ahora bien, la auto-pertenencia significa que un nominal es miembro de un nominal idéntico. Por lo tanto, que todos los X sean X y, por consiguiente, miembros de "X", no constituye auto-pertenencia; esta es simplemente la definición de pertenencia a una clase de primer orden por parte de un no-clase.

Vimos que, empíricamente, al menos con ejemplos comunes, “X” (o la clase de X) nunca es en sí mismo un X, ni por lo tanto un miembro de “X”. Por ejemplo, “perros” no es un perro, ni por lo tanto un miembro de “perros”.

Sugerí que esto podría generalizarse en un postulado inductivo, si no se presentaban ejemplos que demostraran lo contrario. Mi propósito aquí es demostrar que todos los casos aparentes de autopertenencia son ilusorios, debidos únicamente a la imprecisión del lenguaje.

Que “X” sea una clase X, y por lo tanto miembro de “clases X”, no es una auto-pertenencia en sentido literal, sino que es simplemente la definición de pertenencia a

una clase de segundo orden por parte de una clase de primer orden. Por ejemplo, “perros” es una clase de perros, o un miembro de “clases de perros”, o miembro de la clase de clases de perros.

La inferencia formal, a partir de que todos los X son X, implica que todas las clases X son clases X, y por lo tanto pertenecen a las "clases X" (o a la clase de clases de X), y por ende, no nos proporciona un ejemplo de lo que entendemos estrictamente por pertenencia a sí misma; es simplemente una tautología. Por ejemplo, todas las clases de perros pertenecen a las "clases de perros".

Afirmar que una clase X puede ser X, y por lo tanto un miembro de “X”, es simplemente una afirmación más amplia que afirmar que “X” puede ser X, y no solo parece igual de absurdo y sin fundamento empírico, sino que en cualquier caso no constituiría formalmente la autopertenencia. Por ejemplo, afirmar que “los retrievers” son perros.

En cuanto a decir de cualquier X que *es* “X”, en lugar de un miembro de “X”; o decir que *es* alguna otra clase de X, y por lo tanto un miembro de “clases de X”, tales afirmaciones simplemente no parecen estar de acuerdo con las intenciones de las definiciones de clases y clases de clases, y en cualquier caso no son auto-pertenencia.

Surge entonces la pregunta: ¿pertenece la categoría «X-classes» a la categoría «X-classes»? La respuesta es no; ni siquiera en este caso existe autopertenencia. La impresión de que «X-classes» podría pertenecer a sí misma se debe a que se refiere a X, aunque de forma menos directa que «X». Por ejemplo, «dog-classes» se refiere a «retrievers», «terriers» e incluso «perros»; y, por lo tanto, aunque solo indirectamente, se refiere a los perros.

Sin embargo, en términos más formales, la expresión «clases X» no satisface la condición definitoria para

pertenecer a ellas, que sería que «todas las clases X son X». Del mismo modo, la afirmación de que «X» pertenece a las clases X se basa en la premisa de que «todas las X son X». Como se demostrará a continuación, esto implica que la impresión anterior no puede considerarse una generalidad formal, sino, en el mejor de los casos, una verdad contingente en algunos casos; por consiguiente, pierde toda su fuerza y credibilidad.

Si decimos que *para cualquier* X, todas las clases de X son X, implicamos que para todo X, “X” (que es una clase de X) es X; pero ya hemos presentado casos empíricos que demuestran lo contrario; por lo tanto, la conexión no puede ser general ni formal. Así pues, solo podemos afirmar que quizás *para algún* X, todas las clases de X son X; pero con respecto a esa eventualidad, no se han presentado ejemplos.

Puesto que no tenemos fundamentos sólidos (ejemplos específicos) para suponer que "X" o "las clases X" sean alguna vez miembros de sí mismas, y la sugerencia está plagada de dificultades; y solo encontramos ejemplos creíbles en los que no eran miembros de sí mismas, estamos justificados en presumir, por generalización, que: *ninguna clase de nada, o clase de clases de nada, es nunca miembro de sí misma.*

Solo se me ocurre una posible excepción a este postulado: las «cosas» (o «clases de cosas»). Pero sospecho que, en este caso, en lugar de afirmar que la clase es miembro de sí misma, deberíamos considerar que la definición de pertenencia es insuficiente. Es decir, aunque este género supremo *sea* una cosa, no es «miembro de» nada.

2. La paradoja de Russell

La paradoja de Russell es un ejemplo moderno de doble paradoja, descubierta por el lógico británico Bertrand Russell.

Preguntó si la clase de "todas las clases que no son miembros de sí mismas" es o no miembro de sí misma. Si "clases que no son miembros de sí mismas" no es miembro de "clases que no son miembros de sí mismas", entonces sí lo es; y si "clases que no son miembros de sí mismas" es miembro de "clases que no son miembros de sí mismas", entonces también es miembro de "clases que son miembros de sí mismas". Por lo tanto, nos encontramos ante una contradicción en ambos casos.

En cambio, la clase de "todas las clases que son miembros de sí mismas" no plantea una dificultad similar. Si "clases que son miembros de sí mismas" no es miembro de "clases que son miembros de sí mismas", entonces es miembro de "clases que no son miembros de sí mismas"; pero si "clases que son miembros de sí mismas" es miembro de "clases que son miembros de sí mismas", no se deduce ninguna antinomia. Por lo tanto, aquí tenemos una única paradoja unida a una posición consistente, y se puede llegar a una conclusión definitiva: "clases que son miembros de sí mismas" es miembro de sí misma.

Ahora bien, toda incongruencia que surja en el conocimiento debe considerarse una oportunidad para avanzar, un estímulo para la investigación y el descubrimiento de algún detalle previamente desconocido. Entonces, ¿cuál es la lección oculta de este enigma?

Como demostraré, la paradoja de Russell parte esencialmente de una ambigüedad; se asemeja más al sofisma de la paradoja del Barbero que a la paradoja del Mentiroso. Porque *la cuestión no radica en si la autopertenencia es posible o no* . Russell creía que

algunas clases, como las "clases", se incluyen a sí mismas; aunque discrepo de ello, mi desacuerdo no justifica la disolución de la paradoja de Russell. Pues no es el concepto de autopertenencia lo que produce una inconsistencia bidireccional, sino el de no autopertenencia; y todos coinciden en que al menos algunas clases (si no todas, como creo) no se incluyen a sí mismas: por ejemplo, "perros" no es un perro.

Lo que ha desconcertado a tantos lógicos con respecto a la paradoja de Russell ha sido la suposición de que podemos formar conceptos a voluntad con solo formular una definición verbal. Pero este punto de vista carece de justificación. Las palabras deben tener un significado demostrable; en la mayoría de los casos, lo tienen; pero en algunos casos, se aíslan o se combinan sin prestar atención a sus requisitos estructurales intrínsecos. No podemos, por ejemplo, usar la palabra «mayor» sin especificar «¿que qué?»; muchas palabras están unidas y no se pueden reordenar al azar. El hecho de que comúnmente, en el discurso cotidiano, usemos las palabras de forma imprecisa para evitar construcciones tediosas no otorga a los lógicos la misma libertad.

3. Impermutabilidad

La solución al problema es tan sencilla que resulta graciosa, aunque debo admitir que estuve bastante perplejo durante un tiempo. Es simplemente que: ***las proposiciones de la forma 'X (o "X") es (o no es) un miembro de "Y" (o "clases Y")' no se pueden permutar*** . *El proceso de permutación es aplicable a algunas formas, pero no a todas.*

a. En algunos casos, cuando se trata de relaciones relativamente simples, la relación puede adjuntarse al predicado original para formar un nuevo predicado, en una proposición del tipo «S es P», donde «es» tiene un

significado estrictamente clasificatorio. Así, «X no es Y» es permutable a «X no es Y», o «X es algo que no es Y»; «X tiene (o carece de) cualidad Y» es permutable a «X es una cosa que tiene (o carece de) cualidad Y»; «X hace (o no hace) Y» es permutable a «X es una cosa que hace (o no hace) Y». En tales casos, este artificio no produce ningún error.

Pero en otros casos, la permutación no es factible, porque falsea las propiedades lógicas de la relación en cuestión. Vimos ejemplos claros e indudables de esto en el estudio de las modalidades.

Por ejemplo, la forma «X puede ser Y» no es permutable a «X es algo capaz de ser Y», porque al hacerlo cambiamos el sujeto de la relación «puede ser» de «X» a «algo», y también transformamos un potencial «puede ser» en un real «es (capaz de ser)». Como resultado de tales manipulaciones verbales, surgen errores formales. Así, se cree que «X es Y, y todos los Y son capaces de ser Z» concluye «X es capaz de ser Z», cuando en realidad las premisas son perfectamente compatibles con la contradictoria «X no puede ser Z», ya que «X puede convertirse en Z» es una conclusión alternativa válida, como vimos anteriormente.

Asimismo, se puede demostrar que «X puede convertirse en Y» no es permutable a «X es algo que puede convertirse en Y», ya que entonces el silogismo «X es Y, todos los Y son cosas que pueden convertirse en Z, por lo tanto X es algo que puede convertirse en Z» parecería válido, mientras que su conclusión correcta es «X puede ser o convertirse en Z», como se vio anteriormente. Por lo tanto, la modalidad es un tipo de factor relacional que no es permutable. Aunque comúnmente decimos «X es capaz o incapaz de Y», ese «es» no tiene las mismas propiedades lógicas que el «es» en una proposición normal «S es P».

b. La paradoja de Russell nos revela la valiosa información de que la cópula 'es miembro (o no miembro) de' tampoco está abierta a permutaciones a 'es algo que es miembro (o no miembro) de'.

El pronombre «es» original es parte integral de la relación y no tiene el mismo significado que un «es» aislado. La relación «es o no es miembro de» es un todo indivisible; no se puede simplemente separar donde se desee. El hecho de que consista en una secuencia de palabras, en lugar de una sola, es una consecuencia del lenguaje; el hecho de que se puedan separar sus constituyentes verbales no significa que la relación objetiva en sí misma pueda separarse de la misma manera.

La permutación es un proceso que utilizamos, cuando es posible, para sortear las dificultades inherentes a una relación especial; sin embargo, en este caso, no podemos eludir las exigencias particulares de las relaciones de pertenencia mediante este artificio. La paradoja de Russell nos obliga a seguir los procesos inferenciales descritos anteriormente; nos indica que no existen otros procesos legítimos y prohíbe los atajos conceptuales.

La inmutabilidad de «es (o no es) miembro de» implica que no se puede formar una clase de «clases que son miembros de sí mismas» ni una clase de «clases que no son miembros de sí mismas». No se trata de términos, sino de relaciones. Por lo tanto, la paradoja de Russell se disuelve por completo al negar la legitimidad conceptual de sus términos. No existe forma de que podamos formar tales conceptos; implican una permutación ilícita. Las conexiones entre los términos son, por consiguiente, puramente verbales e ilusorias.

La definición de pertenencia es «si algo *es* X, entonces es miembro de “X”» o «si todos los X *son* Y, entonces “X” es miembro de las “clases Y”». La paradoja de Russell nos hace conscientes de que el «es» en la condición debe

ser un «es» normal y solitario; no puede ser un «es» aislado de una cadena de palabras como «es (o no es) miembro de». Si no se cumple esta condición antecedente, no se puede aplicar la regla consecuente. En nuestro caso, la condición *no se cumple* , por lo que la regla no se aplica.

c. He aquí, pues, cómo surge formalmente la paradoja de Russell, paso a paso. Indicaremos las permutaciones mediante corchetes como este: {}.

Sea "X" cualquier clase, de cualquier orden:

(i) Si "X" es un miembro de "X", entonces "X" es {un miembro de sí mismo}. Llamemos Y a la porción encerrada; entonces "X" es Y, lo que define la autopertenencia.

(ii) Si "X" no es miembro de "X", entonces "X" es {no es miembro de sí mismo}. Llamemos a la porción encerrada nonY; entonces "X" es nonY, define la no pertenencia a sí mismo.

A continuación, apliquemos las definiciones generales de pertenencia y no pertenencia a los conceptos de Y y no-Y que acabamos de formar:

(iii) todo lo que no es Y, es nonY, y por lo tanto es un miembro de "nonY".

(iv) todo lo que es Y, no es miembro de "noY", puesto que solo las cosas que no son Y, son miembros de "noY".

Ahora bien, la doble paradoja:

(v) si "nonY" no es miembro de "nonY":

— entonces, al poner "nonY" en lugar de "X" en (ii), "nonY" es {no es un miembro de sí mismo}, lo que significa que no es nonY;

— entonces, por (iii), "nonY" es un miembro de "nonY", lo cual contradice la premisa inicial.

(vi) si "nonY" es un miembro de "nonY":

— entonces, al poner “nonY” en lugar de “X” en (i), “nonY” es {un miembro de sí mismo}, lo que significa que es Y;

— entonces, por (iv), “nonY” no es un miembro de “nonY”, lo cual contradice la premisa inicial.

De todos los procesos utilizados para desarrollar estos argumentos, solo uno tiene una validez incierta (no establecida): la permutación de «es miembro de sí mismo» a «es {miembro de sí mismo}», o de «no es miembro de sí mismo» a «es {no miembro de sí mismo}». Dado que todos los demás procesos son válidos, la fuente de la antinomia debe ser dicha permutación. QED

d. La existencia de relaciones inmutables sugiere que no podemos considerar que todas las relaciones residan de alguna manera *dentro* de las cosas relacionadas, como un componente intrínseco de sus identidades. Nos vemos obligados a considerar algunas relaciones, como la modalidad o la pertenencia, como vínculos que se sitúan fuera de los términos, que no son partes reales de su ser.

Así, por ejemplo, que «este S puede ser P» no implica ontológicamente que exista alguna «marca» programada en la identidad real de este S que registre que «puede ser P». Por esta razón, no se puede presumir que la cláusula verbal {puede ser P} sea una unidad; no hay nada que le corresponda en la realidad de este S, la relación potencial no proyecta una sombra real.

Por lo tanto, debe existir una realidad para la «existencia potencial», al margen de la «existencia actual». Cuando decimos que «este S puede ser P», consideramos que esta potencialidad es P como parte intrínseca de la «naturaleza» de este S. Pero el S al que nos referimos se extiende en el tiempo, en el pasado, el «presente» y el futuro; también posee existencia «potencial» y es más amplio que el S actual.

Lo mismo puede decirse de no poder, deber o no poder. Así, las modalidades naturales (y, por ende, temporales) se refieren a diferentes grados o niveles de existencia.

De igual modo, la inmutabilidad de las relaciones de pertenencia significa que son externas a sus términos, sin dejar huella en ellos, incluso cuando son reales.

Parece una postura razonable, porque si cada relación de algo con todo lo demás implicara algún rasgo correspondiente dentro de ese algo, entonces cada cosa en el mundo tendría que contener un número infinito de mensajes, un mensaje por cada una de sus relaciones con las demás. Mucho más sencillo es considerar las relaciones (al menos, aquellas que son inmutables) como si tuvieran una existencia separada de sus términos, como otros contenidos del universo.

4. La paradoja del barbero

La paradoja del barbero [19]se puede formular así: «Si un barbero afeita a todos los habitantes de su pueblo que no se afeitan, ¿se afeita o no se afeita? Si se afeita, no se afeita; si no se afeita, sí se afeita».

Esta doble paradoja surge de la confusión entre las expresiones "no se afeita él mismo" y "es afeitado por otra persona".

Podemos dividir a la gente de cualquier ciudad en tres grandes grupos: (a) personas que no se afeitan, sino que son afeitadas por otros; (b) personas que no se afeitan ni son afeitadas por otros; (c) personas que se afeitan y no

---

[19] Esta paradoja fue propuesta por B. Russell (en su obra de 1918-19, *La filosofía del atomismo lógico*) como ilustración de la paradoja de Russell; sin embargo, no la reivindicó como propia, afirmando que le había sido «sugerida» por otra persona. No obstante, Russell consideraba que «en esta forma la contradicción no es muy difícil de resolver», ya que basta con negar el sujeto (es decir, afirmar que tal barbero no existe).

son afeitadas por otros. La premisa dada es que nuestro barbero afeita a todas las personas que pertenecen al grupo (a). Se sugiere tácitamente, pero no se implica formalmente, que nadie pertenece al grupo (b), por lo que nadie se deja barba ni necesita afeitarse. Pero, en cualquier caso, la premisa no nos dice nada sobre el grupo (c).

A continuación, subdividamos cada uno de los grupos anteriores en dos subgrupos: (i) personas que afeitan a otros y (ii) personas que no afeitan a otros. Es evidente que cada una de las seis combinaciones resultantes es lógicamente aceptable, puesto que quién me afeita no influye en a quién puedo afeitar. Obviamente, solo el grupo (i) se refiere a los barberos, y nuestra premisa puede interpretarse como que nuestro barbero es el único barbero de la ciudad.

Ahora bien, podemos abordar la pregunta planteada. Nuestro barbero no puede pertenecer al grupo (a)(i), ya que nadie lo afeita. Podría pertenecer al grupo (b)(i) si se le permitiera dejarse barba o si fuera lampiño; pero supongamos que no, a modo de ejemplo. Esto aún le permite pertenecer al grupo (c)(i), lo que significa que se afeita a sí mismo (en lugar de que otros lo afeiten), aunque también afeita a otros.

Por lo tanto, no existe una doble paradoja. La doble paradoja surgió únicamente porque asumimos erróneamente que «afeita a todos los que *no* se afeitan» excluye a «afeita a algunos (como él mismo) que *sí* se afeitan». Pero *«X afeita a Y» no contradice formalmente «X afeita a los que no son Y»* ; no hay base para suponer que la cópula «afeitar» sea obvertible, de modo que «X afeita a Y» implique «X no afeita a los que no son Y».

Si la premisa se reformulara como 'él afeita todos esos *y Si solo* se incluye la opción «aquellos que no se afeitan» (para excluir «él se afeita»), aún tendríamos una salida

diciendo «él no se afeita en absoluto». Si la premisa se ampliara y restringiera aún más insistiendo en que «de alguna manera se afeita o es afeitado», sería simplemente contradictoria (una paradoja).

Se podrían añadir más detalles a lo anterior, como considerar a las personas que se afeitan en otras ciudades o distinguir entre siempre, a veces/a veces no y nunca. Pero creo que la idea principal queda clara. La lección que se desprende de la «paradoja» del barbero es que, sin categorizaciones claras, pueden surgir ambigüedades (como la que existe entre «afeita» y «está afeitado»), lo que crea la ilusión de una doble paradoja.

5. La paradoja del catálogo maestro

Una clase puede considerarse como un sobre imaginario que envuelve de forma flexible a todos sus supuestos miembros, por muy dispersos que estén en el espacio y el tiempo, excluyendo todo lo demás. Surge entonces la pregunta: ¿puede el sobre figurado de la clase «clases» envolverse también a sí mismo, o no?

Al analizar la **paradoja de Russell** [20], debemos concluir que *no todos los «objetos-palabra» son «cosas»* ; las medidas de las cosas no pueden considerarse cosas en sí mismas. Dado que la clasificación es una expresión de nuestra medición de las cosas, no puede considerarse una cosa. Hacerlo genera una paradoja que debemos evitar.

En otras palabras, el problema radica en que la forma iterativa («clase de clases») no es idéntica a la forma simple («clase»), salvo de forma muy superficial y verbal; por lo tanto, la primera no puede subsumirse lógicamente bajo la segunda. Existe una modificación suficientemente significativa de la relación sujeto-predicado, causada por

---

[20] Véase *Lógica futura* , capítulos 43-45, sobre lógica de clases.

la iteración del mismo término, que excluye el reflejo de subsunción. La paradoja que surge si no refrenamos este impulso es precisamente lo que nos enseña a ejercer dicho control.

Cabe señalar que la palabra «cosas» tiene muchos significados. A veces, nos referimos a todos los objetos posibles del pensamiento. Otras veces, queremos excluir las palabras [21]. A veces, queremos excluir clases; o, más específicamente, como se acaba de señalar, clases de clases; lo mismo ocurre con los conceptos o los conceptos de conceptos. A veces, la palabra «cosas» incluye solo objetos materiales, cualquiera que sea su categoría. Otras veces, nos referimos a «entidades [22]» (cuerpos materiales, mentales o espirituales, o sustancias delimitadas, cuyos casos individuales suelen ser sujetos de proposiciones) en contraposición a sus «propiedades» (los predicados de lugar, tiempo, cualidad, acción, cantidad, relación, etc.). A veces, en el discurso cotidiano, nos referimos a «cosas» en contraposición a «personas», es decir, aquí «cosas» significa entidades inanimadas o no volitivas. Y existen aún más sentidos de la palabra.

Por lo tanto, siempre que los lógicos se refieren a "cosas", deberían intentar primero aclarar qué es lo que se incluye bajo ese epígrafe.

Por cierto, aún peor que el concepto de «autopertenencia» es la noción de « *clases que parecen contradecir lo que incluyen* ». Esta última parece inconcebible, al menos en apariencia. Así, por ejemplo: «ninguna relación» es una especie de relación; «no clases» es, en cierto sentido, una clase. Debe haber alguna falacia en tales términos, que

---

[21] Aunque, por supuesto, esta distinción puede resultar paradójica, ya que la palabra "palabra" se refiere a palabras.

[22] Por supuesto, la palabra "entidad" a veces se usa de forma más general, haciendo referencia a cualquier cosa existente.

necesita ser aclarada. ¿Quizás el problema radica en una hipérbole o un término inapropiado?

La respuesta a esta pregunta sería que aquí estamos nuevamente tratando con *clases de clases* , y estas no necesariamente son coherentes externamente con sus clases miembro. Así, la clase de no-relaciones aún implica una relación. La clase de no-clases es, no obstante, una clase. La clase de clases vacías o nulas sí tiene miembros. La clase de clases sin sentido o autocontradictorias no es ni sin sentido ni autocontradictoria. Y así sucesivamente.

Bertrand Russell ilustra su paradoja haciendo referencia a:

(a) un catálogo de todos los libros que se mencionan a sí mismos, y

(b) un catálogo de todos los libros que *no* se mencionan a sí mismos.

El caso (a) no presenta ningún problema: el catálogo puede incluirse a sí mismo sin contradecir su propia definición; mientras que, si no se incluye, traiciona dicha definición [23]. El caso (b), en cambio, sí presenta un problema: si no se incluye a sí mismo, de acuerdo con su propia definición, se vuelve susceptible de ser incluido en sí mismo [24]; pero, si se incluye, contradice su propia definición. Esta última es la doble paradoja que se analiza.

Ahora bien, mi primera objeción sería la siguiente. El título del catálogo (e incluso, tal vez, una breve descripción de su contenido, un resumen) podría figurar dentro del propio libro, pero dicho libro no podría ni podría incluir una reproducción completa del mismo (por no hablar de todos los demás libros que enumera o reproduce), por la sencilla razón de que la tarea sería

---

[23] Es decir, el catálogo no reúne los requisitos para ser incluido en sí mismo, pero eso no afecta a su exhaustividad.

[24] De modo que, si no se incluye inmediatamente en sí mismo, ya no se puede afirmar que esté completo.

infinita (un libro dentro de un libro dentro de un libro... etc., o lo mismo en plural).

Por lo tanto, el libro no es un elemento *independiente* ; en sentido estricto, solo se mencionan en él palabras *que* lo describen. Los términos inclusión o pertenencia, tal como se usan aquí, tienen un significado muy limitado. Así pues, la plausibilidad del ejemplo de Russell es muy superficial y engañosa; está siendo falaz y sofista, sugiriendo algo imposible.

Además, cada libro se «incluye a sí mismo» en el sentido de que consta únicamente de su contenido. Pero si se concibe un libro como un conjunto de otros libros, definidos por alguna afirmación (por ejemplo, todos los libros en inglés), el libro no puede incluirse a sí mismo en el sentido de que dicho contenido sea *solo una parte del* mismo. Esto no solo implicaría una regresión infinita (un libro con otros libros más él mismo, estos a su vez con otros libros más él mismo, y así sucesivamente), y un tamaño infinito, sino que constituiría una contradicción en la propia definición. El libro no puede ser a la vez todo su contenido y solo una parte de él.

Desde esta perspectiva, al definir el libro como "el catálogo de todos los libros que no se incluyen a sí mismos", la paradoja de Russell es similar a la paradoja del mentiroso, ya que el libro proyectado es una entidad que no tiene una dimensión finita; nunca se puede concretar.

Una segunda objeción sería la siguiente. Incluso si consideramos la construcción de Russell como una simple lista de libros, definida como «el catálogo de todos los libros que no *se mencionan* a sí mismos», la definición es absurda, ya que *no puede realizarse lógicamente* . Sencillamente no podemos escribir un libro que enumere todos los libros que no se mencionan a sí mismos (*Lord Jim de Conrad* , *Notre Dame de Hugo* , etc.), dada la

contradicción que supone incluir o no el libro en sí. Por lo tanto, este concepto es necesariamente una clase nula y carece de sentido.

La lógica no se ha visto desconcertada por la paradoja, sino que simplemente se le ha enseñado que el concepto propuesto es erróneo e inutilizable; por lo tanto, debe descartarse o, al menos, modificarse. No hay nada dramático en la paradoja; representa una de las funciones de la lógica. Podríamos intentar proponer un concepto modificado, como se muestra a continuación. Quizás deberíamos referirnos a una biblioteca.

(a) Consideremos un catálogo de todos los libros de una biblioteca, que se colocará en esa misma biblioteca. Si el libro se menciona a sí mismo, no hay problema. Si el libro no se menciona como perteneciente a la biblioteca, simplemente está incompleto y debe ampliarse; o su título es incorrecto y debe modificarse («todos los libros excepto este»); o debe quedar fuera de la biblioteca.

b) Ahora bien, con respecto a un catálogo de todos los libros *que no* están en nuestra biblioteca: un libro de este tipo no puede mencionarse a sí mismo y, al mismo tiempo, estar en la biblioteca. Si queremos conservarlo en nuestra biblioteca, debemos eliminar cualquier mención de sí mismo. Si queremos que se mencione a sí mismo, debemos excluirlo de la biblioteca. Estas son alternativas prácticas que no presentan ningún problema.

Desde esta perspectiva, al buscar una expresión práctica para ella, la paradoja de Russell se asemeja más a la paradoja de Barber.

6. La paradoja de Grelling

Para desarrollar su paradoja [25], Kurt Grelling [26]etiqueta una palabra como "homológica" si posee la cualidad a la que se refiere (por ejemplo, la palabra "corto" es corta, o la palabra "polisílabo" es polisílaba), o "heterológica" si carece de la cualidad a la que se refiere (por ejemplo, "largo" carece de longitud, o de nuevo "monosílabo" no es monosílabo). Luego se pregunta si estas dos palabras, en sí mismas, deben categorizarse de una u otra manera, argumentando:

- Si “heterológico” es homológico, entonces es heterológico (predicados contradictorios).
- Si “heterológico” es heterológico, entonces es homológico (predicados contradictorios).

Pero es un error de interpretación aplicar estas palabras a sí mismas. En su caso, las referencias son *demasiado abstractas* para tener concomitantes visibles o audibles. *Ninguno de* los términos es aplicable a ninguna de ellas.

Nótese primero que las aparentes contradicciones en la predicación, en ambos sentidos, se aplican solo a la palabra “heterológico”. Pues, utilizando un razonamiento similar con respecto a la palabra “homológico”, aunque podría parecer más coherente decir que “homológico” es homológico que decir que es heterológico, la secuencia de predicados parecería coherente en ambos sentidos, es decir:

- Si “homológico” es homológico, entonces es homológico (predicados consistentes).

---

25 Esta paradoja se inspiró en la paradoja de Russell.

26 Alemania, 1886-1942. La paradoja también se conoce como la paradoja de Grelling-Nelson, porque fue presentada en un artículo de 1908 escrito conjuntamente con Leonard Nelson (Alemania, 1882-1927).

- Si “homológico” es heterológico, entonces es heterológico (predicados consistentes).

Esto podría interpretarse como que el término homológico está mejor construido, mientras que el término heterológico presenta una falla estructural. Pero este no es el verdadero problema aquí.

La cuestión fundamental reside en distinguir entre los términos «homológico» y «heterológico» y sus respectivos significados, a saber, homológico y heterológico. Cuando nos referimos a una palabra como tal, tradicionalmente la colocamos entre comillas; y cuando nos referimos a su significado asignado, la usamos simplemente. En las proposiciones anteriores, de las que aparentemente surge una paradoja, los sujetos son las palabras como tales (entre comillas) y los predicados son los significados de dichas palabras.

Desde esta perspectiva, no hay fundamento para afirmar que «heterológico» implica que «heterológico» es homológico, o viceversa. La inferencia es muy superficial, pues confunde la palabra en sí (entendida como sujeto) con su significado (entendido como predicado). Es decir, las comillas en el sujeto no se usan con intención, sino que ocultamos el significado subyacente como sujeto.

¿Cómo dedujimos los consecuentes de los antecedentes? ¿Podíamos ver a simple vista que la primera tesis implica la segunda? Analicemos con más detalle las proposiciones hipotéticas en cuestión:

Si en el antecedente ponemos el énfasis en *la propiedad* a la que se refiere la palabra “heterológico”, es decir, alguna cualidad supuesta llamada heterologicalidad, formularíamos las paradojas de la siguiente manera:

- Si la palabra “heterológico” tiene la propiedad a la que se refiere (es decir, es heterológico), entonces

aparentemente carece de la propiedad a la que se refiere (es decir, es homológico).

- Si la palabra "heterológico" carece de la propiedad a la que se refiere (es decir, es homológica), entonces aparentemente tiene la propiedad a la que se refiere (es decir, es heterológica).

Si, por otro lado, en el antecedente ponemos el énfasis en la palabra "heterológico", *que tiene o carece* de la propiedad a la que se refiere, formularíamos las paradojas de la siguiente manera:

- Si la palabra "heterológico" tiene la propiedad a la que se refiere (es decir, es homológico), entonces aparentemente carece de la propiedad a la que se refiere (es decir, es heterológico).
- Si la palabra "heterológico" carece de la propiedad a la que se refiere (es decir, es heterológico), entonces aparentemente tiene la propiedad a la que se refiere (es decir, es homológico).

En cualquiera de estos casos, el consecuente se construye comparando el sujeto «heterológico» con el predicado antecedente «heterológico» u «homológico»; si son la misma palabra, «inferimos» «homológico» como nuestro predicado consecuente, mientras que si difieren verbalmente, «inferimos» «heterológico». Pero, en realidad, al realizar estas comparaciones entre el sujeto y el predicado antecedente, no hemos detectado ninguna cualidad en la palabra «heterológico» como tal, sino que nos hemos referido tácitamente a su significado subyacente y lo hemos contrastado con el predicado hipotético.

En otras palabras, la afirmación de que lo «heterológico» es homológico (o, de hecho, que lo «homológico» es heterológico) no es tan contradictoria como parece a primera vista; podría ser coherente. En realidad, es indeterminada y, por lo tanto, carece de sentido.

Más precisamente, para resolver la paradoja, debemos recordar cómo se originaron nuestros términos. Podemos afirmar que "corto" es corto *simplemente viendo u oyendo la palabra* (suponiendo que cualquier sílaba, independientemente de cómo se escriba o pronuncie, se considere corta). Pero en el caso de un término como "heterológico", no podemos determinar si la palabra posee o carece de la propiedad a la que se refiere, porque dicha propiedad no es una cualidad concreta (visible o audible) de la palabra, sino algo abstracto que aplicamos a los componentes visibles o audibles de las palabras. Si la cualidad buscada no es visible ni audible, es incognoscible y no hay manera de saber qué predicado se aplica.

Es decir, nuestras definiciones iniciales de esos términos, que mencionan “una palabra que posee o carece de cierta cualidad a la que se refiere”, no son claras ni precisas, porque no especifican, como deberían, que las cualidades a las que se refieren son fenomenológicas, es decir, aspectos perceptibles de la palabra. Si la palabra designa algo que no está incluido en sus aspectos físicos, los términos homológico y heterológico simplemente no se aplican. Aplicarlos es recurrir a artificios verbales. Por lo tanto, ninguno de estos predicados es aplicable a ninguna de estas palabras en sí.

Podría objetarse que las palabras poseen atributos no fenoménicos. Por ejemplo, solemos considerar una palabra útil o inútil. En tal caso, podríamos preguntarnos: ¿es útil o no la palabra «útil»? Sí, respondería. Por lo tanto, «útil» es homológica. Del mismo modo, «inútil» es útil, por lo que «inútil» es heterológica. Desde esta perspectiva, cabe dudar de la exactitud de lo que acabamos de proponer: que «homológico» y «heterológico» son términos que presuponen predicados concretos (en lugar de abstractos).

Pero a esta objeción se podría replicar que la utilidad de una palabra es, en última instancia, algo concreto: una palabra es útil si marca una diferencia práctica perceptible en el desarrollo del conocimiento. En ese caso, nuestra definición podría modificarse ligeramente, especificando que los términos homológico o heterológico solo son aplicables cuando podemos vincularlos, *directa o indirectamente, a alguna propiedad concreta.*

En resumen, *estos términos deben referirse a algo distinto de sí mismos antes de poder ser utilizados* . La falacia implicada es similar a la de la paradoja del mentiroso, donde el término «esto» se usa en referencia a sí mismo, mientras que solo adquiere significado cuando tiene algo más a lo que referirse. Dichos términos son *relacionales* y, por lo tanto, no pueden referirse a otras relaciones de forma circular o hasta el infinito: necesitan, en última instancia, estar anclados a algún término no relacional.

Nótese, por cierto, que si sustituyéramos la palabra «short» por «shortissimo», manteniendo el mismo significado, la palabra cambiaría de estatus y se convertiría en heterológica, puesto que «shortissimo» no es shortissimo. Por otro lado, independientemente de la palabra que sustituyamos por «heterológico», la paradoja de Grelling al respecto sigue siendo evidente. Esta prueba demuestra que, en este último caso, no se trata simplemente de la palabra en sí, sino de su significado subyacente. En cuanto a la palabra «útil», podríamos decir que es útil por su contenido, o a lo sumo por el hecho de ser una palabra (una unidad del lenguaje), y no por su forma o sonido específicos.

8. La paradoja de Russell (revisión)

De A Fortiori Logic, apéndice 7.5.

La lógica es lo que nos ayuda a transformar percepciones concretas dispersas en conceptos abstractos bien ordenados. El conocimiento humano, o la opinión, se basa en la experiencia, la imaginación y la intuición racional. Esta última es una especie de «experiencia» en un sentido más amplio, una experiencia no fenoménica, que podríamos llamar «intuición» lógica. Por esta razón, los antiguos, tanto en Occidente como en Oriente, denominaban a la razón el «sexto sentido» o «sentido común», es decir, el órgano sensorial que une los otros cinco sentidos, aquellos que nos ponen en contacto empírico con la experiencia fenoménica: colores, formas, sonidos, olores, sabores, sensaciones táctiles, etc., ya sean percibidos físicamente o imaginados mentalmente. Los cinco sentidos sin el sexto producen un sinsentido caótico (carecen de sentido, no se les puede dar sentido); y, a la inversa, el sexto sentido es inútil sin los otros cinco, porque carece de información sobre la que tener intuiciones racionales. La imaginación reorganiza los datos y razonamientos experienciales del pasado, posibilitando la formación de nuevas ideas y teorías que posteriormente se ponen a prueba con referencia a nuevas experiencias y razonamientos.

1. Elementos de la lógica de clases

La lógica se desarrolló inicialmente como una ciencia principalmente en referencia al discurso natural, dando

como resultado lo que hoy conocemos como lógica de predicados. En el discurso humano natural, nosotros (tú, yo y todos los demás) pensamos y discutimos rutinariamente cosas que hemos percibido, o eventualmente concebido, mediante proposiciones categóricas que involucran un sujeto (por ejemplo, S) y un predicado (por ejemplo, P) que están relacionados entre sí mediante la cópula "es". Dichas proposiciones tienen la forma "S es P", que puede ser singular o plural, y en este último caso general (o universal) o particular, y positiva o negativa, y además puede involucrar varios modos y categorías de modalidad [27].

Una proposición de la forma «S es P» es en realidad una doble predicación: nos dice que algo que es S también es P; por lo tanto, S y P son en realidad predicados, aunque uno (el sujeto S) tiene prioridad en el pensamiento para «predicar» al otro (el predicado P) [28]. Principalmente, S se refiere a algún fenómeno concreto (ya sea físico, mental o espiritual), es decir, una entidad individual o un conjunto de entidades, y P a una propiedad de este o de ellos. Por ejemplo, «Juan es un hombre» y «Todos los hombres son seres humanos» son, respectivamente, una predicación singular (sobre un hombre, Juan) y una plural (sobre todos los hombres).

Además, en el discurso natural, el sujeto de nuestros pensamientos pueden ser predicados *como tales* , es decir, predicados en su calidad de predicados; un ejemplo es: «“hombres” puede ser el sujeto o predicado de una

---

[27] No es necesario entrar en detalles sobre estas distinciones, ya que son bien conocidas. También existen muchas sutiles diferencias entre los distintos tipos de términos que pueden aparecer en las proposiciones como sujetos o predicados; pero simplifiquemos el asunto.

[28] El término "predicación" se refiere a la cópula y al predicado juntos, como si fueran una acción del hablante (o la afirmación hecha) sobre el sujeto.

proposición». Esto último ocurre en el discurso específicamente filosófico (o lógico o lingüístico); por ejemplo, en el presente ensayo.

Ahora bien, los lógicos a lo largo de los siglos, y especialmente en los tiempos modernos, han encontrado que el discurso natural resulta algo inadecuado para sus necesidades y han desarrollado gradualmente un lenguaje más artificial, el de las "clases" [29]. Este tipo de discurso es exactamente paralelo al discurso natural, pero es un poco más abstracto y descriptivo para facilitar el discurso filosófico (o lógico o lingüístico) y hacerlo más preciso. En este lenguaje, en lugar de decir "este S *es* P", decimos "este S *es un miembro (o instancia) de* P" (nótese bien el alargamiento de la cópula de "es" a "es un miembro (o instancia) de"). Si "este S" simboliza un individuo concreto, entonces "P" aquí se llama " *clase* "; pero si "este S" simboliza una clase abstracta, entonces "P" aquí se llama " *clase de clases* ".

Una clase, entonces, es una abstracción, una construcción mental en la que agrupamos figurativamente algunas cosas concretas (ya sean físicas, mentales o espirituales). Si bien podemos clasificar mentalmente las cosas temporalmente sin nombrarlas, normalmente les asignamos un nombre (es decir, una palabra o frase distintiva) porque esto facilita la memoria y la comunicación. Nombrar no es la esencia de la clasificación, pero sí facilita enormemente la clasificación a gran escala. El nombre de una clase de cosas no las

---

[29] La siguiente explicación de la lógica de clases se basa en mi presentación en *Lógica del Futuro* , capítulos 43-45. La palabra «clase» proviene del latín *classis* , que se refiere a un «grupo llamado al servicio militar» (Merriam-Webster). Desconozco si los antiguos usaban esa palabra en su sentido lógico, o alguna similar, en su discurso, pero sin duda pensaban según la lógica de clases. Ejemplos de pensamiento de clases son la distinción de Aristóteles entre especies y géneros y el árbol de Porfirio.

representa como un símbolo, sino que más bien las dirige hacia ellas o llama nuestra atención sobre ellas; es decir, es un instrumento de intención.

Una clase, en su sentido primario, es una clase de cosas en general; una clase, en su sentido secundario, es más específicamente una clase *de clases* . Por lo tanto, la pertenencia es de dos tipos: pertenencia de elementos que no son clases a una clase, o pertenencia de clases a una clase de clases. Alternativamente, podemos hablar de clases de primer orden y clases de segundo orden para distinguir estos dos tipos. No existen otros órdenes de clases. Cuando pensamos o hablamos de cosas más concretas, hablamos en lógica de clases de primer orden; cuando pensamos o hablamos de clases de primer orden, hablamos en lógica de clases de segundo orden, y esto último también se aplica a las clases de segundo orden, ya que, al fin y al cabo, también son clases. Los dos órdenes de clases no deben confundirse con la jerarquía de clases dentro de cada orden.

La relación entre clases de clases y clases es análoga a la relación entre clases y concretos; es una relación de subsunción. Cuando una clase inferior (es decir, de primer orden) pertenece a una clase superior (es decir, de segundo orden), no se *deduce* que los miembros de la clase inferior también pertenezcan a la clase superior; de hecho, si pertenecen a una, ciertamente no pertenecen a la otra. Así, por ejemplo, tú y yo, aunque pertenecemos a la clase «hombres» porque somos hombres, no pertenecemos a la clase «clases de hombres» porque no somos «hombres». Asimismo, la clase «hombres» no es un hombre, sino que pertenece a la clase «clases de hombres». Los miembros de la clase «clases de hombres» (o, dicho más brevemente, «clases de hombres»), que es una clase de clases, son, además de la clase amplia

«hombres», las clases más específicas «jardineros», «ingenieros», «sabios», «neuróticos», etc.[30]

Por otro lado, la jerarquización se refiere a clases *dentro de un orden dado* que comparten instancias, no solo por superposición parcial, sino de tal manera que todos los miembros de una clase son miembros de la otra (y en algunos casos, pero no en todos, viceversa). Por ejemplo, dado que todos los hombres son animales, aunque no todos los animales son hombres, la clase 'hombres' es *una subclase (o especie) de* la clase 'animales', y la clase 'animales' es *una superclase (o género) de* la clase 'hombres'. Si dos clases tienen las mismas instancias, ni más ni menos, se puede decir que son clases *coextensivas* (una clase que sirve como especie y género en algún contexto se dice que es *sui generis*). Si dos clases solo comparten algunas instancias, se puede decir que son clases que *se intersecan (o se superponen)* , pero no están organizadas jerárquicamente (por ejemplo, 'jardineros' e 'ingenieros'). Si dos clases del mismo orden no tienen instancias en común, se puede decir que son clases *mutuamente excluyentes* .

Es importante comprender y tener presente la distinción entre jerarquía y orden. Dado que tú y yo somos hombres, cada uno de nosotros es miembro de la clase 'hombres'; esto es subsunción por una clase de primer orden de sus instancias concretas. Dado que todos los hombres son animales, la clase 'hombres' es una subclase de la clase

---

30 Nótese que decir o escribir la palabra «hombres» sin comillas se refiere a un *predicado* . Cuando queremos referirnos a la *clase correspondiente* , decimos «la clase de hombres» o «la clase hombres»; si escribimos, podemos escribir lo mismo con o sin comillas, o simplemente «hombres» entre comillas. Cuando tratamos de clases de clases, decimos «la clase de clases de hombres» o «la clase de hombres-clases» o «la clase hombres-clases», y podemos escribir lo mismo con o sin comillas, o simplemente «clases de hombres» o «hombres-clases» entre comillas.

'animales'; esto es jerarquía entre dos clases de primer orden. Dado que 'hombres' es una clase de animales, es miembro de la clase 'clases de animales' (o 'clases animales'); esto es subsunción por una clase de segundo orden (es decir, una clase de clases) de sus instancias de clase de primer orden (es decir, meras clases). Dado que todas las 'clases de hombres' son 'clases de animales', la clase 'clases hombres' es una subclase de la clase 'clases animales'; esto es jerarquía entre dos clases de segundo orden, es decir, entre dos clases de clases. La relación entre clases de primer orden y clases de segundo orden nunca es de jerarquía, sino siempre de subsunción; Es decir, los primeros son siempre miembros (instancias) de los segundos, nunca subclases. Las jerarquías solo se dan entre clases del mismo orden.

Así pues, en la lógica de clases, tenemos dos planos de existencia que considerar. En el nivel básico se encuentra el plano relativamente objetivo de los fenómenos empíricos (ya sean de naturaleza física, mental o espiritual); por encima de este, en nuestra mente, se halla el plano relativamente subjetivo de las ideas (concebidas como insustanciales, pero que poseen aspectos fenoménicos: imágenes mentales o físicas, palabras habladas o escritas y las intenciones que transmiten), que comprende ideas sobre fenómenos empíricos e ideas sobre dichas ideas. Las clases se desarrollan para facilitar el estudio de los fenómenos empíricos, y las clases de clases se desarrollan a su vez para facilitar el estudio de las clases, pues las clases (incluidas las clases de clases) son, por supuesto, fenómenos empíricos en sí mismas. La clasificación es una invención humana útil para la ordenación cognitiva de las cosas observadas a través de nuestros sentidos, nuestra imaginación o nuestras intuiciones introspectivas. Si bien las clases son producto de la mente, esto no significa que sean arbitrarias: se

forman, organizan y controlan mediante nuestra facultad racional, es decir, con la ayuda de la lógica.

Evidentemente, para ser considerada una clase, debe tener al menos un miembro (en cuyo caso, el único miembro es único). Generalmente, una clase tiene dos o más miembros, incluso innumerables. Una clase es finita si incluye un número específico de instancias; si el número de instancias es difícil de enumerar, se dice que la clase es abierta (es decir, infinita o al menos indefinida). Lo que une a las instancias de una clase es que poseen alguna propiedad distintiva en común; la clase se define por esta propiedad (que, por supuesto, puede ser una compleja combinación de muchas propiedades). Una clase sin instancias se denomina clase nula (o vacía); esto significa que su propiedad definitoria es ficticia, por lo que, estrictamente hablando, no es una clase.

Así pues, cabe señalar que el término «clase» es algo ambiguo, ya que puede referirse a una clase de primer orden (una clase de no clases, es decir, de cosas distintas de las clases) o a una clase de segundo orden (una clase de clases, es decir, una construcción mental que agrupa dos o más de dichas construcciones mentales). Una clase (de primer orden) no es, ni puede ser, una clase de clases (es decir, una clase de segundo orden). Existe, por supuesto, una clase denominada «no clases»; sus instancias son principalmente todas las cosas concretas, que no son clases en sí mismas; por ejemplo, tú y yo somos no clases. «No clases» es simplemente una clase, no una clase de clases, puesto que no incluye ninguna clase. Por lo tanto, se puede decir que «no clases» es una clase de primer orden, pero no califica como una clase de segundo orden.[31]

---

[31] Cabe señalar que, si bien los términos positivos son relativamente fáciles de traducir al lenguaje de la lógica de clases, los términos negativos presentan una verdadera dificultad. Por ejemplo,

El ámbito de las clases de clases es muy limitado como objeto de estudio en comparación con el ámbito de las clases simples. ¿Qué distinciones podemos establecer entre clases? No muchas. Podemos distinguir entre clases y clases de clases, entre clases finitas y abiertas, entre clases positivas y negativas [32], y quizás algunas más, pero no mucho más.

2. Una aparente doble paradoja

Bertrand Russell (Gran Bretaña, 1872-1970) propuso una distinción entre «una clase que es miembro de sí misma» y «una clase que no lo es». Si bien toda clase es necesariamente coextensiva consigo misma (y en este sentido está incluida en sí misma), no se deduce que toda clase sea miembro de sí misma (evidentemente, algunas lo son y otras no). Esta distinción puede demostrarse como legítima mediante ejemplos convincentes. Así, la clase «clases positivas» es miembro de sí misma, puesto que se define por una propiedad positiva; mientras que la clase «clases negativas» no lo es, puesto que también se

---

mientras que el término «hombres» se refiere únicamente a no clases, su antítesis estricta, el término «no hombres» en su sentido más amplio, incluye tanto no clases (es decir, cosas concretas distintas de los hombres) como clases (es decir, cosas más abstractas). Asimismo, mientras que el término «clases finitas» se refiere únicamente a clases, su antítesis estricta, el término «clases no finitas» en su sentido más amplio, incluye tanto clases abiertas (abstractas) como no clases (concretas). Por lo tanto, en aras de la coherencia, debemos admitir que algunos términos abarcan tanto no clases como clases (incluidas clases de clases). En la práctica, esto significa que debemos utilizar *disyunciones* que revelen las alternativas implícitas. Esto, por supuesto, complica considerablemente la lógica de clases.

[32] Las clases positivas se definen por alguna propiedad positiva y las clases negativas por una negativa. Por ejemplo, «hombres» se define en referencia a animales racionales (positivo), mientras que «solteros» se define en referencia a hombres que aún no se han casado (negativo).

define positivamente (aunque con referencia general a la negación). De igual modo, la clase «clases finitas» no es miembro de sí misma, puesto que tiene innumerables miembros; mientras que la clase «clases abiertas» sí lo es, puesto que también tiene innumerables miembros.

¿Qué ocurre con la clase «clases»? ¿Es miembro de sí misma o no? Dado que «clases» es una clase, debe ser miembro de «clases», es decir, de sí misma. Esto se afirma sin tener en cuenta la distinción entre clases de primer y segundo orden. Si formulamos la pregunta con mayor precisión, la respuesta debe matizarse. La clase «clases de primer orden», al ser una clase de clases y no una clase mera, no puede ser miembro de sí misma, sino solo miembro de «clases de segundo orden»; los miembros de las «clases de primer orden» son todas clases meras. Por otro lado, dado que la clase «clases de segundo orden» es una clase de clases, es miembro de sí misma, es decir, miembro de «clases de segundo orden». Por lo tanto, la clase «clases de segundo orden» se incluye a sí misma y a la clase «clases de primer orden», de modo que cuando decimos que la clase más amplia «clases» es miembro de sí misma, nos referimos más específicamente a que es miembro de la clase más restringida «clases de clases». En cuanto a la clase «no-clases», dado que es una clase y no una no-clase, no es miembro de sí misma. Cabe señalar, sin embargo, que la paradoja de Russell no distingue entre clases de primer y segundo orden, sino que se centra en las «clases» indiscriminadamente.

Russell preguntó **si «la clase de todas las clases que no son miembros de sí mismas» es o no es miembro de sí misma** . Parecía lógicamente imposible responder a la pregunta, porque en ambos casos se producía una contradicción. Pues si la clase «clases que no son miembros de sí mismas» *no es* miembro de la clase «clases que no son miembros de sí mismas», entonces sí es miembro de «clases que no son miembros de sí

mismas» (es decir, de sí misma); y si la clase «clases que no son miembros de sí mismas» *es* miembro de «clases que no son miembros de sí mismas», entonces también es miembro de «clases que son miembros de sí mismas» (es decir, de su contradictoria). Esto parecía una doble paradoja alucinante.

La solución del problema . La búsqueda del conocimiento es una empresa humana y, por lo tanto, se basa en el método de ensayo y error. El conocimiento es mucho más inductivo que deductivo; la deducción es solo una de las herramientas de la inducción. Existen verdades absolutas en el conocimiento humano, pero son escasas y distantes entre sí. Cuando formulamos una teoría, siempre es esencialmente una hipótesis, que posteriormente podría necesitar ser revisada o descartada. Mientras parezca útil y sólida, y lo sea más que cualquier teoría competidora, la adoptamos; pero si alguna vez resulta ser refutada por algunos hechos o produce antinomia, estamos obligados a reformularla o descartarla. Este es el principio de inducción. Cuando nos encontramos con una contradicción, debemos «revisar nuestras premisas» y modificarlas según sea necesario. En el caso que nos ocupa, dado que la paradoja de Russell demuestra que nuestra concepción de la lógica de clases es errónea, debemos retroceder y averiguar dónde nos equivocamos. Así pues, repasemos cuidadosamente nuestros pasos. Definimos una clase y la pertenencia a una clase transformando la predicación en clasificación, diciendo en efecto:

> Si algo es X, entonces es miembro de la clase 'X', y no es miembro de la clase 'nonX'.
>
> Si algo no es X, entonces no es miembro de la clase 'X', sino miembro de la clase 'nonX'.

¿De dónde proviene esta definición? No es un absoluto que nos haya sido impuesto cognitivamente. La

inventamos nosotros mismos; fue una convención mediante la cual concebimos la idea de clases y la pertenencia a ellas. El conocimiento puede perfectamente avanzar sin recurrir a esta idea, y así ha sido durante milenios y continúa haciéndolo en la mente de muchas personas. Es una idea con una historia, que fue incorporada al arsenal de técnicas de razonamiento por lógicos de épocas relativamente recientes. Estos lógicos observaron que ellos mismos y otros razonaban mediante la clasificación, y se dieron cuenta de que se trataba de un artificio útil, distinto de la predicación, pero basado de alguna manera en ella. Por lo tanto, propusieron formalmente la definición anterior y procedieron a estudiar el tema con mayor detalle para maximizar su utilidad. Así nació la «lógica de clases» o «lógica de clases».

Sin embargo, en cierto momento, el lógico Bertrand Russell se percató de una inconsistencia inherente en nuestra concepción de la clasificación, que ponía en seria duda todo el entramado de la lógica de clases. Ese fue el descubrimiento de la paradoja que lleva su nombre. Fue un hallazgo trascendental, pues no hay nada más importante para el desarrollo del conocimiento, y especialmente para el desarrollo de la rama del conocimiento conocida como lógica formal, que el mantenimiento de la coherencia. Cada descubrimiento de una inconsistencia estimula el perfeccionamiento de nuestro conocimiento. Russell merece un gran reconocimiento por este hallazgo, aunque tuvo muchas dificultades para resolver la paradoja de manera totalmente convincente. Intentemos aquí ir más allá, profundizando en los procesos de pensamiento involucrados en la clasificación. ¿Qué es la clasificación, más precisamente?

Si analizamos con mayor detenimiento nuestra definición anterior de una clase 'X' y la pertenencia de las cosas a

ella por el simple hecho de ser X, debemos preguntarnos: ¿qué logra concretamente esta definición? ¿Acaso solo estamos sustituyendo la frase 'es miembro de' por la cópula 'es' y la clase 'X' por el predicado X? Si esto es lo que estamos haciendo, no tiene sentido, pues es obvio que cambiar *el nombre* de una relación o un término no lo afecta en absoluto. Las palabras son secundarias al conocimiento; lo que importa es su intención subyacente, su significado. Si las palabras cambian, pero no el significado, nada de gran importancia ha cambiado. No, aquí no nos limitamos a cambiar las palabras utilizadas, sino que proponemos una imagen mental.

Nuestra idea de clasificación es la de *entidades mentales llamadas clases, en las que se agrupan y contienen figurativamente* cosas distintas a las clases (o clases inferiores, en el caso de clases de clases) . Cuando decimos que las cosas pertenecen a la clase 'X', queremos decir que la clase 'X' es una especie de caja *en* la que estas cosas se almacenan, mediante la imaginación (en un momento dado, ya sea temporal o permanentemente). Es decir, nuestra 'definición' de clasificación es en realidad una convención formal utilizada para instituir esta imagen. Su verdadero significado es el siguiente:

> Si algo es X, entonces pertenece *a* la clase 'X' y está fuera de (es decir, *no pertenece a*) la clase 'nonX'.
>
> Si algo no es X, entonces está fuera de (es decir, no está en) la clase 'X', pero está en la clase 'nonX'.

Es evidente que «ser» algo y «estar en» (dentro, dentro de) otra cosa *no son lo mismo* . Nuestra definición convencional (es decir, por consenso) establece que si X se predica de algo, entonces podemos pensar que ese algo está contenido en la entidad mental llamada clase «X». Pero este principio no es absoluto; no es una proposición

que indique que estar sujeto a la predicación de X *implique natural y necesariamente* ser miembro de la clase «X». Porque la idea misma de clasificación, y por lo tanto esta definición de lo que constituye una clase y la pertenencia a ella, es una invención humana. Esta invención bien puede ser, y de hecho lo es, muy útil, pero sigue estando sujeta a las leyes de la naturaleza. Si descubrimos que la forma en que la hemos concebido, es decir nuestra definición, conduce inevitablemente a una contradicción, debemos ajustarla de tal manera que dicha contradicción ya no pueda surgir. Esta es nuestra forma de razonar y actuar en todas las situaciones similares.

Como mostraremos a continuación, dado que la contradicción es consecuencia de la implicación definitoria antes mencionada, debemos modificar dicha implicación. Es decir, debemos decretar que tenga límites. Por supuesto, no podemos simplemente decir vagamente que tiene límites; debemos definir con precisión estos límites para que se restablezca el valor práctico de nuestro concepto de clasificación. Podemos hacerlo al darnos cuenta de que nuestra definición de clasificación con referencia a algo que «está en» otra cosa significa que *la lógica de clases se concibe como relacionada con la lógica geométrica* . Esto es obvio, cuando reflexionamos sobre el hecho de que a menudo «representamos» las clases como figuras geométricas (en particular, círculos) y sus miembros como puntos dentro de esas figuras. Esta práctica no es accidental, sino que es la esencia misma de nuestra idea de clasificación. La clasificación es imaginar que colocamos ciertos elementos, identificados por poseer alguna propiedad común y distintiva, en un contenedor etiquetado [33].

[33] Se trata de una "representación" pictórica, una imagen analógica que no debe tomarse literalmente.

Examinemos ahora el concepto de auto-pertenencia a la luz de estas reflexiones. ¿Qué es la idea de auto-pertenencia? Es la presuposición de que una clase puede ser miembro de sí misma. Pero, ¿es realmente concebible esa noción? Si dejamos de lado por un momento la cuestión de la lógica de clases y reformulamos la pregunta en términos de lógica geométrica, vemos que es absurda. ¿Puede un contenedor contenerse a sí mismo? Por supuesto que no. No existe ningún ejemplo conocido de un contenedor que se contenga a sí mismo en el mundo físico; y, de hecho, ni siquiera podemos imaginar visualmente un contenedor que se contenga a sí mismo. Así pues, la idea de auto-contención no tiene base empírica, ni siquiera en la esfera mental. Es solo una conjunción fantasiosa de dos palabras, sin base experiencial. Por esta razón, la idea nos parece ilógica y podemos afirmar con seguridad como un "axioma" universal y eterno que la auto-contención es imposible. Un término sin sentido como "la colección de todas las colecciones" es necesariamente un término vacío; no estamos obligados a aceptarlo, de hecho, lógicamente no se nos permite hacerlo; Solo podemos hablar consistentemente de "la colección de todas *las demás* colecciones" [34].

Un contenedor, por supuesto, siempre es coextensivo consigo mismo, es decir, ocupa exactamente el espacio que ocupa. Pero tal «coextensión» no es contención, y mucho menos autocontención, pues en realidad (salvo en sentido figurado) no se refiere a dos cosas, sino solo a una; no hay ningún «co-» en ello, simplemente se extiende, una sola vez. Hablamos de contención cuando

---

[34] Para ilustrarlo con un ejemplo concreto: una bolsa de canicas (ya sea sola o, peor aún, con las canicas dentro) no se puede meter dentro de sí misma, aunque la bolsa en su conjunto, junto con todo su contenido, se pueda hacer rodar como una canica y, por lo tanto, se la pueda llamar canica.

un objeto más pequeño cabe dentro de uno más grande (o, en el límite, cuando *otro* objeto de igual tamaño cabe perfectamente dentro de un objeto determinado). El concepto de contención se refiere a dos objetos, no a uno. Debe haber dos objetos distintos; no basta con etiquetar el mismo objeto de dos maneras. Imaginar la «autocontención» es imaginar que un objeto entero puede de alguna manera caber dentro de sí mismo como un objeto más pequeño (o que de alguna manera puede convertirse en dos, con uno de los dos dentro del otro). Esto es *inconcebible* . Una cosa entera no puede ser una parte (ya sea una parte completa o parcial) de sí misma; nada puede ser a la vez entero y parte. Una sola cosa no puede ser dos cosas (ya sean del mismo o diferente tamaño) a la vez; nada puede existir simultáneamente como dos cosas.

No se puede decidir por convención que algo sea a la vez un todo y una parte, ni que una cosa sea dos. No se puede convocar algo naturalmente imposible. Solo se puede convocar algo naturalmente posible, aunque sea innecesario. Por lo tanto, el concepto de autocontención carece de sentido; es un concepto inevitablemente vacío, porque presupone que algo imposible es posible. No existe la autocontención; un contenedor jamás puede contenerse a sí mismo. Si esto es cierto, entonces, por supuesto, también lo es que ninguna clase se incluye a sí misma, ya que (como hemos visto) la clasificación es esencialmente una idea geométrica. Dado que un contenedor no puede contenerse a sí mismo, se deduce que la respuesta a la pregunta de si una clase puede ser miembro de sí misma es indudablemente y definitivamente: No. Porque decir de cualquier clase que es miembro de sí misma implica que un contenedor puede ser su contenido. Así como no existe ningún contenedor que sea su contenido, tampoco *existe ninguna clase que sea miembro de sí misma.*

Esta es una idea revolucionaria para la lógica de clases. Se aplica a todas y cada una de las clases, no solo a la clase «clases que no son miembros de sí mismas», que dio origen a la paradoja de Russell. Además, cabe destacar que aquí negamos la posibilidad de que una clase sea miembro de sí misma, pero no la posibilidad de que *no* lo sea. Cuando decimos que ningún contenedor se contiene a sí mismo, implicamos que es cierto que todo contenedor no se contiene a sí mismo. Del mismo modo, cuando decimos que ninguna clase es miembro de sí misma, implicamos que es cierto que toda clase no es miembro de sí misma. Esto significa que, si bien reconocemos el objeto de la paradoja de Russell, a saber, la clase «clases que no son miembros de sí mismas», rechazamos la noción de que tal clase pueda ser , ni *siquiera* hipotéticamente por un instante, miembro de sí misma (y, por lo tanto, paradójica), pues, afirmamos, ninguna clase es jamás miembro de sí misma.

¿Cómo es posible, se preguntarán ustedes? ¿Acaso no hemos demostrado ya con ejemplos que algunas clases son miembros de sí mismas? ¿No hemos acordado, por ejemplo, que la clase «clases», al ser una clase, tiene que ser miembro de la clase «clases», es decir, de sí misma? ¿Cómo podemos negar algo tan obvio? Seguramente objetarán además: si la clase «clases que no son miembros de sí mismas» no es miembro de sí misma, entonces innegablemente es miembro de sí misma; y si es miembro de sí misma, entonces innegablemente no es miembro de sí misma. Para responder a estas preguntas legítimas, volvamos a nuestra definición de clasificación y a lo que dijimos sobre ella. Como indiqué anteriormente, nuestra definición de clases y pertenencia a ellas tiene la forma de una implicación convencional. Dice:

> Si y solo si algo es X, entonces es miembro de la clase 'X'.

Ahora bien, dado que esta implicación convencional nos lleva inexorablemente a la paradoja, debemos revisarla, es decir, limitar su alcance, es decir, especificar las condiciones exactas en las que «funciona» y en las que deja de «funcionar». Lo esencialmente erróneo, como hemos visto, es que sugiere que una clase puede ser miembro de sí misma. Por ejemplo, dado que la clase «clases» es una clase, entonces es miembro de «clases»; en este ejemplo, la variable X tiene el valor clase y la variable «X» tiene el valor «clases». Pero, como hemos demostrado, la afirmación de que una clase puede ser miembro de sí misma implica lógicamente algo geométricamente imposible; a saber, que un contenedor puede ser su propio contenido. Por lo tanto, para evitar que surja la paradoja de Russell, necesitamos evitar las consecuencias indeseadas de nuestra definición. Dado que nuestro concepto de clasificación es problemático tal como está, ¿cuáles son las condiciones que debemos especificar para delimitarlo y así resolver el problema?

La respuesta a esta pregunta es que cuando el sujeto y el predicado de la cláusula antecedente son uno y el mismo, entonces la cláusula consecuente no debe estar implícita. Es decir, si la cláusula antecedente tiene la forma “si la clase 'X' es X”, entonces la cláusula consecuente “entonces la clase 'X' es un miembro de 'X' (y por lo tanto de sí misma)” *no se deduce* . Este “no se deduce” es una convención, al igual que el “se deduce” en general. Lo que hemos hecho aquí es simplemente trazar una línea, diciendo que el consecuente *generalmente* sigue al antecedente, *excepto en el caso especial* en que el sujeto y el predicado en el antecedente son “el mismo” (en el sentido de que el predicado X es aplicable a la clase “X”, que a su vez se basa en el predicado X). Esta es lógicamente una medida bastante aceptable, sin duda. Si se encuentra que una proposición general inducida tiene excepciones, entonces es legítimo e incluso obligatorio

hacerla menos general, retrocediendo lo justo para permitir dichas excepciones.

Dado que la definición inicial de clasificación era una convención general, es perfectamente admisible, al descubrir que esta convención nos lleva a una contradicción, acordar una convención ligeramente más restrictiva. Así, mientras que, en la gran mayoría de los casos, sigue siendo cierto que si algo es X, entonces pertenece a la clase 'X', y más específicamente, si una clase (por ejemplo, 'Y') es X, entonces (es decir, 'Y') pertenece a la clase de clases 'X', ***excepcionalmente, en el caso especial en que la clase que es X es la clase 'X' (es decir, donde 'Y' = 'X'), no podemos afirmar que pertenece a 'X'*** , pues esto implicaría que pertenece a sí misma, lo cual es imposible, ya que esto implica que un contenedor puede ser su propio contenido. Cabe destacar que no negamos que, por ejemplo, la clase 'clases' *sea* una clase; solo negamos *la implicación* que normalmente se da de que la clase 'clases' *pertenezca a* la clase 'clases'. Podemos seguir diciendo alegremente "es" (porque eso es mera predicación), pero aquí no se nos permite convertir ese "es" en "es miembro de" (porque eso constituiría una clasificación ilícita).

De esta forma, se evita que surja la paradoja de Russell. Es decir, con respecto a la clase «clases que no son miembros de sí mismas»: en primer lugar, es perfectamente legítimo suponer que la clase «clases que no son miembros de sí mismas» no es miembro de sí misma, puesto que sabemos con certeza (por lógica geométrica) que ninguna clase es miembro de sí misma; pero *no es* legítimo afirmar que este hecho (es decir, que no es miembro de sí misma) implique que sí lo es, puesto que tal implicación ha sido convencionalmente excluida. En segundo lugar, *no es* legítimo suponer, ni siquiera a efectos de argumentación, que la clase «clases que no son miembros de sí mismas» es miembro de sí misma, puesto

que ya sabemos (por lógica geométrica) que ninguna clase es miembro de sí misma, y por lo tanto no podemos establecer mediante tal suposición que no lo es, aunque de todos modos sea cierto que no lo es.

Como puede verse, nuestra corrección de la definición de clasificación, haciéndola menos general que la original, al especificar la situación concreta en la que no debe aplicarse la implicación, logra eliminar la paradoja de Russell. Podemos afirmar que la clase «clases que no son miembros de sí mismas» no es miembro de sí misma, pero no podemos afirmar que lo sea; por lo tanto, se bloquean ambos aspectos de la doble paradoja. En el primer aspecto, hemos bloqueado la inferencia de no ser «miembro de sí misma» a serlo; en el segundo aspecto, hemos prohibido la suposición de ser «miembro de sí misma», aunque la inferencia de no serlo sería inofensiva. En consecuencia, la respuesta a la pregunta planteada por Russell —a saber: «¿La clase de todas las clases que no son miembros de sí mismas es miembro de sí misma o no?»— es que esta clase no es miembro de sí misma, y que el hecho de que esta clase no sea miembro de sí misma no la convierte, contrariamente a las apariencias, en miembro de sí misma, porque ninguna clase es miembro de sí misma.

Así pues, si bien es cierto que la clase «clases» es una clase, no se deduce que sea miembro de sí misma; si bien es cierto que la clase «clases de clases» es una clase de clases, no se deduce que sea miembro de sí misma; si bien es cierto que la clase «clases positivas» es una clase positiva, no se deduce que sea miembro de sí misma; si bien es cierto que la clase «clases abiertas» es una clase abierta, no se deduce que sea miembro de sí misma; si bien es cierto que la clase «clases que no son miembros de sí mismas» es una clase que no es miembro de sí misma, no se deduce que sea miembro de sí misma. En cuanto a la clase «clases miembros de sí mismas», no

tiene ningún miembro. Cabe destacar que la restricción de clasificación que hemos introducido aquí tiene un alcance muy limitado; apenas afecta a la lógica de clases, ya que se refiere a unos pocos casos muy límite.

Lo anterior es, creo, la *resolución correcta y definitiva* de la paradoja de Russell. Reconocimos la existencia de un problema, la paradoja de Russell. Diagnosticamos la causa del problema, la suposición de que *la auto-pertenencia* es posible. Demostramos que la auto-pertenencia es inconcebible, ya que implica que un contenedor puede contenerse a sí mismo; esto no fue un ajuste arbitrario, nótese bien, sino una apelación a la razón. Propusimos una solución al problema, una que lo aborda con precisión y lo elimina quirúrgicamente. Nuestro remedio consistió en *desacoplar* la predicación de la clasificación en todos los casos donde se asume la auto-pertenencia, y solo en tales casos. Esta solución al problema es puro sentido común y no un vuelo de especulación; es simple y elegante; es convincente y sin complicaciones; no modifica esencialmente el concepto de pertenencia a una clase, sino que solo limita un poco su aplicación; introduce una restricción, pero una que está claramente circunscrita y bastante pequeña; Esto no genera daños colaterales en áreas de la lógica de clases, ni en la lógica en general, que no sean problemáticas, y por lo tanto no requiere adaptaciones adicionales de la doctrina lógica. Cabe destacar, además, que nuestra solución no recurre a ningún sistema oscuro de lógica simbólica moderna, sino que se desarrolla completamente utilizando el lenguaje común y conceptos y procesos ampliamente conocidos y aceptados.

3. Un poco de historia

Veamos ahora brevemente parte de la historia de la paradoja de Russell y cómo él y otros lógicos simbólicos modernos la abordaron [35].

Georg Cantor ya había encontrado en 1895 una antinomia en su propia teoría de conjuntos. En 1902, cuando Gottlob Frege (Alemania, 1848-1925) estaba a punto de publicar el segundo volumen de sus *Grundgesetze* , Russell le informó de dicha paradoja. Frege quedó totalmente convencido y no pudo ver cómo salir de las autocontradicciones inherentes a «la clase de clases que no se pertenecen a sí mismas». Lo consideró muy grave y afirmó: «Lo que está en cuestión es… si la aritmética puede tener siquiera un fundamento lógico». Frege intentó primero solucionar el problema sugiriendo que podría haber «conceptos sin clases correspondientes», o bien ajustando uno de sus «axiomas» de tal manera que:

> "Se puede decir que dos conceptos tienen la misma extensión si, y solo si, todo objeto que se incluyera en el primero *pero que no fuera a su vez una extensión del primero,* se incluyera igualmente en el segundo y viceversa" [36].

Es evidente que la sugerencia inicial de Frege de que podrían existir «conceptos sin clases correspondientes» puede interpretarse como una anticipación de mi disociación entre predicación y clasificación en casos específicos. Sin embargo, Frege no especificó en qué casos debe producirse dicha disociación. Esto se evidencia en su siguiente sugerencia, que, si bien apunta

---

35 Me refiero aquí principalmente al relato de William y Martha Kneale en *El desarrollo de la lógica* , cap. XI.1-2.

36 Kneale y Kneale, pág. 654. Las cursivas son suyas.

de manera tentadora a la dificultad de la noción de auto-pertenencia, no la rechaza de plano, sino que intenta atenuarla. Habla de dos conceptos en lugar de uno, y trata de excluir convencionalmente la extensión de cada uno del otro, aunque, por supuesto, continúa incluyendo los objetos que se encuentran dentro de la extensión; esto demuestra que no se ha percatado de que la auto-inclusión por extensión es ni siquiera concebible.

Cabe destacar que la paradoja de Russell se refiere a que una clase determinada (es decir, entre todas las clases que no pertenecen a sí mismas) sea o no sea miembro de *sí misma* , no de *otra* clase. Frege intenta resolver esta paradoja no con referencia a una sola clase, sino a un par de clases iguales, aunque (hasta donde sé) no ha demostrado que las clases coextensivas den lugar a una paradoja comparable a la de Russell. Por consiguiente, su intento de solución al problema no es pertinente. Además, Frege parece haber pensado que si todos los elementos que pertenecen a una clase (por ejemplo, 'Y') pertenecen a otra clase (por ejemplo, 'X'), entonces se puede suponer que *la clase* 'Y' pertenece a la clase 'X'; y viceversa en caso de coextensión. Esto se sugiere por su intento de evitar tal suposición, para así evitar (según su criterio) la paradoja de Russell resultante. Pero, en realidad, del hecho de que todos los Y sean X no se deduce que la clase 'Y' sea miembro de la clase 'X'; solo se deduce que la clase 'Y' es una subclase de la clase 'X', o una clase igual si la relación es reversible. Así pues, parece que Frege confundió las relaciones de pertenencia a una clase con la jerarquización de las clases, utilizando un término vago como «caída» para caracterizarlas ambas.

Cabe preguntarse si una clase igual, una subclase o incluso una superclase podría pertenecer a su clase jerárquicamente relacionada. A primera vista, parece posible. Por ejemplo, todas las clases positivas son clases y, por lo tanto, pertenecen a la clase «clases», y la clase

«clases positivas» es una subclase de la clase «clases». Sin embargo, aunque no todas las clases son positivas (algunas son negativas), la clase «clases» es una clase positiva (definida por una proposición positiva) y, por consiguiente, pertenece a la clase «clases positivas». Pero si bien este ejemplo sugiere que una superclase podría pertenecer a su subclase (y, por lo tanto, aún más, una clase igual o una subclase podría pertenecer a su clase jerárquicamente relacionada), podríamos expresar una duda mediante una *analogía* , como quizás pretendía Frege. Podríamos argumentar inductivamente, por generalización, que si una clase no puede ser miembro de sí misma, entonces quizás tampoco pueda ser miembro de ninguna clase coextensiva (como sugiere Frege), e incluso tal vez de una subclase o una superclase. La cuestión aquí radica en si las instancias a las que se refiere la primera clase pueden considerarse presentes dos veces en la segunda (como miembros de ella por derecho propio y como constituyentes de un miembro). Así pues, Frege podría haber planteado una cuestión válida, que podría conducir a mayores restricciones en la lógica de clases. Sin embargo, esto no nos concierne en el presente contexto, ya que (como ya se explicó) no es directamente relevante para la resolución de la paradoja de Russell.

Russell describió su paradoja en su libro *Principios de Matemáticas* , publicado poco después. Aunque inicialmente se inclinó por el segundo enfoque de Frege, más tarde prefirió el primero, proponiendo que podría haber "algunas funciones proposicionales que no determinaran clases genuinas". Nótese aquí nuevamente la incapacidad de identificar la fuente y la solución precisas del problema. Posteriormente, Russell pensó que "el problema nunca podría resolverse por completo hasta que todas las clases fueran eliminadas de la teoría lógica". Esto, en mi opinión, sería tirar al bebé con el agua del baño: una reacción exagerada. Pero luego descubrió (o

más bien, creyó descubrirlo, o se convenció a sí mismo de que lo había hecho) que la misma paradoja podía generarse sin "hablar de clases", es decir, con referencia a meros predicados, es decir, en términos de lógica de predicados en lugar de en términos de lógica de clases. Como lo expresan Kneale y Kneale [37]:

> En lugar de la clase que supuestamente contiene todas las clases que no son miembros de sí misma, consideremos la propiedad de ser una propiedad que no se ejemplifica a sí misma. Si esta propiedad se ejemplifica a sí misma, entonces no puede ejemplificarse a sí misma; y si no se ejemplifica a sí misma, entonces debe ejemplificarse a sí misma. Claramente, la naturaleza del problema es la misma que en la paradoja original, y sin embargo no se habla de clases.

Pero incluso si las clases no se mencionan explícitamente aquí, es evidente que se insinúan tácitamente. ¿Cómo se "ejemplificaría" una propiedad? Presumiblemente, la propiedad X sería "una propiedad que se ejemplifica" si resulta ser una de las cosas que poseen la propiedad X. Es decir, X ejemplifica X si pertenece a la clase de cosas que son X. No podemos hablar de propiedades sin recurrir a la predicación; y una vez que predicamos, podemos (dada la definición inicial de clasificación) clasificar sin duda. Por lo tanto, este intento no es más que un juego de palabras; se pretende lo mismo, pero se disfraza con otras palabras. Es deshonesto. Además, la forma en que supuestamente se evoca la paradoja aquí no da como resultado una paradoja.

---

[37]William y Martha Kneale, pág. 655.

No podemos decir, ni siquiera hipotéticamente, "si esta propiedad [es decir, la propiedad de ser una propiedad que no se ejemplifica a sí misma] se ejemplifica a sí misma", pues eso ya es contradictorio. Para reconstruir una paradoja de Russell en términos de "propiedad", tendríamos que hablar de " *la propiedad de todas las propiedades* que no se ejemplifican a sí mismas"; pues entonces tendríamos un nuevo término para analizar, como hicimos en la lógica de clases. Pero claramente, este nuevo término es bastante artificial y carece de sentido. Aquí de nuevo, debemos referirnos a " *la clase de todas las propiedades* que no se ejemplifican a sí mismas", y en ese caso, volvemos a la lógica de clases. Así pues, cabe señalar que, si bien Russell tenía razón al intentar determinar si su paradoja era un problema específico de la lógica de clases o uno que también ocurría en la lógica de predicados, y afirmó haber establecido que ocurría en ambos campos, en realidad (como acabamos de demostrar) no lo logró. En realidad, la paradoja *era* específica de la lógica de clases; y le habría convenido más admitir el hecho que intentar ignorarlo.

En respuesta a ciertas críticas de sus colegas, Russell finalmente «aceptó que las paradojas se debían a círculos viciosos y estableció como principio para evitarlos que "todo aquello que involucra a todos los elementos de una colección no debe ser uno de ellos"». Así, puede decirse que Russell admitió el principio que yo también he utilizado, a saber, que una colección no puede incluirse a sí misma como uno de los elementos recopilados, aunque, en realidad, la forma en que lo formuló sugiere que lo concibió como una convención diseñada para bloquear círculos viciosos incomprensibles más que como un absoluto lógico (nótese que dice «no debe» en lugar de «no puede»). Consideraba las paradojas de la teoría de conjuntos como «esencialmente del mismo tipo que la antigua paradoja de Epiménides (o del Mentiroso)». Esto

sugiere que, en esta etapa, veía su propia paradoja como debida, de alguna manera, a la autorreferencia. A primera vista, parece que la proposición «la clase de todas las clases que no son miembros de sí mismas es (o no es) miembro de sí misma» contiene algún tipo de autorreferencia, ya que la frase «miembro de sí misma» se repite (positiva o negativamente, en singular o plural) en el sujeto y el predicado [38]. Sin embargo, no se puede afirmar que la autorreferencia sea precisamente el problema.

Unos años más tarde, en un artículo publicado en 1908, Russell ofreció una explicación más elaborada de la paradoja de Russell basada en su «teoría de tipos». Russell argumentó entonces que «ninguna función puede tener entre sus valores algo que la presuponga, pues si lo tuviera, no podríamos considerar definidos los objetos denotados ambiguamente por la función hasta que esta lo fuera, mientras que, a la inversa… la función no puede ser definida hasta que sus valores lo sean» [39]. En otras palabras, la pregunta «la clase de todas las clases que no son miembros de sí mismas, ¿es o no es miembro de sí misma?» es intrínsecamente errónea, porque el tema permanece siempre fuera de nuestro alcance. No podemos comprenderlo hasta que resolvamos si es o no miembro de sí misma, y no podemos hacer esto último hasta que resolvamos lo primero; por lo tanto, el enigma es irresoluble, es decir, la pregunta es irrespondible. En efecto, el tema es un término cognitivamente imposible de formular, debido al dilema que implica su definición para cualquier pensador.

---

[38] Cabe señalar que si la autorreferencia fuera el meollo del problema, entonces la proposición "la clase de todas las clases que son miembros de sí mismas es (o no es) miembro de sí misma" sería igualmente problemática, aunque aparentemente no dé lugar a una paradoja similar.

[39] Citado por los Kneales, pág. 658.

Cabe señalar que el propósito de la explicación de Russell era, en efecto, invalidar la clase negativa «clases que no son miembros de sí mismas», puesto que esta es la clase que da origen a la doble paradoja que intentaba remediar. La clase positiva «clases que son miembros de sí mismas» claramente no produce una doble paradoja: si suponemos que *no es* miembro de sí misma, se produce una autocontradicción, pero aún podemos afirmar sin autocontradicción que sí lo *es* . De hecho, si la explicación de Russell fuera correcta, la clase positiva debería ser tan ilícita como la negativa. Porque si afirmamos la imposibilidad de que una clase se refiera a algo que *aún no está resuelto* , como hizo Russell con respecto a la clase negativa, entonces debemos admitir que esta característica también se encuentra en la clase positiva, y debemos rechazarla también. Russell no parece haberse dado cuenta de esto, es decir, de que su solución no diferenciaba técnicamente las dos clases y, por lo tanto, podía aplicarse a ambas. Por esta razón, su intento de resolver la paradoja de Russell mediante la circularidad o el infinito debe considerarse un fracaso. En mi teoría, en cambio, la clase positiva (la de la autopertenencia) es inválida (y vacía), puesto que es geométricamente impensable, mientras que la clase negativa (la de la no autopertenencia) sigue siendo perfectamente legítima (e incluso se manifiesta), como cabría esperar según el principio de que todas las afirmaciones (incluida la de autopertenencia) deben ser negables.

En resumen, Russell concluyó que una función no podía ser un valor en sí misma y propuso diferenciar función y valor como dos «tipos» que no podían mezclarse indiscriminadamente. Sin embargo, esta teoría, a mi parecer, es demasiado general y complica considerablemente las cosas. Como hemos visto, no podemos negar que, por ejemplo, «clases» *sea* una clase; lo máximo que podemos hacer es negar que esto implique

que «clases» *sea un miembro de* sí misma. Esto implica negar la autopertenencia, no la autopredicación ni la autorreferencia. En cuanto a la noción de «tipos» y, posteriormente, la de «órdenes dentro de los tipos», no deben confundirse con las ideas más tradicionales de jerarquías y órdenes de clases, que ya hemos expuesto anteriormente en este ensayo. En realidad, la semejanza entre los conceptos de Russell y estos últimos confiere a su teoría una apariencia de credibilidad; pero esta apariencia es ilusoria: se trata de conjuntos de conceptos muy diferentes. La noción de «tipos» de Russell es altamente especulativa y dista mucho del sentido común; si bien podría parecer que resuelve la paradoja de Russell, sus ramificaciones van mucho más allá e, incidentalmente, invalidan ideas tradicionales que no parecen problemáticas [40]. En resumen, es una medida improvisada y poco convincente.

Cada paradoja que encontramos es, por supuesto, una señal de que nos estamos desviando del camino correcto. En consecuencia, puede decirse que la paradoja de Russell fue una señal para Frege, Russell y otros lógicos modernos de que algo fallaba en su perspectiva. Se esforzaron por encontrar la raíz del problema, pero aparentemente no pudieron localizarla con precisión. Toda la complejidad de su enfoque simbólico y axiomático de la lógica no les sirvió de ayuda, sino que, por el contrario, les dificultó la solución. Esto demuestra que, antes de cualquier intento de simbolización y axiomatización, es esencial que los lógicos comprendan plenamente el tema en cuestión en términos de lenguaje ordinario y mediante el sentido común. Que yo sepa, la solución al problema que se propone en este ensayo es original, es decir, no se encuentra en ningún otro lugar. Si esto es cierto, entonces

---

[40] Para empezar, véase la crítica de Kneales a la "teoría de los tipos" en el capítulo XI.2.

la teoría de la lógica de clases desarrollada por los lógicos simbólicos modernos, que aún constituye el núcleo de lo que se enseña en las universidades hoy en día, necesita ser revisada y actualizada exhaustivamente.

4. Un poco de autocrítica

En cuanto a la solución de la paradoja de Russell que propuse hace más de dos décadas en mi *libro Lógica del futuro* , es necesario aclarar lo siguiente. Si bien mantengo, en general, mi teoría de la lógica de clases (en los capítulos 43-44), debo ahora distanciarme un tanto de mi intento de solución de la paradoja de Russell (en el capítulo 45).

Cabe reconocer que, en ese trabajo anterior, expresé un gran escepticismo respecto a la noción de autopertenencia; pero no logré descartarla por completo. Declaré: «Intuitivamente, al menos para mí, la sugerencia de que algo pueda ser a la vez contenedor y contenido es difícil de aceptar», e incluso postulé, a modo de generalización a partir de varios casos examinados, que «ninguna clase de nada, ni ninguna clase de clases de nada, es jamás miembro de sí misma», con la posible excepción de las «cosas» o las «clases de cosas» (aunque podría decirse de estas clases que no son miembros de ninguna clase, y mucho menos de sí mismas [41]); pero aun así, no rechacé la autopertenencia por principio, ni utilicé ese rechazo para explicar y resolver la paradoja de Russell, como hago en el presente ensayo.

Esto se evidencia, por ejemplo, en mi aceptación de la idea de que las "clases auto-miembros" son miembros de sí mismas. La razón por la que lo hice fue que "la

[41] Cabe señalar que, en este contexto, se me ocurre la idea de que la definición de pertenencia podría fallar ocasionalmente. Sin embargo, en aquel momento no profundicé en esa idea, ya que no analicé entonces las implicaciones formales de tal fallo.

auto-pertenencia, sea posible o no, no es la cuestión". Superficialmente, esto es cierto: la paradoja de Russell se refiere a la "clase de todas las clases que no son miembros de sí mismas", y no a la "clase de todas las clases que sí lo son". Pero, de hecho, como he demostrado hoy, esto no es cierto; la aceptación de la auto-pertenencia es la verdadera causa de la paradoja de Russell, y la no-pertenencia no es problemática en sí misma.

En cualquier caso, al no haber descartado debidamente la auto-pertenencia, recurrí a la única solución al problema que me pareció prometedora en aquel momento: el rechazo de la «permutación» de «es (o no es) miembro de sí mismo» a «es (o no es) {miembro de sí mismo}» (nótese la adición de llaves). Es decir, propuse la prohibición lógica de cambiar la *relación* de auto-pertenencia o no auto-pertenencia por un *término predecible* . Ahora veo que esto fue un error; fue una acción *tardía* en el proceso de pensamiento que condujo a la paradoja de Russell. Fue un intento superficial, que trataba un síntoma en lugar de la enfermedad. Lo hice, por supuesto, porque pensé que este era «de todos los procesos utilizados para desarrollar estos argumentos, el único de validez incierta (no establecida)». Pero, en realidad, no era la única causa posible del efecto; hubo un proceso *anterior* , de mayor significado, a saber, la transición de «es» a «es miembro de». En aquel momento no me percaté de este proceso anterior, y mucho menos de su vulnerabilidad; y por esa razón, no lo ataqué.

Evidentemente, iba por buen camino al buscar un punto en el proceso de pensamiento que impidiera el desarrollo de la paradoja de Russell. Sin embargo, mi error fue elegir ese punto demasiado tarde. De hecho, el punto correcto se encuentra antes, como se defiende en este ensayo. La paradoja de Russell no surge de una permutación ilícita, sino de la transformación ilícita de un predicado en una clase en casos donde se produciría una afirmación de

autopertenencia. Y si bien la solución propuesta es, en cierto modo, «convencional», el defecto que pretende corregir es muy real: la autopertenencia es, de hecho, imposible y, por lo tanto, nunca puede asumirse como verdadera. Mi solución anterior solo impedía la paradoja de Russell; no impedía la autopertenencia, que es la verdadera causa de la paradoja. Por consiguiente, la solución que propongo en este ensayo es más profunda y precisa.

9. Más información sobre la paradoja de Russell

De Temas (publicados solo en línea hasta el momento).

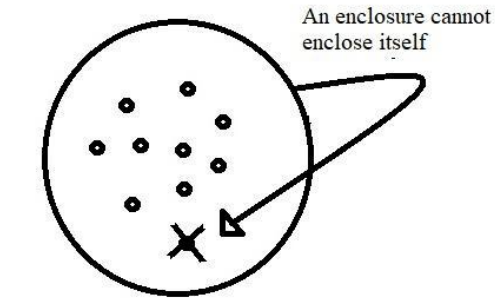
1. Mi resolución de la paradoja de Russell

Mi solución a la paradoja de Russell es, sencillamente, que *ninguna* clase es, ni puede ser jamás, miembro de sí misma. La autopertenencia es *inconcebible* . Por lo tanto, no sorprende que la adopción irreflexiva de la idea de autopertenencia por parte de Russell lo condujera a la paradoja. El concepto de pertenencia a una clase difiere significativamente del de predicación. La predicación es una forma natural de pensar; la pertenencia a una clase es artificial, inventada por los lógicos. El concepto de pertenencia a una clase es útil, pero tiene limitaciones. La idea de pertenencia a una clase es muy antigua, ya implícita en los discursos de Aristóteles; pero los lógicos la utilizaron de forma más consciente en la época moderna.

La idea original es que si cualquier cosa individual (llamémosla X) 'tiene' alguna propiedad (llamémosla cualidad Y), entonces está 'subsumida' bajo la especie o género de cosas que tienen Y; es decir, ***si X es Y, entonces X es un Y*** (nótese el uso de «a» para indicar que Y se refiere a una especie o género, a diferencia de una propiedad). Esta idea condujo posteriormente al concepto de «clase con miembros», definida por: ***si X es un Y, entonces X es un miembro de la clase «Y».*** (nótese el uso de comillas para referirse a una clase como distinta de un

predicado). Aquí se pretenden implícitamente los complementos: “si X no es Y, entonces X no es un Y” y “si X no es un Y, entonces X no es un miembro de la clase 'Y'”, nota.

Pero conviene aclarar que estas dos ideas, aunque relacionadas, son bastante distintas. Si bien se puede decir de cualquier cosa que sea Y que es *una* Y, no se puede decir de ninguna clase “Y” que resulte ser una Y que sea miembro de la clase “Y” (es decir, miembro de sí misma). Pues la idea de pertenencia a una clase es, figurativamente, la idea de un recinto (la clase) en el que se incluye un grupo de cosas (los miembros) que son similares de alguna manera (es decir, que tienen alguna propiedad en común, por ejemplo, Y); y es *imposible imaginar visualmente* que el recinto se incluya dentro de sí mismo (es decir, que la clase pueda ser miembro de sí misma). Se puede decir con palabras, por supuesto; pero no se puede dibujar en papel.



A class cannot be a member of itself

Así pues, la idea de subsunción bajo una especie o género (de «Y» bajo Y) y la de pertenencia a una clase (de «Y» en «Y») son solo parcialmente equivalentes y no deben confundirse. Por ejemplo, si bien la clase «clases» es, en

efecto, una clase, lógicamente no puede ser miembro de la clase «clases». Las caracterizaciones «es una clase» y «es miembro de la clase 'clases'» no son intercambiables. Sin embargo, la clase «clases» sí puede ser miembro de otras clases, en particular de la clase «clases no auto-miembros». De hecho, ninguna clase puede ser miembro de «clases auto-miembros», porque no existe tal clase (como se acaba de explicar); y, por consiguiente, *todas* las clases son miembros de «clases no auto-miembros».

Tenga esto en cuenta: aunque algunas clases (como la clase "clases") se *subsumen* a sí mismas, no se puede decir razonablemente que sean *miembros de* sí mismas. Esto significa que, cualquiera que sea el valor de Y: ***ya sea que la clase "Y" sea o no sea una Y, "Y"*** **no es** ***un miembro de la clase "Y",*** es decir, no es un miembro de sí misma. Por lo tanto, la proposición mediante la cual definimos inicialmente la pertenencia a una clase, a saber, si X es una Y, entonces X es un miembro de la clase "Y", tiene *una excepción formal* , a saber, el caso en que la clase "Y" es a la que se refiere el término X (que de otro modo es general).

(Aunque aquí no hemos tratado las subclases ni las clases superiores, cabe señalar que dicha excepción no impide que la clase «Y» sea miembro de la *clase superior* «clases Y», la cual, por definición, incluye a «Y» junto con todas sus subclases. Esto se debe a que «Y» y las «clases Y» se encuentran en niveles diferentes, por lo que esto no constituye una autopertenencia. Más adelante se abordará este tema).

Otra peculiaridad de la lógica de clases es que, así como las "clases" son una clase, las "no clases" pueden considerarse una clase sin contradicción. En la lógica de predicados, esto no es posible; es decir, una no clase (algo que no es una clase) no puede considerarse una clase: solo

una clase puede ser considerada una clase. Esto demuestra, una vez más, la divergencia entre estas dos formas de pensar. Por otro lado, cabe señalar que las "clases" no pueden considerarse consistentemente una no clase, y de igual modo, las "no clases" no pueden considerarse una no clase.

En cuanto a la paradoja de Russell, se alega que la clase «todas las clases que no son miembros de sí mismas» constituye una doble paradoja, porque: si decimos que no es miembro de sí misma, entonces sí lo *es* , y si decimos que *sí lo es,* entonces no lo es. Pero resulta evidente, a la luz de lo expuesto anteriormente, que el problema reside en el consecuente de la primera proposición condicional y en el antecedente de la segunda, es decir, en la *falsa* afirmación de que implican que «todas las clases que no son miembros de sí mismas» pueden ser miembros de sí mismas. En realidad, *ninguna* clase es miembro de sí misma, por lo que ambas proposiciones condicionales son erróneas (nonsequiturs).

En otras palabras, la resolución de la paradoja de Russell es que la clase "todas las clases que no son miembros de sí mismas", *al igual que cualquier otra clase* , *no es* miembro de sí misma, y esta afirmación *no* genera ninguna autocontradicción.

2. Por qué las resoluciones de Russell son inadecuadas

Ahora bien, Bertrand Russell formuló una explicación algo similar de su paradoja, diciendo:

> “Ninguna colección (en su totalidad) puede contener miembros que se definan en términos de sí misma; específicamente, ninguna colección

> existente puede ser jamás una parte constitutiva de sí misma" [42].

¿En qué se diferencia esta formulación de mi negación de la posibilidad lógica de la auto-pertenencia? La formulación de Russell obviamente se refiere indirectamente a la colección misma; es decir, en otras palabras: una colección no puede contenerse a sí misma como miembro porque dicha pertenencia solo podría activarse *definiendo la colección en términos de sí misma* , y eso no es factible. ¿Dónde está la "definición" aquí? Es la formulación misma del término, a saber, "todas las clases que son miembros de sí mismas" en nuestro caso.

La definición propuesta implica el término que se está definiendo; es decir, hay un elemento de autorreferencia, de circularidad. Esto puede llamarse autodefinición; y es obviamente inútil si no sabemos de antemano a qué se refiere por otros medios. Pero, ¿es la definición de la clase en sí misma el problema en cuestión? Yo diría que no: este término es simplemente el producto final de un proceso de razonamiento; no surge de la nada. Es este proceso de razonamiento el que debe examinarse y evaluarse.

¿La clase de "todas las clases que son miembros de sí mismas" *se produce por* autodefinición? No, el término predicativo subyacente "todas las clases que *son* clases" no es problemático: en efecto, todas las clases son clases. Esta predicación genera el término clasificatorio "todas las clases que *son miembros de* 'clases'", que tampoco es problemático en sí mismo. Lo problemático es que este último término, utilizando, por ejemplo, el hecho de que

---

42 Cita de *Principia Mathematica* de Nicholas Rescher en *Paradojas: sus raíces, alcance y resolución* , pág. 172. Este "principio del círculo vicioso" fue formulado anteriormente por Henri Poincaré.

"clases" es una clase y, por lo tanto, aparentemente un miembro de sí misma, da lugar inductivamente al término "clases que son miembros de *sí mismas* ". Así pues, la génesis de este último término problemático no se produce por mera autodefinición, como parece pensar Russell, sino por la traducción ilegítima (supuesta traducción) de una predicación perfectamente legítima en una proposición de pertenencia a una clase.

Dicha traducción es conocida y se ha demostrado su ilegitimidad por el hecho de que da lugar a la paradoja de Russell. Mi propuesta de solución se demuestra fácilmente con referencia al término negativo «todas las clases que *no son* miembros de sí mismas». Si la cuestión fuera la autodefinición, como sugiere Russell, entonces este término *también* sería problemático, puesto que también se refiere a sí mismo (aunque sea negativamente); ¡sin embargo, este término no es problemático! [43]Claramente, entonces, la cuestión en cuestión no es la autodefinición, sino algo completamente distinto. Se trata de la inferencia ingenua que acabo de describir.

Más precisa es la afirmación final de Russell, a saber, que «ninguna colección existente puede ser jamás una parte constitutiva de sí misma»; pero esto no es una explicación en sí misma: es una consecuencia evidente de la paradoja de Russell, es decir, una afirmación de que obviamente

---

[43] Alternativamente, si se considera problemático lo segundo, entonces no se puede considerar problemático lo primero. En cualquier caso, uno de los dos debe ser aceptable para que el diálogo sea posible. Menciono esta alternativa solo porque, *a primera vista,* podría parecer que el término negativo «todas las clases no son miembros de sí mismas» es el problemático, ya que aparentemente produce una paradoja (mientras que el positivo aparentemente no). En realidad, este término negativo no es problemático, y es el término positivo «todas las clases son miembros de sí mismas» el que causa la paradoja. Tenga esto en cuenta.

existe algún tipo de problema con que cualquier «colección existente» sea «una parte constitutiva de sí misma», lo que implica que tal colección no puede existir. Esto no es una explicación, sino una observación que necesita una explicación, y la explicación que ofrece Russell gira en torno a la autodefinición, como hemos visto. Pero la autodefinición no está, de hecho, implicada genéticamente.

Diría que, si bien Russell vislumbró un problema inherente a la idea de autopertenencia (o, alternativamente, quizás, de no autopertenencia) por parte de una clase, no comprendió con precisión la razón de dicho problema. Lo identificó erróneamente como autodefinición, es decir, como autorreferencia en la definición. Pero aunque esta característica está presente y es significativa, no es la clave del problema.

En realidad, el problema solo puede resolverse mediante el análisis conceptual más detallado que he formulado. El problema surge de la suposición de que la subsunción bajo un concepto y la pertenencia a una clase son ideas totalmente paralelas (efectivamente intercambiables); simplemente no lo son. Es esta suposición errónea previa la que, en última instancia, da origen a la paradoja de Russell. Una vez comprendido esto, resulta fácil ver por qué la «autopertenencia» es inconcebible y debemos afirmar categóricamente que ninguna clase es miembro de sí misma, y en ese caso, la paradoja de Russell deja de existir.

Más adelante, Russell, aparentemente insatisfecho con la solución que había propuesto anteriormente para su doble paradoja, propuso su Teoría de los Tipos como medio para lograr este fin. En su forma más simple, esta teoría se centra en las diferencias entre los elementos individuales que pertenecen a las clases, las clases mismas, las clases de clases que a su vez pertenecen a estas últimas, y así

sucesivamente. Estos son "tipos" sucesivos, que forman una jerarquía de elementos desde los objetos concretos hasta derivados cada vez más abstractos de los mismos.[44]

En resumen, la teoría postula que una proposición puede ser universalmente verdadera para un tipo inferior, pero no incluir un tipo superior. Por ejemplo, se puede decir que todas las clases pertenecen a la clase «clases», sin que esto implique que «clases» también pertenezca a ella, simplemente porque (según la teoría) es de un tipo superior a todas las clases que incluye. Así pues, «todas las clases» no significa literalmente *todas* las clases; puede (debido a la diferencia de «tipo») excluir la clase «clases» misma. Por supuesto, esta afirmación es bastante arbitraria y solo oculta la paradoja de Russell tras una cortina de humo.[45]

Esta idea tiene algo de cierto, pero también genera cierta confusión. En efecto, existe una jerarquía entre elementos individuales concretos, agrupaciones abstractas de elementos (clases) y grupos más abstractos de grupos de elementos (clases de clases). Sin embargo, la paradoja de Russell no surge cuando intentamos incluir una clase en una clase de clases, es decir, una subclase en una clase superior. Y, por supuesto, la paradoja de Russell no surge cuando intentamos incluir un elemento individual en una clase. La paradoja de Russell solo surge cuando intentamos incluir una clase en sí misma; o (por extensión) una clase de clases en sí misma.

La paradoja de Russell surge *únicamente* al intentar incluir una clase, o una clase de clases, *en sí misma* . Por

---

[44] No me extenderé aquí en un análisis detallado de esta teoría. Remito a los lectores a mis análisis anteriores, más exhaustivos, que se encuentran en mis libros *Lógica del futuro* (capítulos 43-45) y Lógica a fortiori (Apéndice 7.5).

[45] Remito a los lectores a la excelente exposición y crítica de la paradoja de Russell y sus intentos de resolución, en el capítulo 11 de * *The Development of Logic** , de William y Martha Kneales.

lo tanto, la jerarquía no es el problema aquí, ¡y la teoría de tipos no puede ofrecer una solución! Es solo incidental que una clase solo puede agrupar objetos concretos, o que una clase de clases solo puede agrupar clases menos abstractas. La razón de la paradoja no reside en las diferencias de tipos, sino en el hecho de que una clase o una clase de clases no puede, ni siquiera en la imaginación, incluirse a sí misma.

Cabe señalar que, en nuestro ejemplo anterior, la clase «clases» (es decir, la clase de todas las clases) es en realidad una clase de clases, puesto que las unidades que agrupa son todas clases. Sin embargo, sigue siendo una clase por derecho propio y, como tal, aparentemente se subsume a sí misma. La solución no reside en que no pueda subsumerse a sí misma, sino simplemente en que no puede ser miembro de sí misma. Esto no es una cuestión de jerarquía (tipo), sino de transformar la predicación en clasificación.

3. Por qué la resolución de Rescher es inadecuada

La solución que propone Nicholas Rescher para la paradoja de Russell difiere en cierta medida, aunque no demasiado, del primer intento de Russell. Aplica al caso en cuestión un principio general que él mismo ha formulado, denominado Principio de Introducción Exitosa. Señala que «solo cuando algo se identifica correctamente, puede servir como objeto de un debate significativo»; y añade: «Con un pseudo-objeto identificado de forma inapropiada… se abre de par en par la puerta a la contradicción y la paradoja» [46]. En su

[46] Op. cit., págs. 172-73. Si bien este principio SIP es intuitivamente sólido y relevante para el caso en cuestión, cabe señalar que no es una verdad general. En realidad, gran parte del discurso humano se construye mediante términos más o menos vagos.

opinión, la paradoja de Russell se debe, pues, a un fallo en la identificación adecuada inherente a una clase definida en términos de sí misma.

Así pues, Rescher se centra en el problema de la autodefinición, al igual que Russell. Su enfoque difiere, como él mismo señala, en que mientras Russell considera la autodefinición una cuestión ontológica, Rescher la ve como una cuestión comunicativa. Para Russell, no puede existir una clase que se defina en referencia a sí misma; mientras que para Rescher, ni siquiera podemos empezar a usar o discutir tal clase, porque es como si aún no se hubiera dicho nada. Ambas ideas son, en mi opinión, ciertas. Pero no resuelven la paradoja de Russell, como ya se ha explicado. Repito, la cuestión no es de autodefinición, sino de pasar de un discurso predicativo a uno clasificatorio. Una clase como «clases» *es ,* en efecto, una clase, pero «clases» no puede ser *miembro de* la clase «clases». Se trata de dos sistemas de pensamiento diferentes, que no se comportan de forma idéntica en todas las circunstancias.

No podemos crear conceptos a voluntad, simplemente juntando palabras en una definición. Una definición es

---

Relativamente pocos términos, si acaso alguno, están definidos de forma completa y definitiva. Esto se debe a que nuestro conocimiento es esencialmente inductivo, en lugar de deductivo (como muchos lógicos imaginan). Usamos una palabra (cuando la necesitamos) y una definición operativa (cuando la tenemos) como herramientas pragmáticas. Nuestra comprensión de su referencia es siempre provisional y temporal. A medida que nuestro conocimiento crece y nuestros análisis se refinan, nuestros términos (con suerte) ganan mayor precisión y se vuelven más concisos. Por lo tanto, el principio propuesto por Rescher no debe tomarse de forma demasiado general. El debate significativo se produce constantemente con palabras imperfectamente definidas; esa es la norma, no la excepción, y, a pesar de ello, solemos evitar muy bien las contradicciones. El significado no es esencialmente una cuestión verbal, sino de intuición y comprensión.

una hipótesis sujeta a verificación lógica. En el caso de la definición de pertenencia a una clase, es decir, la proposición de que “si X es Y, entonces X pertenece a la clase 'Y'”, se produce una doble paradoja en el caso particular de “si 'Y' es Y, entonces 'Y' pertenece a 'Y'”. Por lo tanto, la regla que intentamos formular (es decir, la mencionada definición de pertenencia a una clase) debe tener una excepción: cuando X representa la clase “Y”.

La paradoja de Russell es valiosa porque nos enseña que debe contemplarse una excepción en la fórmula definitoria. Inicialmente, asumimos que la definición general es válida; pero al descubrir que conduce a una autocontradicción y una paradoja, nos vemos obligados a retractarla parcialmente. Lógicamente, nos vemos forzados a rechazar por completo la noción de autopertenencia. Esta retractación es muy limitada y precisa, pero restablece la coherencia en la lógica de clases. Nuestra definición hipotética se ha ajustado para evitar la inconsistencia. Esta es la resolución correcta de la paradoja de Russell.

4. Otras encarnaciones de la paradoja

Cabe mencionar que Russell se dio cuenta de que el concepto de clase estaba en juego de alguna manera, pero luego intentó demostrar que la misma paradoja podía formularse dentro de la predicación. Como lo expresaron Kneale y Kneale [47]:

> En lugar de la clase que supuestamente contiene todas las clases que no son miembros de sí misma, consideremos la propiedad de ser una propiedad que no se ejemplifica a sí misma. Si esta propiedad se ejemplifica a sí misma, entonces no

47 *El desarrollo de la lógica* , pág. 655.

> puede ejemplificarse a sí misma; y si no se ejemplifica a sí misma, entonces debe ejemplificarse a sí misma. Claramente, la naturaleza del problema es la misma que en la paradoja original, y sin embargo no se habla de clases.

Las implicaciones aquí son que algunas propiedades se "ejemplifican" a sí mismas, mientras que otras no. Es decir, una propiedad Y: en algunos casos, Y es Y (es decir, tiene la propiedad Y) y en otros, Y no es Y (es decir, carece de la propiedad Y). Si Y es Y, entonces Y se ejemplifica a sí misma; y si Y no es Y, entonces Y no se ejemplifica a sí misma. Sin embargo, ¿es cierto alguna vez que Y es Y? ¿Se puede señalar algo que tenga precisamente a sí mismo como predicado? Yo diría que no. Claramente, "Y es Y" aquí no se refiere a una tautología predicativa, como "blanco es blanco". Más bien, lo que queremos decir con "Y es Y" aquí es que "blancura es blanco", es decir, la Y- *idad* (*la propiedad* Y) es Y (tiene la propiedad Y). Esta es una afirmación muy diferente, que no se deduce de la anterior.

Y la misma crítica que se le puede aplicar a la versión clasificatoria se puede utilizar. Bien podríamos decir que si X es Y, entonces X "ejemplifica" la cualidad de Y (es decir, es una de las cosas que posee dicha cualidad), pero en el caso especial en que X se reemplaza por la cualidad de Y (es decir, la propiedad Y), debemos hacer una excepción y negar que "la cualidad de Y es Y" implique "la cualidad de Y ejemplifica la cualidad de Y (es decir, ella misma)". Para evitar una eventual paradoja de Russell, debemos al menos evitar dicha implicación, si no rechazar por completo el consecuente (afirmando que nunca es cierto que la cualidad de Y ejemplifique la cualidad de Y), si no rechazar siquiera el antecedente

(afirmando que nunca es cierto que la cualidad de Y sea Y).

Russell afirmó que si «la propiedad de no autoejemplificación» se ejemplifica a sí misma, entonces carece de «la propiedad de no autoejemplificación» y, por lo tanto, no puede ejemplificarse; y si no se ejemplifica a sí misma, entonces posee «la propiedad de no autoejemplificación» y, por lo tanto, debe ejemplificarse. En mi opinión, esto significa que debemos rechazar la idea misma de autoejemplificación y afirmar que *nada* se ejemplifica a sí mismo (incluso si encontramos alguna cualidad Y que sea Y). En ese caso, podemos afirmar sin contradicción alguna que «la propiedad de no autoejemplificación» no se ejemplifica a sí misma.

Lo mismo puede expresarse con otros términos: algunos conceptos son «instancias» de sí mismos, mientras que otros no. Es decir, podemos definir la instanciación diciendo: si X es un Y, entonces X es una instancia de Y (es decir, X es un individuo dentro del grupo llamado Y); pero debemos hacer una excepción en el caso especial en que X se reemplaza por Y, y negar que si Y es un Y, entonces Y es una instancia de Y. Es evidente que Y aquí se refiere específicamente a un concepto, y no a cualquier cosa. Un ejemplo sería: «el concepto "concepto" es un concepto; por lo tanto, el concepto "concepto" es una instancia de sí mismo».

Esto es claramente muy similar a decir: «la clase de todas las clases es en sí misma una clase; por lo tanto, la clase de todas las clases es un miembro de sí misma»; y de ello surge inevitablemente una paradoja similar, que puede neutralizarse de manera parecida. Es decir, tras reflexionar, debemos afirmar que *nada* es una instancia de sí mismo, de modo que, de hecho, «el concepto de no-autoinstanciación » no es una instancia de sí mismo.

Se puede seguir un razonamiento similar con respecto a la «inclusión» y cualquier otra relación similar.

Claramente, ninguna de estas cuestiones se relaciona con la predicación en sí, sino con ideas más geométricas (clasificar, tener/pertenecer, instanciar, incluir) mediante las cuales intentamos representarla. Estas relaciones se conciben como metáforas de la predicación; pero resulta que estas analogías no son perfectas. La lección que nos enseña la paradoja de Russell es que todas estas relaciones derivadas tienen excepciones (como ya se detalló anteriormente). La paradoja de Russell no afecta a la predicación en sí; la predicación es inmune a ella. Si bien Russell tenía razón al investigar si su paradoja afectaba a la predicación, se equivocó al concluir que sí; no la afecta.

5. Acerca de la paradoja del barbero

Hay un pueblo con un solo barbero. Algunos hombres se afeitan solos; el resto no, pero todos son afeitados por el barbero. ¿Y el barbero? ¿Se afeita él mismo o no? Si lo hace, es uno de los que se afeitan solos (sin problema); si no lo hace, es uno de los que son afeitados por el barbero, es decir, por sí mismo (aparente contradicción). Obviamente, en su caso, las dos clases de hombres se superponen: él se afeita solo y es afeitado por el barbero. Esto no es problemático: simplemente nos indica que la supuesta dicotomía entre los dos grupos de hombres no es realmente una sola; su caso es una excepción lógicamente implícita. Es decir, en su caso, a diferencia de otros, ser afeitado por el barbero no implica no afeitarse él mismo.

¿Cómo se compara esto con la paradoja de Russell, de la cual la paradoja del barbero se presenta como una ilustración? Resuelvo la paradoja de Russell diciendo que ninguna clase es miembro de por sí misma; de ello se deduce que podemos afirmar que la clase de «clases que

no pertenecen a sí mismas» no necesariamente pertenece a sí misma, sin que esto genere una contradicción. Resuelvo la paradoja del barbero de forma diferente, al afirmar que, a diferencia de todos los demás hombres del pueblo, el barbero puede afeitarse a sí mismo y ser afeitado por el barbero (él mismo). En la paradoja de Russell, una de las clases en conflicto se elimina por completo. En la paradoja del barbero, el conflicto entre las clases se atenúa en un único caso (el del propio barbero). Dado que ambas paradojas no se resuelven de la misma manera, deben considerarse lógicamente distintas.

10. La paradoja de la confirmación de Hempel

De Reflexiones lógicas y espirituales 1.8.

1. Análisis tradicional

Carl Gustav Hempel [48] expuso una supuesta "paradoja de la confirmación", que sugería que una lógica inductiva formal totalmente consistente es imposible. Esto se conoce comúnmente como "la paradoja del cuervo" y puede describirse de la siguiente manera:

a) La observación de que algunos cuervos son negros (algunos A son B) confirma la hipótesis de que todos los cuervos son negros (todos A son B).

Esta última proposición puede oponerse a Todas las cosas que no son negras no son cuervos (Todo lo que no es B no es A).

b) A continuación, consideremos la observación de que Algunas manzanas son verdes (Algunas C son D). Esto se puede convertir en Algunas cosas verdes son manzanas (Algunas D son C).

De esta proposición se deduce que algunas cosas que no son negras no son cuervos (algunas cosas que no son B no son A), puesto que las cosas verdes no son negras y las manzanas no son cuervos.

Ahora bien, así como Algunos cuervos son negros (Algunos A son B) confirma la hipótesis de que Todos los cuervos son negros (Todos los A son B), así también

48 Alemania-EE. UU., 1905-1997.

Algunas cosas que no son negras no son cuervos (Algunos que no son B no son A) confirma la hipótesis de que Todas las cosas que no son negras no son cuervos (Todos los que no son B no son A).

Esta proposición inducida puede, a su vez, oponerse a «Todos los cuervos son negros» (o «Todos los A son B»), y aquí radica la dificultad, pues parece que la mera observación de algunas manzanas verdes basta para confirmar la hipótesis de que «¡Todos los cuervos son negros!». Cabe destacar que para llegar a este resultado ni siquiera fue necesario observar ningún cuervo negro.

c) De lo anterior se deduce que podemos igualmente bien, utilizando el mismo proceso lógico, dado que algunas manzanas son verdes, confirmar la hipótesis de que todos los cuervos son rosas, o de cualquier otro color (excepto el verde).

d) Esto es en sí mismo un misterio: ¿cómo pueden las manzanas decirnos algo sobre los cuervos? Intuitivamente, esto debe considerarse una *falacia lógica* . Además, en el caso de los cuervos negros, su existencia cuenta con respaldo empírico, como ya se indicó; por lo tanto, la inferencia de que todos los cuervos son negros a partir de las manzanas verdes aún parece razonable. Sin embargo, en el caso de los cuervos rosados, nunca hemos observado ninguno; por lo tanto, la inferencia de que todos los cuervos son rosados a partir de las manzanas verdes parece completamente injustificable.

Además, si sabemos por observación que algunos cuervos son negros, ¿cómo podemos concluir que todos los cuervos son rosas? Incluso si no afirmamos que todos los cuervos son negros, sino solo que todos son rosas, en tales circunstancias estaríamos sosteniendo proposiciones contradictorias: la particular de que algunos cuervos son negros y la general de que todos los cuervos son rosas.

Además, incluso si nunca hemos observado el color de ningún cuervo, según el proceso inductivo anterior

podemos llegar a conclusiones contradictorias como: Todos los cuervos son negros, Todos los cuervos son rosas, Todos los cuervos son naranjas, etc. Este resultado también contradice el principio de no contradicción.
Además, lo mismo puede aplicarse no solo a las manzanas verdes, sino también a las de otros colores (excepto negras o rosas, etc., según el caso), e incluso a otras cosas (que no sean cuervos) que no sean manzanas. En ese caso, casi todo es posible en el ámbito del conocimiento, ¡y *podemos afirmar o negar a nuestro antojo casi cualquier cosa sobre casi todo* !
Según las presentaciones tradicionales [49], esta es la paradoja de Hempel. De dicho análisis se desprende que los procesos inductivos de confirmación de hipótesis (como las generalizaciones derivadas directamente de la experiencia o indirectamente de sus derivados lógicos) presentan fallos fundamentales. El análisis implicado es bastante *formal* , es decir, puede realizarse en términos de símbolos como A, B, C, D, y por lo tanto, tiene validez universal.
De ello se deduce que la inducción inevitablemente dará lugar a diversas absurdidades: aparentes falacias, numerosas contradicciones y, en última instancia, implicará la arbitrariedad de todo conocimiento humano. Claramente, Hempel descubrió aquí un serio desafío para la lógica inductiva y la lógica en general.

2. Análisis novedoso

Como mostraré a continuación, *el análisis anterior es inexacto en algunos aspectos importantes* . Demostraré que, si bien Hempel descubrió un problema formal interesante para que los lógicos lo consideraran y resolvieran, este problema no da lugar a lo que

[49] Véase, por ejemplo, el artículo de Wikipedia en http://en.wikipedia.org/wiki/Raven_paradox.

llamaríamos una paradoja. Es decir, se pueden extraer valiosas lecciones de la paradoja de Hempel (como podemos seguir denominándola convencionalmente), pero no plantea a la lógica ningún dilema insuperable.

a) La primera operación descrita anteriormente es el proceso inductivo de *generalización , de uso común* . Una proposición particular (Algunos A son B) se transforma en una general (Todos los A son B). Lo particular respalda lo general del mismo modo que la evidencia positiva confirma una hipótesis. Su relación lógica es aductiva. Aquí, «algunos» significa «al menos algunos, posiblemente todos y posiblemente solo algunos», y al generalizar, optamos por la hipótesis «todos» en lugar de la hipótesis «solo algunos».

Sin embargo, sería un error considerar que el hecho de que algunos A sean B por *sí solo* justifica inductivamente que todos los A son B. Tal generalización es una inferencia inductivamente admisible *siempre que hayamos buscado y no hayamos encontrado hasta ahora ningún* A que *no sea* B. Porque si *hubiéramos* encontrado (por observación directa o por algún razonamiento) que algunos A no son B, ciertamente no habríamos generalizado. Además, si *posteriormente* encontramos un A que no es B, tendríamos que *particularizar* todos los A son B y volver a algunos A son B.

Esta condición *indispensable* para la generalización, a saber, estar atento a los casos contradictorios y ajustar el juicio en consecuencia, no se destaca ni se menciona en la presentación anterior, nótese bien. Sin embargo, esta es una regla conocida y aceptada del pensamiento científico al menos desde la época de Francis Bacon, quien enfatizó la importancia del «caso negativo» en la inducción. Ignorar esta condición inevitablemente conduce a contradicciones tarde o temprano.

Respecto a la contraposición de «Todos los A son B» a «Todos los que no son B no son A», se trata, por

supuesto, de un acto deductivo. Aun así, debemos tener presente que la conclusión «Todos los que no son B no son A» se debe únicamente a la inferencia inductiva previa de «Todos los A son B» a partir de «Algunos A son B». No se requiere ninguna observación para la deducción, pero seguimos obligados a verificar el acto inductivo anterior, es decir, a estar atentos a posibles casos de A que no sean B.

b) Ahora bien, admitamos que algunos C son D, como se indicó anteriormente. Algunos C son D se convierte fácilmente en Algunos D son C. Sin embargo, Algunos D son C no implica formalmente que algunos no B no sean A; aquí interviene tácitamente alguna inferencia silogística, que debería quedar clara. Evidentemente, damos por sentado que *el verde no es negro* y *las manzanas no son cuervos* , de donde se construyen los dos silogismos sucesivos siguientes:

Primera figura, EIO:

Ninguna cosa verde es una cosa negra (Ninguna D es B)

Algunas manzanas son verdes (Algunas C son D)

Por lo tanto, algunas manzanas no son negras (algunas C no son B).

Tercera figura, AII:

Todas las manzanas no son cuervos (Todas las C no son A)

Algunas manzanas no son negras (Algunas C no son B)

Por lo tanto, algunas cosas que no son negras no son cuervos (algunas cosas que no son B no son A).

De ahí que, por generalización, obtenemos: Todas las cosas que no son negras no son cuervos (Todo lo que no es B no es A); y luego, por contraposición: Todos los cuervos son negros (Todo lo que es A es B). Cabe señalar que las premisas que llevaron a esta conclusión general no incluyen «Algunos cuervos son negros»; es decir, esta conclusión no se basa en ninguna observación empírica de cuervos negros.

Cabe señalar también que podríamos haber obtenido el mismo resultado con las premisas «Ningún cuervo es verde (ningún A es D)» y «Ninguna manzana es negra (ningún C es B)». Asimismo, cabe destacar que, si bien los silogismos implicados son procesos deductivos, todas estas premisas tácitas requieren observaciones y generalizaciones previas (es decir, inducciones) para su adopción.

Además, es importante señalar que estos silogismos no podrían construirse si los colores de los cuervos y las manzanas en cuestión fueran iguales (ambos verdes o ambos negros), o si los cuervos y las manzanas no fueran clases mutuamente excluyentes. También asumimos que un cuervo no puede tener más de un color (por ejemplo, ser parcialmente negro y parcialmente verde, o a veces uno y a veces el otro); y lo mismo ocurre con una manzana.

El siguiente paso fue generalizar «Algunos no B son no A» a «Todos los no B son no A». Pero, de nuevo, la generalización solo está permitida si no tenemos evidencia o inferencia de ninguna otra fuente que indique que algunos no B no son no A. Es decir, en nuestro ejemplo, debemos tener presente que es posible que algunas cosas que no son negras no sean cuervos, es decir, que sean cuervos, lo que significa que aún podríamos encontrar algunos cuervos no negros (albinos).

Aquí también debemos asegurarnos, de acuerdo con el principio fundamental de aducción de Bacon, de que no exista ninguna observación contradictoria que obstaculice nuestro *ímpetu expansivo* . Esto es aún más necesario, puesto que aquí la premisa de generalización Algunos no B no son A se obtuvo indirectamente por deducción a partir de productos de inducción previos, mientras que anteriormente nuestra premisa Algunos A son B se observó (supuestamente) directamente.

Cabe señalar además que estas dos generalizaciones tienen la condición de que algunos A no son B o algunos no B no son no A, respectivamente, no sean verdaderas; y estas condiciones son idénticas, ya que ambas proposiciones son lógicamente equivalentes por contraposición. Esto significa que, en ambos casos, ya sea que razonemos directamente a partir de cuervos negros o indirectamente a partir de manzanas verdes, existe la misma condición implícita para la generalización: que, según nuestra experiencia o razonamiento hasta la fecha, no han aparecido cuervos que no sean negros.

Por lo tanto, cualquiera que sea la generalización que elijamos, la condición de que no existan instancias conocidas de A que no sean B permanece inalterada, y la dependencia de la verdad de que todos los A son B respecto de esta condición no cambia. Nótese también que la misma condición se mantiene antes y después de tales generalizaciones. Es decir, incluso después de este proceso inductivo, si descubrimos nuevas pruebas en contrario, lógicamente podemos y debemos *retractarnos* . Como se indicó anteriormente en c) y d): utilizando el mismo proceso lógico, dado que algunas manzanas son verdes, podemos confirmar igualmente la hipótesis de que todos los cuervos son rosas, y muchas otras hipótesis descabelladas que entran en conflicto entre sí [50]. Obviamente, algo estamos haciendo mal y debemos tomar medidas para prevenir tales consecuencias absurdas o corregirlas si se producen. A continuación, explicaré la solución al problema.

---

[50] Para demostrar que las proposiciones con predicados diferentes son contradictorias, utilizamos el silogismo. Por ejemplo, «Todos los cuervos son negros» y «Todos los cuervos son rosas» son incompatibles, porque sabiendo que «Ninguna cosa negra es rosa», obtenemos, mediante silogismo (primera [figura], EAE): «Ningún cuervo es rosa», lo cual es contrario a «Todos los cuervos son rosas».

La generalización nunca es un proceso irreversible. Por lo tanto, si alguna generalización conduce a contradicciones, tenemos la libertad, e incluso la obligación, de particularizar. La pregunta, por supuesto, sigue siendo: ¿ *en qué dirección precisa y hasta dónde debemos remontarnos?* Aun así, esto significa que no existe una «paradoja» en la lógica inductiva como en la deductiva; casi todo (con la excepción de la lógica misma, especialmente las leyes del pensamiento sobre las que se fundamenta) es y seguirá siendo «negociable».

En la deducción, una contradicción es un hecho mucho más grave, ya que se presume que el proceso que la origina es necesario. En cambio, *en la inducción, sabemos desde el principio que la relación entre la(s) premisa(s) y la conclusión es condicional; por lo tanto, se espera que surjan contradicciones, y es precisamente la función de la lógica inductiva determinar cómo responder a ellas.*

Abordar las contradicciones es una rama de la lógica inductiva, denominada ***armonización*** o resolución de conflictos. Esto no es algo raro ni excepcional, sino que ocurre constantemente en el desarrollo del conocimiento. *A veces, los conflictos se resuelven antes de que se manifiesten; otras veces, después. Si los prevemos, los evitamos; de lo contrario, realizamos las rectificaciones posibles y necesarias.*

La particularización de una proposición general se denomina retractación. En un sentido más amplio, la retractación implica el rechazo o la modificación de una teoría a la luz de nuevas evidencias. Por ejemplo: hasta ahora, solo he visto cuervos negros y he asumido que todos son negros; mañana, podría observar un cuervo blanco y cambiar mi opinión sobre los posibles colores de los cuervos.

Hempel, evidentemente o al menos aparentemente, desconoce este aspecto crucial del razonamiento inductivo, pues de lo contrario no habría considerado

paradójicas las contradicciones que surgen durante la inducción. Sin embargo, la situación que describe resulta interesante en este contexto, por razones que, creo, no llegó a comprender.

Después de la primera generalización, partiendo de Algunos cuervos son negros (Algunos A son B), si posteriormente descubrimos que Algunos cuervos no son negros (Algunos A no son B), *simplemente volvemos a nuestra observación inicial* de que Algunos cuervos son negros (Algunos A son B). En cambio, después de la segunda generalización, partiendo de Algunas cosas que no son negras no son cuervos (Algunos no B no son A), si posteriormente descubrimos la misma evidencia contradictoria, no podemos simplemente negar que Todos los cuervos son negros (Todos los A son B).

¿Por qué? Porque *esto aún nos dejaría con parte de nuestra generalización* , a saber, la afirmación de que algunos cuervos son negros (algunos A son B). Es decir, esperaríamos que «Todos los A son B» más «Algunos A no son B» produjera la conclusión armonizadora «Algunos A son B y algunos A no son B». El particular negativo no elimina el particular positivo subyacente a la generalidad positiva; puesto que previamente (debido a dicha generalización) creíamos en la generalidad, ahora tenemos un remanente que explicar.

En el caso de «Todos los cuervos son negros», tal retractación no es relevante, puesto que sabemos por experiencia que algunos cuervos son negros; pero en el caso de «Todos los cuervos son rosas», tenemos un problema grave, pues no existe ni una pizca de evidencia que respalde la afirmación de que algunos cuervos son rosas. En otras palabras, la retractación propuesta no es suficiente en la situación presentada por Hempel, es decir, cuando «Todos los A son B» se deduce de «Algunos no B son no A».

3. El problema y su solución

En mi opinión, este es el meollo del problema revelado por la exploración de Hempel. El problema no es exactamente una paradoja, ya que *la validez de la generalización depende formalmente de que dicho proceso no dé lugar a ninguna contradicción eventual* .[51] Que a partir de la observación de algunas manzanas verdes podamos, por generalización, inferir que todos los cuervos son negros y todos los cuervos son rosas, y muchas otras conclusiones contradictorias, resulta divertido, pero no alarmante. Porque en tales situaciones de autocontradicción, podemos, mediante la rectificación, encontrar maneras de armonizar nuestro conocimiento. El problema es temporal.

Por otro lado, lo que Hempel ha revelado es que no siempre podemos retractarnos simplemente mediante la particularización de las tesis en conflicto. La particularización parece aceptable en algunos casos (por ejemplo, con los cuervos negros), pero en otros produce resultados inaceptables (por ejemplo, con los cuervos rosados), porque la consecuencia lógica de dicha retractación carece de fundamento empírico.

---

[51] Una paradoja es una tesis que se contradice formalmente o que conduce deductivamente a proposiciones contradictorias. A partir de una sola paradoja de este tipo, podemos concluir que la tesis en cuestión es falsa; la lógica como tal no se ve cuestionada, ya que la contradicción implicada es meramente condicional. Una doble paradoja, en cambio, representa una seria amenaza para la lógica; en este caso, tanto la tesis *como* su contradicción son paradójicas, por lo que la contradicción es *incondicional* . En tal caso, la lógica no puede declarar ninguna de ellas verdadera o falsa, sino que debe encontrar entre ellas un non sequitur (como en la paradoja del Barbero) o un término sin sentido (como en la paradoja del Mentiroso). Es decir, la lógica debe cuestionar una o más de las implicaciones involucradas, y/o uno o más de los términos o tesis implicados. El escenario de Hempel no da lugar a una paradoja incondicional/doble.

Supongamos que, utilizando el método de Hempel, partiendo de manzanas verdes, obtenemos las generalidades «Todos los cuervos son rosas» y «Todos los cuervos son naranjas». Estas dos conclusiones son contradictorias. Digamos que decidimos resolver el conflicto negándolas ambas; aún nos quedan dos proposiciones: «Algunos cuervos son rosas» y «Algunos cuervos son naranjas».

Estas dos proposiciones no son contradictorias —y, demos por sentado, que ninguna tiene base empírica—, ¡pero de alguna manera se han arraigado en nuestro conocimiento! Se introdujeron a partir de generalizaciones basadas en ejemplos concretos, pero no desaparecieron cuando abandonamos dichas generalizaciones. Nos quedamos con ellas, incluso aunque los complejos procesos que las originaron hayan sido revocados.

Resulta impensable que tales *detalles* (sean verdaderos o falsos) surjan de la observación, aparentemente inconexa, de manzanas verdes (o cualquier otra cosa). Creo que este es el problema fundamental que Hempel descubrió. El problema no radica en el conflicto de generalidades ni entre proposiciones generales y particulares, por lo que no se trata de una paradoja. El problema tiene que ver con el «daño colateral» al conocimiento, derivado de la corrección incompleta de errores.

Sugiero la siguiente solución: cuando generalizamos de "Algunos A son B" a "Todos los A son B", y luego descubrimos que algunos A no son B, particularizamos "Todos los A son B" volviendo a "Algunos A son B". Ese es el procedimiento habitual que todos practicamos.[52]

Por otro lado, cuando obtenemos Todos los A son B mediante la generalización de Algunos no B son no A a Todos los no B son no A (seguido de la contraposición de

---

52 Simbólicamente, A + O = IO.

este último), entonces cuando descubrimos que Algunos A no son B, no podemos simplemente particularizar Todos los A son B de vuelta a Algunos A son B, sino que también debemos retractar el intermedio. premisa de la proposición Todos los A son B, es decir, Todos los no B son no A, y volver a Algunos no B son no A.

En vista de esta última retractación, de hecho ya no tenemos fundamento para afirmar que algunos A son B (esto no se puede deducir de que algunos no B no son A). Sería un error de inducción olvidar la *fuente real* de nuestra creencia en que todos los A son B. Es necesario tener presente la distinción entre los fundamentos inductivos «Algunos A son B» y «Algunos no B no son A», de modo que, en caso de descubrir evidencia contradictoria, a saber, que algunos A no son B, podamos retomar nuestra *posición anterior* en cada caso.

Podemos formular a partir de ahí la siguiente nueva ley de lógica inductiva, que puede llamarse ***ley de retracción conmensurable*** : un producto de generalización como Todos los A son B no puede ser tratado sin tener en cuenta su fuente particular; cuando alguna vez es negado por nueva evidencia, debemos volver al mismo particular inicial y no a *algún otro* particular que estaba implícito en la generalidad cuando parecía verdadera, pero que ahora ya no está implícito en ella puesto que ya no es verdadera.

En otras palabras, si nos topamos con una contradicción como la aquí considerada, debemos comprender que esto no solo desacredita la generalidad previamente inducida, sino que, más profundamente, desacredita el acto inductivo que la originó. Por lo tanto, no debemos retractarnos mediante una mera particularización, sino verificar cuidadosamente si el particular restante tiene algún fundamento independiente y, de no ser así, debemos regresar al *statu quo anterior* para asegurarnos de que no quede ningún particular sin confirmar.

Esta parece una instrucción perfectamente razonable: revertir y eliminar todo rastro de un acto inductivo que se consideró ilícito, es decir, que nos condujo a un callejón sin salida lógico.

Todo esto significa que, utilizando los procedimientos lógicos habituales, jamás caeríamos en una situación contradictoria (por ejemplo, afirmar paradójicamente que todos los cuervos son negros y todos los cuervos son rosas). Es común que las generalizaciones puedan generar resultados incompatibles; lidiamos con estos conflictos a diario sin dificultad. Cuando surgen tales conflictos, lógicamente debemos armonizarlos. Si no podemos encontrar una forma específica de resolver el conflicto, este se resuelve de manera genérica, es decir, se ponen en duda todas las generalizaciones implicadas.

En una situación donde dos o más proposiciones se ponen en duda por un conflicto mutuo, naturalmente daremos más crédito a aquella que tenga una base empírica directa (como «Todos los cuervos son negros») que a aquella que simplemente surgió de una proyección indirecta (como «Todos los cuervos son rosas»). No es necesario tratar todas las proposiciones en conflicto con la misma duda, sino que podemos ser selectivos con respecto a su origen inductivo.

Respecto a la evidencia que respalda las tesis contradictorias, es evidente que, si no disponemos de datos sobre cuervos negros o rosados, no sabríamos qué generalizar, y ambas serían problemáticas. Sin embargo, si hemos observado algunos cuervos negros y nunca hemos visto ninguno rosado, optaríamos naturalmente por la generalización de que todos los cuervos son negros (todos los A son B). Por otro lado, si hemos observado tanto cuervos negros como rosados, no haríamos ninguna generalización y simplemente combinaríamos ambos casos.

En lo que respecta a los procesos inductivos empleados, la generalización directa se favorecería naturalmente sobre la indirecta propuesta por Hempel. Si el conflicto en cuestión puede resolverse por medios ordinarios, por ejemplo, recurriendo a consideraciones empíricas, no es necesario analizar el proceso. Sin embargo, en los casos en que no disponemos de otros medios de decisión, el proceso sería, naturalmente, el objeto de la revisión.

Una posible objeción a la ley de retractación proporcional sería que, en la práctica, rara vez logramos recordar las fuentes exactas de nuestras generalizaciones. Tal desconocimiento podría ocurrir y causar estragos del tipo que Hempel describió en nuestro conocimiento.

Sin embargo, cabe señalar que, *en la práctica, casi nunca nos encontramos en la situación descrita* por Hempel. ¿Con qué frecuencia se generaliza a partir de una proposición como «Algunos no B son no A»? La respuesta estadística es «probablemente nunca»: la paradoja de Hempel es solo *una remota posibilidad formal* que los lógicos deben considerar, pero su impacto práctico es prácticamente nulo.

Además, es improbable que lleguemos a una proposición de la forma «Algunos no B son no A», salvo mediante el razonamiento descrito anteriormente, es decir, a través de otros términos como C y D. No podemos *observar directamente* que algunos no B son no A. La observación se relaciona principalmente con fenómenos positivos; puede referirse a fenómenos negativos, pero solo indirectamente. Esto sugiere que, si nos encontráramos con una situación de paradoja de Hempel, probablemente seríamos conscientes de cómo surgió.

Otra observación que vale la pena hacer es que la solución anterior al problema planteado en la paradoja de Hempel puede caracterizarse como heurística; es un trabajo de reparación por ensayo y error. Pero ya he propuesto en mi

trabajo *Lógica futura*[53] un detallado, sistemático, Tratamiento formal de la inducción, mediante factorización y revisión de fórmulas. Creo que está exento del problema de Hempel, ya que todas las posibilidades formales están incluidas en las fórmulas factoriales desarrolladas.

Con respecto a las soluciones a la paradoja de Hempel ofrecidas por otros lógicos, por ejemplo las de Goodman y Quine descritas en el artículo de Wikipedia mencionado anteriormente:

"Nelson Goodman sugirió añadir restricciones a nuestro razonamiento, como por ejemplo, no considerar nunca un ejemplo como apoyo para 'Todos los P son Q' si también apoyara 'Ningún P es Q'… Goodman, y más tarde otro filósofo, [WV] Quine, utilizaron el término *predicado proyectable* para describir aquellas expresiones, como *cuervo* y *negro* , que *sí* permiten la generalización inductiva; los predicados *no proyectables* son, por el contrario, aquellos como *no negro* y *no cuervo,* que aparentemente no la permiten. Quine sugiere que es una cuestión empírica qué predicados, si los hay, son proyectables; y señala que en un dominio infinito de objetos, el complemento de un predicado proyectable siempre debería ser no proyectable. Esto tendría como consecuencia que, aunque «Todos los cuervos son negros» y «Todas las cosas no negras no son cuervos» deben tener el mismo apoyo, ambos derivan todo su apoyo de los cuervos negros y no de los no-cuervos no negros."

Considero que estas propuestas son razonables y compatibles con la mía. Sin embargo, creo que la mía es un poco más precisa al identificar el problema y su solución.

---

[53] Publicado por primera vez en 1990, pocos años antes de la muerte de Hempel. Véase la parte VI.

La sugerencia de Goodman de restringir la inducción a partir de una proposición si dicho proceso produce conclusiones contradictorias es lógicamente sólida. Sin embargo, su instrucción no puede seguirse de forma preventiva, sino solo *después* de descubrir que el proceso produce conclusiones contradictorias. Por lo tanto, no se trata de una medida preventiva, como él parece considerarla, sino de una corrección a posteriori. En consecuencia, puede considerarse similar a la ley de retractación proporcional que propongo anteriormente. La única diferencia radica en que él no parece haber distinguido entre el conflicto de generalidades y los particulares subyacentes.

En cuanto a los «predicados no proyectables», coincido en que los términos negativos (complementos) plantean un problema general en la inducción. Si bien la lógica deductiva no distingue entre términos positivos y negativos, la fenomenología sí distingue entre la presencia y la ausencia de fenómenos positivos. Mientras que podemos observar un fenómeno positivo (como un cuervo negro) sin tener en cuenta su negación, no podemos mencionar un término negativo (como «no negro» o «no cuervo») antes de pensar en el fenómeno positivo correspondiente y buscarlo, sin llegar a encontrarlo.[54]

Por lo tanto, un término verdaderamente negativo nunca puede ser verdaderamente empírico. Su contenido nunca es «He visto algo», sino siempre «He buscado diligentemente algo y no lo he encontrado». Un término negativo es «empírico» en un sentido menor y más derivado que uno positivo. Implica ya una generalización, desde «no se pudo encontrar» hasta «no estaba allí para ser encontrado».

De esta observación se deduce que la generalización a partir de términos negativos, como «Algunos no B son no

[54] Véase mi ensayo sobre este tema en *Reflexiones* (parte I, capítulo 9).

A», solo puede llevarse a cabo con una cautela y un escepticismo excepcionales. El escenario de Hempel justifica aún más esta cautela. Incluso podríamos vernos tentados, como solución radical, a prohibir siempre la generalización cuando se trata de un sujeto verdaderamente negativo. Si algún tipo de discurso tiene ciertas consecuencias ilógicas probables, los lógicos harían bien en formular una ley lógica preventiva de este tipo.

Otra tentación es negar cualquier contenido significativo a las proposiciones de la forma «Algunos no B son no A». Esta proposición está formalmente implícita en «Todos los A son B» y es compatible con «Algunos A son B», «Ningún A es B» y «Algunos A no son B», pero ¿realmente nos aporta alguna información? De hecho, dado que no se puede inferir nada sobre A o B (como sujetos) a partir de «Algunos no B son no A», ¿qué información contiene dicha proposición? ¿No sería concebible afirmar tal proposición utilizando cualquier par de términos elegidos al azar? Este tipo de duda podría usarse para justificar aún más la prohibición de generalizar a partir de tales proposiciones.[55]

Sin embargo, dado que es posible una solución menos radical, a saber, la ley de retractación proporcional propuesta anteriormente, quizás no sea necesario llegar tan lejos. En lugar de prohibir de antemano ciertos procesos dudosos en todas las circunstancias, prefiero

---

[55] Estas cuestiones se aclaran si consideramos la eventual negación de «Algunos noB son noA», es decir, la forma «Ningún noB es noA», que implica «Todos los noB son A». En caso de que esta última proposición sea verdadera, tendríamos un término negativo (noB) incluido en uno positivo (A). Esto podría interpretarse como que casi todo el mundo (excepto las cosas que son B) se incluye en A. Para que esto ocurriera, A tendría que ser un concepto muy amplio. Un concepto así sería muy excepcional y prácticamente carente de significado. Por lo tanto, podemos decir que «Algunos noB son noA» es casi siempre verdadero, y al mismo tiempo, poco informativo.

permitirlos en caso de que funcionen ocasionalmente y preparar el mecanismo correctivo adecuado para cuando no funcionen.

En resumen, creo que hemos demostrado de forma convincente que la supuesta paradoja de Hempel no representa una dificultad insuperable para la lógica; se la presenta como algo más compleja de lo que realmente es. Aun así, constituye una aportación interesante para que los lógicos reflexionen sobre ella.

11. La paradoja de la predicción de Goodman

De Reflexiones lógicas y espirituales 1.9.

1. El supuesto problema

Nelson Goodman [56]propuso un «enigma de inducción» (como él lo denominó [57]) o « paradoja de la predicción» (como otros lo han caracterizado) , que parecía demostrar una dificultad formal en la generalización. Esto puede enunciarse de la siguiente manera:

"Goodman… introdujo el color grue, que se aplica a todas las cosas examinadas antes de un tiempo determinado *t,* por si acaso son verdes, pero también a otras cosas por si acaso son azules y no se examinaron antes del tiempo *t* . Si examinamos esmeraldas antes del tiempo *t* y encontramos que la esmeralda *a* es verde, la esmeralda *b* es verde, y así sucesivamente, cada una confirmará la hipótesis de que todas las esmeraldas son verdes. Sin embargo, las esmeraldas *a, b, c,* etc., también confirman la hipótesis de que todas las esmeraldas son grue. En este caso, las esmeraldas *a, b, c, ...* examinadas después del tiempo *t* deberían ser grue **y, por lo tanto, azules** ." (Énfasis mío)[58]

---

56 Estados Unidos, 1906-98.

57 O, dicho de forma más pretenciosa, "el nuevo problema de la inducción".

58 Aquí cito: http://en.wikipedia.org/wiki/Nelson_Goodman . En otro lugar, se nos informa que "se aplica a todas las cosas examinadas antes de *t* por si acaso son verdes, pero a otras cosas por si acaso son azules" es la propia formulación de Goodman en su

Según sus defensores, la importancia de este artificio radica en que, si bien "green" y "grue" tienen la misma forma lingüística y, por lo tanto, deberían estar sujetos a los mismos procesos lógicos (en este caso, el proceso inductivo de generalización), internamente son tipos de conceptos bastante diferentes, ya que el primero implica una similitud entre sus instancias pasadas y futuras, mientras que el segundo sugiere un cambio de color a lo largo del tiempo, de modo que el resultado es paradójicamente bastante diferente si generalizamos con referencia a uno u otro.

2. La solución lógica

Sin embargo, como demostraré formalmente a continuación, esto no es más que un truco, pues aunque la generalización sea igualmente válida para el verde y el grisáceo, no se deduce que podamos inferir que todas las esmeraldas son azules a partir de la proposición general de que todas son grises. Es decir, la conclusión «y por lo tanto azules» en la presentación anterior es una deducción errónea.

Para exponer este sencillo error, el escenario planteado debe reformularse con mayor cuidado (los símbolos X, A, B, C son míos):

- Digamos que examinamos todas las esmeraldas disponibles (X) hasta cierto momento (t), y encontramos que todas son verdes (A), normalmente concluimos por generalización que Todas las esmeraldas son verdes (Todas las X son A), aunque sabemos [por experiencia previa con la inducción en general] que la siguiente esmeralda que encontremos, después del tiempo t, bien

---

presentación original en *Fact, Fiction, and Forecast* (http://en.wikipedia.org/wiki/Grue_%28color%29).

podría resultar ser azul (B) [¡o incluso de algún otro color [59]!].

- Sigamos ahora a Goodman e introduzcamos un nuevo concepto "verde" (C) que se definirá como la clase que agrupa todas las cosas que fueron examinadas antes de un cierto tiempo t y se encontró que eran verdes (A) y todas las cosas que *no* fueron examinadas antes del tiempo t y que resultan *ser* azules (B) [o de hecho, de algún otro color].[60]
- Aplicando esta definición, todas las X (esmeraldas) examinadas antes del tiempo t que resultaron ser A (verdes) también son C (grises); es decir, por silogismo podemos inferir que algunas X son C. En cuanto a los casos restantes de X, aquellos que no se examinan hasta después del tiempo t [si es que se examinan], *cada uno resultará ser A (verde) o B (azul) [o incluso de algún otro color]* ; en ese sentido, estas últimas X también son C. Por lo tanto, la conclusión de que todas las X son C parece razonable.

---

59 Esta es mi propia interpolación, para hacer más precisa la tesis de Goodman. Pues no hay razón para suponer a priori que solo se puedan encontrar esmeraldas azules. Simplemente estamos especulando sobre la posibilidad de que existan esmeraldas azules, sin basarnos en observaciones específicas; por lo tanto, cualquier otro color es igualmente probable (o improbable). Sin embargo, mi refutación de Goodman funciona igual de bien sin este comentario adicional.

60 Nótese que se afirma que estas últimas cosas simplemente "no se examinarán hasta el momento t [aún, si es que alguna vez se examinan]"; esto no debe confundirse (como han hecho algunos comentaristas) con "se examinarán después del momento t", pues, independientemente de cuántas cosas lleguemos a examinar, es evidente que nunca lograremos (ni sabremos que hemos logrado) una enumeración completa de todas ellas en el universo. Nótese también que el concepto grue se define aquí como un predicado general para cualquier sujeto eventual ("cosas"), en lugar de específicamente para las esmeraldas.

- Pero ciertamente no se deduce lógicamente de lo anterior que se pueda encontrar alguna esmeralda de otro color que no sea verde, es decir, que cualquier X sea B [es decir, azul o cualquier otro color que no sea verde]. Porque, entendida correctamente, la categoría C *no se* formula como una disyunción de A o B que necesariamente actualice algunos casos de XA y algunos casos de XB.
- Si observas con atención, verás que C incluye, por un lado, elementos que ya *se sabe* que son A (esmeraldas verdes ya observadas) y, por otro, una gama de elementos cuya calificación aún *se desconoce* , es decir, A o B (azul) [o incluso algún otro color]. Esto último es una disyunción de resultados *posibles* , no de resultados *inevitables* . Por lo tanto, inferir XB como un resultado *real* sería una *falacia lógica* .
- El hecho de que desconozcamos si algún X futuro será A no nos permite inferir, a partir de esta disyunción de posibilidades, que algún X futuro sea necesariamente B. Tampoco sabemos aún si algún X futuro será B. Es muy posible que la afirmación «Todos los X son A» («Todas las esmeraldas son verdes») siga siendo válida indefinidamente después del tiempo t, al igual que antes del tiempo t (como se predijo en la generalización ordinaria inicial).
- Las premisas «Todos los X son C» y «Todos los C son A o B» conducen, en efecto, a la conclusión silogística «Todos los X son A o B». Sin embargo, la disyunción «A o B» no puede interpretarse de forma distinta en la premisa mayor y en la conclusión. Dado que la disyunción en la premisa no es extensional, la disyunción en la conclusión no puede considerarse

extensional [61]. Hacerlo equivaldría a cometer la falacia de los cuatro términos.

Por lo tanto, queda claro en nuestra exposición que *la introducción del concepto «grue» no ha alterado en absoluto las posibilidades inductivas que ofrecen los datos presentados* . La conclusión inductiva correcta permanece inalterada por las artimañas de Goodman. Lo único que ha logrado Goodman es ocultar hábilmente su razonamiento *deductivo falaz* (una interpretación errónea del tipo de disyunción implicada); todo es mera sofistería.

En medio de la densa retórica de Goodman, se da la impresión de que predecir la presencia de esmeraldas azules es tan sencillo como predecir la de esmeraldas verdes. Pero esa no es en absoluto la conclusión lógica según la lógica inductiva. ¿Por qué? Porque, en el caso de la hipótesis de que las futuras esmeraldas observadas serán verdes, contamos con datos concretos que la respaldan: todas las esmeraldas observadas, tanto presentes como pasadas, han resultado ser verdes.

Si bien, en apoyo de la hipótesis de Goodman sobre la aparición de esmeraldas azules, hasta el momento carecemos de evidencia empírica, lo único que podemos afirmar es que no es inconcebible que algún día se encuentren esmeraldas azules, pero esto no implica que vayan a aparecer. El hecho de que no sea inconcebible no justifica una predicción concreta; simplemente significa que es imaginable dentro del contexto actual del conocimiento.

Es decir, lo único que tenemos es un principio epistemológico general que nos invita a mantener una

---

[61] Es decir, la base de la disyunción dada es «Algunos C *podrían ser* B», mientras que la base correspondiente de la disyunción supuestamente inferida es «Algunos X *son* B». Pero imaginar que algo sucede no prueba que deba suceder en algún momento. La conclusión no se deriva de la premisa.

mente abierta ante cualquier resultado posible, basándonos en la experiencia pasada relacionada con todo tipo de objetos, y en la constatación de que la novedad aparece ocasionalmente. Pero esta apertura mental científica no equivale a una predicción positiva de cambios específicos. Se trata simplemente de un llamado, en nombre del realismo, a evitar los prejuicios y la rigidez.

3. El concepto de 'grue'

Una pregunta que deberíamos hacernos es si el concepto de "grue" de Goodman está bien definido.

El concepto común de «verde azulado» se formularía simplemente como «verde y/o azul». Podría ser útil introducir este concepto, quizás para destacar que el verde y el azul están cerca en la gama de colores, o que algunas cosas son en parte verdes y en parte azules, o a veces verdes y a veces azules, o que algunos tonos intermedios son difíciles de clasificar como claramente verdes o claramente azules. Al fin y al cabo, la línea divisoria entre estos colores es bastante arbitraria.

Dado que algunas esmeraldas son verdes, podríamos deducir que algunas son de color grisáceo. Sería igualmente válido inferir, a partir de ahí, que todas las esmeraldas son verdes o que todas son de color grisáceo. Esto no implicaría ninguna contradicción inherente, ya que afirmar que todas las esmeraldas son de color grisáceo no implica (ni excluye) que alguna sea azul. Que todas las esmeraldas sean de color grisáceo es formalmente compatible con la posibilidad de que todas sean verdes. Por lo tanto, en realidad no existe ninguna «paradoja de la predicción».

El concepto “grue” de Goodman no difiere de este concepto ordinario en lo que respecta a dicha implicación

lógica. Su diferencia no radica en la disyunción (verde o azul), puesto que dicha disyunción es bastante común; por ejemplo, el concepto "de color" significa (aproximadamente) "rojo, naranja, amarillo, verde, azul, índigo o violeta". La diferencia significativa en el concepto de Goodman reside en la implicación de condiciones temporales y epistémicas. Esto cumple la función retórica de confundir las cuestiones.

Definir el concepto de "verde" como la clase de todas las cosas examinadas antes del tiempo t y que resultaron ser verdes, y todas las cosas *no* examinadas antes de t que resultan *ser* azules, implica una especie de contradicción. Si aún no he examinado las cosas después del tiempo t, ¿cómo puedo afirmar con certeza que alguna de ellas es azul? Solo podría hacer tal afirmación *a posteriori* , después de haber examinado algunas de las cosas después del tiempo t y haberlas encontrado azules.

Alternativamente, tendría que ser dicho *por un «tercero» que observa la situación* , que ha examinado algunas de las cosas antes del tiempo t y las ha encontrado azules, y que observa mi situación antes de que yo haga lo mismo. Pero en cuanto a todos los observadores actuales en conjunto, no pueden adoptar lógicamente tal hipótesis sobre cosas que resultan ser azules aunque aún no se haya observado que lo sean. Solo podemos hablar de forma coherente sobre cosas que *podrían llegar* a ser azules. Por esta razón, el concepto de «gris» de Goodman no está bien formulado.

El concepto de "verde" se define principalmente como la unión de cosas verdes y azules; sin embargo, de esta definición no se deduce que si algunas cosas (como las esmeraldas) son verdes, otras *(* es decir, otras esmeraldas) deban ser azules. Decir que un tipo de cosa (las esmeraldas) es "verde" no implica que sus ejemplos deban abarcar toda la gama de posibilidades incluidas en este

concepto. El concepto de "verde" sigue siendo legítimo siempre que encontremos los predicados que agrupa (verde, azul) presentes en diversos tipos de cosas (esmeraldas, el mar, etc.).

Así pues, todo predicado ordinario implica cierta incertidumbre en cuanto a su aplicación a sujetos específicos. Además, esta aplicabilidad puede variar con el tiempo: según nuestro contexto de conocimiento y según los cambios que se produzcan en los objetos observados. Por consiguiente, no es necesario incluir tales factores epistémicos y temporales en la definición de ninguno de los conceptos que proponemos. Dichos factores son inherentes a la conceptualización.

La razón por la que Goodman introdujo tales complicaciones en su definición de "grue" fue porque quería refutar (o dar la impresión de que estaba refutando) el proceso de generalización que comúnmente usamos para desarrollar nuestro conocimiento sobre la base de una observación limitada.

Según la lógica inductiva, al observar que algunos X son A, y al no haber encontrado hasta ahora ningún X que no sea A, podemos generalizar y afirmar que todos los X son A. Esto sigue siendo cierto para nosotros mientras no tengamos evidencia de ningún X que no sea A. La generalización implica predicción, es decir, afirmar algo sobre casos de X que aún no hemos observado y que tal vez nunca observemos.

Goodman pretendía demostrar que estamos igualmente justificados al predecir un resultado negativo (es decir, no A, por ejemplo B) que un resultado positivo (es decir, A) [62]. No comprendió la *justificación lógica* de nuestras

---

62 Para ello, necesitaba construir un concepto que incluyera tanto A como no-A, de modo que se pudiera demostrar formalmente que la generalización podía darse en ambos sentidos. Sin embargo, dado que un concepto que incluye contradictorios no es informativo,

generalizaciones [63]. No predecimos arbitrariamente que los casos de X que observemos en el futuro serán A en lugar de no A. Simplemente nos atenemos a *la misma polaridad* (A), porque es la única polaridad para la que tenemos evidencia empírica hasta el momento. En comparación, predecir *la polaridad opuesta* (no A, en este contexto) sería puramente arbitrario, una afirmación descabellada. Específicamente para X, la primera opción cuenta con cierto respaldo empírico, mientras que la segunda carece por completo de él.

Goodman simplemente no se percató de esta diferencia en la justificación entre ambos enfoques, aunque resulta obvia para cualquiera que se detenga a reflexionar. Por lo tanto, no aplicó el principio inductivo de que una hipótesis confirmada siempre debe preferirse a una no confirmada. Además, como vimos, en su afán por invalidar el razonamiento inductivo, ¡cometió uno de los errores más elementales del razonamiento deductivo!

4. ¿Hasta qué punto en la escala?

Detrás del enigma de Goodman subyace otra cuestión importante para la lógica inductiva: ¿hasta qué punto se pueden generalizar legítimamente las cosas en cualquier escala de clasificación? Si para un sujeto dado se ha generalizado un predicado determinado, ¿por qué no generalizar aún más arriba en la escala a un predicado mayor?[64]

Consideremos un sujeto X y dos predicados cualesquiera S y G, relacionados como *especie y género* , es decir, tales que todos los S son G pero no todos los G son S (es decir,

---

incluyó contrarios, es decir, 'A o B' (donde B no es A). Esto disimula ligeramente el problema, pero en realidad no lo cambia.

63 Véase mi *libro Lógica del futuro* , capítulo 50.

64 He abordado este tema (indirectamente, en lo que respecta a la lógica ética) en mi *obra Lógica Judaica* , capítulo 13.3.

algunos que no son S también son G). Cabe destacar que S y G son conceptos comunes, como verde y color.

- Si todos los casos de X que hemos observado hasta ahora son S, y hemos buscado y no encontrado ningún X que no sea S, podemos inferir inductivamente que todos los X son S. Esta generalización sigue siendo válida siempre que no se encuentren casos de X que *no sean* S; pero si finalmente aparece algún X-nonS, la lógica inductiva nos exige revisar nuestro juicio anterior y particularizarlo a algunos X son S y algunos X no son S. Porque la inducción procede condicionalmente [65].

- El mismo razonamiento se aplica a G. [66]Alternativamente, si se admite que todos los S son G, podemos inferir *deductivamente , mediante silogismo (Figura* [1] , AAA), que todos los X son G. Por lo tanto, podríamos postular que, si estamos justificados para generalizar, para un sujeto X dado, hasta el predicado específico S, también estamos justificados para hacerlo en un nivel superior de la escala de clasificación, hasta el predicado más general G. Esto es lógicamente correcto si se comprende y aplica adecuadamente.

- Sin embargo, sería un grave error de juicio [67]inferir de una generalización tan válida que podría haber

---

[65] La aducción o generalización se justifica por *dos* principios esenciales: (1) la confirmación de una hipótesis mediante un ejemplo positivo, y (2) la no refutación de dicha hipótesis por ningún ejemplo negativo. *Ambos* principios deben cumplirse por igual para que el proceso sea lógico. Por supuesto, existen muchas otras condiciones involucradas; véase mi ensayo «Principios de la aducción» en *Fenomenología* (capítulo VII.1).

[66] Es decir, dado que algunos X son G (o deduciendo esto de que algunos X son S), podemos generalizar a que todos los X son G, siempre que no haya ninguna instancia negativa conocida (X-noG) que lo desmienta.

[67] Como vimos, este fue uno de los errores que Goodman cometió al formular su "acertijo". Este error es de naturaleza deductiva, no inductiva.

algunos X que sean G pero no S (incluso si sabemos que hay *cosas)* . *Excepto X* que son G pero no S). En esta etapa, el contenido real Todos los X son G es idéntico (en extensión e implícitamente en intensión) al de Todos los X son S del que se derivó [68]. La diferencia entre ambas afirmaciones radica únicamente en la eventual *particularización correctiva* …

- Supongamos que mañana descubrimos *una X que, aunque sigue siendo G, no es S* (por ejemplo, una esmeralda de un color distinto al verde). En tal caso, tendríamos que particularizar la primera afirmación (la más específica) a «Algunas X son S y algunas X no son S»; pero la segunda afirmación (la más genérica) «Todas las X son G» permanecería inalterada.[69]
- Pero como resultado de dicha particularización, la afirmación «Todos los X son G» tiene *un significado más vago* , ya que G ya no se refiere para nosotros solo a la especie S de G, sino también a alguna otra especie (no S) de la misma. Por lo tanto, aunque la regla inductiva sería generalizar hasta donde podamos llegar, debemos tener en

---

[68] Esto resulta obvio si consideramos que podemos obtener igualmente «Todos los X son G»: (a) por generalización a partir de «Algunos X son G», que deducimos de «Algunos X son S», o (b) por deducción a partir de «Todos los X son S», que generalizamos a partir de «Algunos X son S». En realidad, podría argumentarse que estas dos opciones son ligeramente diferentes, ya que (a) exige que nos aseguremos de que no existan instancias de X que no sean G, mientras que (b) exige que nos aseguremos de que no existan instancias de X que no sean S. Sin embargo, esta diferencia se pone de manifiesto en la etapa subsiguiente de particularización.

[69] Nótese que si descubrimos un X que no es G, necesariamente tampoco es S, dado que todos los S son G. En ese caso, ambas proposiciones generales tendrían que ser particularizadas.

cuenta que cuanto más avancemos en la escala, más diluiremos el significado final de nuestra generalización.[70]

Así, aunque en principio la generalización ascendente es ilimitada, en la práctica procedemos con relativa lentitud para mantener la utilidad noética de nuestras ideas y enunciados. Para dar un ejemplo formal: la proposición «Todos los X son S» podría usarse como premisa menor en un silogismo donde S es el término medio, mientras que la proposición «Todos los X son G» —aunque idéntica en extensión e intensión a la anterior— sería inútil en ese mismo contexto (es decir, con S como término medio).

Además, considerar que "Todos los X son G" es una generalización más rentable que "Todos los X son S", en el sentido de que nos proporciona información sobre más cosas por el mismo precio en términos de datos dados, indica una confusión [71]entre la generalización para un sujeto dado de un predicado más restringido a un predicado más amplio, y la generalización de un predicado dado de un sujeto más restringido a un sujeto más amplio.

Este último caso constituye la forma verdaderamente provechosa de generalización. Supongamos que todos los X son P, y que Y es una superclase de X (es decir, todos los X son Y, aunque no todos los Y son X); entonces, esto consistiría en inducir que todos los Y son P, salvo que se descubra algún Y que no sea P. Las reglas de dicha

---

[70] En este contexto, podríamos comparar el concepto de "grue" de Goodman con el de "oomph" de Feynman. Este último, definido (en tono de broma) como "una especie de tendencia al movimiento", podría parecer útil para "explicar" diversos fenómenos, pero es tan vago que no puede predecir nada y, por lo tanto, carece de valor (p. 19).

[71] Por cierto, Goodman era culpable de ello al formular su "acertijo".

generalización se tratan en detalle en mi obra *Lógica del Futuro,* en el apartado de Inducción Factorial (Parte VI).

12. La paradoja de Sorites

De Temas (publicados solo en línea hasta el momento).

1. ¿Qué es un montón?

La paradoja de Sorites no es una paradoja en el sentido estricto del término, sino una pregunta. A veces se plantea de forma sofística, para que parezca paradójica. Pero puede formularse de manera más directa, en cuyo caso se revela como simple, aunque no exenta de importancia. El término *sorites* es latino y deriva del griego *sōros* , que significa montón.

Una forma de expresar la paradoja de Sorites es: ¿Qué es un «montón» (de guijarros, por ejemplo)? O, ¿cuántos guijarros (por ejemplo) constituyen un «montón» de ellos? La respuesta obvia es que seguramente debe haber al menos un guijarro. Si no hay guijarros, no hay un montón, sino un no-montón. Pero, ¿basta con un guijarro? La respuesta obvia es: no, se necesita *al menos Dos* guijarros forman un montón, ya que montón es un término colectivo, y que además sugiere que los guijarros están apilados uno encima del otro (y no se puede apilar algo que no sea un guijarro sobre otro guijarro, ni un guijarro sobre otro que no lo sea). Un solo guijarro, lógicamente, no constituye un montón; montón es intrínsecamente plural.

Formulada así, la pregunta no es muy problemática. Pero si se formula de la siguiente manera, se vuelve más complicada. Si tenemos muchas piedrecitas (digamos, 100) apiladas, obviamente tenemos un montón. ¿Qué

sucede si quitamos una piedrecita? ¿Seguimos teniendo un montón? Sí, 99 forman un montón. ¿Qué sucede si quitamos una piedrecita más? ¿Seguimos teniendo un montón? Sí, 98 forman un montón. Y así sucesivamente, hasta llegar a números pequeños, momento en el que el sofista se pregunta si dos piedrecitas constituyen un montón, luego una piedrecita, luego ninguna. En este punto, la pregunta parece paradójica, bastante artificial: ¿deberíamos concluir que ninguna piedrecita forma un montón, o que una piedrecita forma un montón? El sofista cree que sí; el verdadero lógico sabe que no.

2. El uso de términos vagos

Evidentemente, el problema aquí, si es que lo hay, radica en la formulación precisa de términos inicialmente vagos. Si no nos detenemos *desde el principio* a reflexionar sobre la intención exacta de un término vago, ya sea de este tipo (un término que sugiere cantidad, por ejemplo) o de cualquier otro, podríamos encontrarnos con dificultades *más adelante* en nuestro discurso. Por lo tanto, debemos detenernos a pensar antes de usar tales términos y anticiparnos a cualquier dificultad que puedan generar. En el presente caso, como hemos visto, el término (montón) es inherentemente plural (y, por lo tanto, inaplicable a menos de dos guijarros). Un sofista prefiere complicar el asunto, para poner en duda la razón y el conocimiento humanos, como suele hacer; pero el asunto es realmente bastante simple.

En algunos casos, sin duda, la definición de un término vago tiene un componente *convencional* . Esto se puede ilustrar con otra versión de la llamada paradoja de Sorites: ¿Cuántos cabellos tiene un hombre para no ser calvo? Obviamente, un hombre sin cabello es definitivamente calvo. Pero, ¿acaso un hombre con un solo cabello, o con muy pocos, no sería considerado calvo en el lenguaje

cotidiano? Quizás sí, pero solo por convención. Si definimos "calvo" *estrictamente* , implica cero cabello; pero si lo definimos *de forma amplia* , puede incluir casos que involucren un número *arbitrario* , aunque preferiblemente pequeño, de cabellos, determinado por convención; por ejemplo, cinco o veinte cabellos [72]. Claramente, si queremos evitar confusiones, nada nos impide referirnos a este sentido más amplio de "calvo" como, más precisamente, "calvo o casi calvo".

Esta cuestión de la vaguedad no tiene nada de especial; no se limita a los términos que dan lugar a la paradoja de Sorites. Por ejemplo, un término relativo como «pequeño» (o su relativo, «grande») es *inherentemente vago* y solo puede usarse con precisión en situaciones específicas mediante una magnitud convencional. Asimismo, al tratar con continuos, puede que necesitemos establecer líneas divisorias arbitrarias. Por ejemplo, no existe una línea divisoria objetiva entre un color y un color adyacente en el espectro, y en algunas circunstancias puede ser necesario que imaginemos una (por ejemplo, con fines legales u otros fines prácticos [73]). Las distinciones convencionales forman parte del pensamiento humano; pero, es importante recalcar, no lo

---

[72] O cualquier cantidad mínima de vello que, subjetivamente, nos parezca tan cercana a la calvicie que sea, en la práctica, calvicie.

[73] Por ejemplo, en la ley judía (*halajá*) se presta mucha atención a las definiciones cuantitativas, especialmente a los máximos y mínimos de duraciones, horas, distancias, longitudes, volúmenes, pesos, temperaturas, valores monetarios, etc. Inicialmente, los rabinos solían expresar estas medidas en términos vagos (por ejemplo, "el volumen de un huevo"), pero posteriormente se requirieron formulaciones más precisas (que diferentes autoridades podían estimar de manera distinta). Sin embargo, algunas medidas siguen siendo subjetivas (por ejemplo, la estimación de cuándo uno se siente saciado después de comer). Para más detalles, consulte: http://halachipedia.com/index.php?title=Reference_of_Measurements_in_Halacha .

constituyen todo. Siempre hay elementos objetivos detrás de las convencionales. Por ejemplo, la línea divisoria entre el azul y el verde estaría en algún punto intermedio entre lo que vemos claramente como azul y lo que vemos claramente como verde; nunca estaría muy alejada de un extremo ni del otro, y mucho menos entre el verde y el amarillo, o entre el azul y el índigo, o aún más lejos.

Volviendo a nuestra supuesta paradoja, conviene hacer algunos comentarios más. Como ya se ha dicho, no hay nada paradójico en el concepto de montón, si se define correctamente. En pocas palabras, un montón puede definirse como objetos materiales colocados unos encima de otros de cualquier forma, lo que implica que debe haber dos o más objetos. Una definición más compleja supondría que un montón debe ser piramidal, es decir, requiere al menos cuatro objetos (tres para la base y uno encima); pero esto no nos supone ninguna dificultad, ya que se puede hablar más específicamente de un montón piramidal. Del mismo modo, la calvicie *estricta sensu* se refiere a la ausencia total de cabello; sin embargo, se aplica a pequeñas cantidades de cabello, aunque solo en sentido figurado.

En nuestro lenguaje cotidiano abundan los términos vagos de este tipo [74], algunos de los cuales requieren una

[74] En efecto, esto es inevitable por dos razones. Primero, muchos objetos concretos son imposibles de definir con precisión. ¿Dónde termina, por ejemplo, una naranja exactamente? ¿El perfume o el calor que emana de la fruta forman parte de ella o no? ¿En qué momentos del tiempo y del espacio pueden considerarse lógicamente separadas dichas emanaciones? Segundo, dado que el conocimiento humano es inductivo, no siempre podemos partir de una definición precisa de un concepto, sino que tendemos a dejarlo abierto, para que se defina con mayor precisión a medida que se desarrolla la experiencia. Desde esta perspectiva, la mayoría de los términos abstractos que utilizamos son abiertos, incluyendo términos que pueden usarse para definir con mayor precisión otros términos; por lo tanto, aquí también la vaguedad es inevitable. Pero tales dificultades

definición convencional por razones prácticas. Por ejemplo, el término «multitud» podría referirse a una reunión de tres o más personas, según el dicho popular «dos son compañía, pero tres son multitud»; o podríamos, por ejemplo, en el software policial para monitorear grandes grupos de personas, optar por un mínimo mayor (digamos, 50 o 500), establecido arbitrariamente como motivo de alarma. La palabra «turba» podría preferirse cuando esta última multitud se descontrola.

En cualquier caso, no existe ningún problema lógico real en tales expresiones cuantitativas inespecíficas. No constituyen un defecto del lenguaje ordinario que nos obligue a construir un «lenguaje ideal» donde todos los términos tengan un único significado preciso [75]. Mucho menos requieren un tratamiento mediante lógicas simbólicas abstrusas (tan apreciadas por muchos lógicos modernos). Al contrario, demuestran la versatilidad y flexibilidad del pensamiento y el habla cotidianos, y evidencian que gran parte de nuestra comunicación lingüística tiene matices no verbales que, en su mayoría, comprendemos con gran facilidad. La mayor parte de nuestro uso diario de términos vagos *no requiere* una

ontológicas y epistémicas no implican una paradoja; simplemente invitan a la reflexión filosófica.

75 El recurso a un «lenguaje ideal» por parte de ciertos lógicos modernos para resolver un problema de lógica es inútil. Incapaces de comprender la forma en que los seres humanos reales abordamos lógicamente ciertas dificultades cognitivas, intentan imponernos una manera de pensar superficial, artificial e impráctica. El papel del verdadero lógico no es imponer lógicas imaginarias, sino comprender nuestros medios lógicos naturales y, a partir de ahí, perfeccionarlos y reforzarlos. El razonamiento humano debería ser la principal preocupación de los lógicos. La manera natural de abordar las paradojas de Sorites es usar las palabras con mayor precisión; por ejemplo, en lugar de llamar «calvos» a las personas con muy poco cabello, llamarlas «casi calvas»; o, aún con mayor precisión, si es necesario para algún propósito práctico, «que tienen (digamos) entre uno y diez cabellos».

definición más precisa. Se utilizan para sugerir cosas de forma aproximada y no pretenden ser afirmaciones exactas y veraces. Si surge la necesidad de precisión y veracidad, se aborda de manera que se preserve la coherencia (mediante una convención explícita si es necesario).

3. Razonamiento con términos vagos

Sin duda, los términos vagos pueden ser peligrosos si intentamos razonar con ellos. Pero a menudo se utilizan sin involucrarlos en procesos argumentativos. Además, razonar con términos vagos no siempre es inválido: existen contextos donde la vaguedad no impide llegar a una conclusión fiable.

En el silogismo categórico (o hipotético), la regla relativa a los términos (o tesis) vagos es la siguiente: El elemento central (término o tesis, según sea el caso) no puede ser vago, ya que no ofrecería garantía de que su intención sea la misma en ambas premisas. Si el elemento central es vago, no podemos estar seguros de la superposición y la conclusión es inválida. Por otro lado, los elementos mayor y menor pueden ser vagos sin afectar el argumento, y esto se aplica a las cuatro figuras. Sin embargo, existe una excepción a esta regla: cuando el término central es vago, pero no tanto como para que no se garantice la superposición, aún se puede formular un argumento válido.

Esto último se evidencia al considerar los dos argumentos siguientes en la tercera figura, donde el término medio garantiza la superposición, aunque ninguna de las premisas sea universal. Las expresiones «la mayoría» (más de la mitad) y «pocos» (menos de la mitad) son vagas, ya que no especifican cifras exactas. Pero observemos la particularidad (en contraposición a

«mayoría» o «minoría») en la conclusión, es decir, la mayor vaguedad de la misma.

La mayoría de M son P
y la mayoría de las M son S;
Por lo tanto, algunos S son P

Pocos M son P (lo que implica que la mayoría de los M no son P).
y la mayoría de las M son S;
Por lo tanto, algunos S no son P

En la apódosis, las reglas relativas a los términos o tesis vagos son las siguientes:

Modus ponens:
Si A es B, entonces C es D, y A es B (afirmando el antecedente); entonces, C es D (el consecuente es afirmado).
El antecedente y la premisa menor no pueden ser vagos; de lo contrario, la conclusión es inválida. Sin embargo, el consecuente podría ser vago, y la conclusión seguiría siendo válida (aunque también vaga).

Modus tollens:
Si A es B, entonces C es D, y C no es D (negando el consecuente); entonces, A no es B (se niega el antecedente).
El consecuente y la premisa menor no pueden ser vagos; de lo contrario, la conclusión es inválida. Sin embargo, el antecedente podría ser vago, y la conclusión seguiría siendo válida (aunque también vaga).

Una vez más, en ambos casos, cabe concebir una excepción si sabemos que las premisas mayor y menor se superponen, aunque no sepamos con precisión cuánto se superponen.

Se pueden formular reglas similares para otras variedades de argumentos, como el dilema o el argumento a fortiori.

Quienes afirman que los términos vagos son inherentemente paradójicos se dedican a buscarle tres pies al gato, motivados por el deseo de impresionarse a sí mismos o a los demás con su supuesta capacidad para encontrar y resolver paradojas (artificiales), o (peor aún) para demostrar que el conocimiento humano es inevitablemente paradójico y, por lo tanto, inútil. Claramente, así como es deshonesto llamar montón a una sola piedra o a ninguna, también lo es llamar calva a una persona con uno o más cabellos [76]. Si se incurre en tales contradicciones desde el principio, inevitablemente se llegará a paradojas. Quienes actúan así no son verdaderos lógicos, sino sofistas. Manipulan y fabrican; no les interesa encontrar caminos hacia el verdadero conocimiento.

---

[76] Como explico en *Lógica a más fuerte* , capítulos 1.4 y 2.5, a veces resulta útil formular términos de manera tan inclusiva que abarquen valores positivos, cero y negativos. Esto se hace con frecuencia en el discurso científico porque facilita ciertos cálculos y gráficos. Sin embargo, es fundamental comprender que tales términos inclusivos son inherentemente innegables, es decir, ya engloban tanto una afirmación como su negación. En este contexto, podríamos optar por ampliar términos como «montón» o «calvo» para incluir sus opuestos, por la razón que sea; pero al hacerlo, debemos ser plenamente conscientes de lo que estamos haciendo. Si no tratamos estos términos con el debido cuidado, no deberíamos sorprendernos si nos vemos obligados a incurrir en contradicciones.

4. Inventar falsas paradojas

Las formulaciones originales de ambos enigmas descritos anteriormente, relativos a un montón y a la calvicie, se atribuyen a Eubúlides de Mileto (*activo* en el siglo IV a. C.), el lógico megarense que también nos legó la paradoja del mentiroso [77]. Fue discípulo de Euclides, maestro de Diodoro Cronos y contemporáneo y rival de Aristóteles. Estos enigmas quizás no se presentaron inicialmente como paradojas, sino más bien como ilustraciones de una pregunta (a saber, ¿dónde debemos trazar la línea?). Esto posiblemente reflejaba una creciente conciencia de que existen términos vagos que pueden requerir una definición arbitraria en ciertas circunstancias. Como se ha demostrado anteriormente, este problema se resuelve con bastante facilidad. Sin embargo, pensadores posteriores intentaron exagerar la situación y presentaron el tema en forma de cadena argumentativa (o sorites, donde la conclusión del argumento precedente sirve de premisa para el siguiente).

Así, el enigma del hombre calvo se convirtió, en su formulación positiva: Sin duda, un hombre con un solo cabello es tan calvo como uno sin cabello; y si un hombre con un solo cabello puede ser considerado calvo, entonces un hombre con dos cabellos también lo es; y si un hombre con dos cabellos es calvo, entonces un hombre con tres cabellos también lo es; etc.; por lo tanto, un hombre con mil cabellos aún puede ser considerado calvo (paradoja).

---

77 Un par de siglos antes, Epiménides de Cnosos declaró: «Los cretenses, siempre mentirosos», a pesar de ser él mismo cretense, aparentemente sin percatarse de la contradicción inherente a tal afirmación. Eubúlides pudo haber notado la paradoja y haber intentado refinarla y reforzarla, ya que no se trataba de una doble paradoja, sino de una que se resolvía fácilmente diciendo que posiblemente no todos los cretenses son mentirosos o que no todos los cretenses mienten siempre (siendo Epiménides una notable excepción).

Alternativamente, el argumento podría plantearse en forma negativa: si un hombre con mil cabellos no debe ser considerado calvo, entonces uno con 999 cabellos tampoco; y si 999 no cumple con el requisito, entonces 998 tampoco; y así sucesivamente… de donde, un hombre con un solo cabello no es calvo; por lo tanto, un hombre sin cabello no es calvo (paradoja).

Evidentemente, estos argumentos son *forzados* : implican premisas muy dudosas y engañosas. En la versión positiva, la premisa falsa es que «un hombre sin pelo puede ser llamado calvo» *implica* que «un hombre con un pelo puede ser llamado calvo». En la versión negativa, la premisa falsa es que «un hombre con un pelo no puede ser llamado calvo» *implica* que «un hombre sin pelo no puede ser llamado calvo». Se trata de la misma afirmación falsa de implicación (en forma contrapuesta). La paradoja se crea al negarse a admitir que «calvo», en sentido estricto, significa «sin pelo»; una negativa que no se basa en una comprensión lógica honesta, sino en un acto deliberado de ilógica. De manera similar, en el caso de un montón, el truco consiste en implicar que una sola piedra, o incluso ninguna, puede considerarse un montón.

Los creadores de estos argumentos no se detienen a reflexionar racionalmente sobre el problema subyacente (es decir, ¿dónde está la línea divisoria?) antes de iniciar un aparente proceso de inferencia, sino que intentan engañarnos y acorralarnos en una paradoja. La cadena de argumentos propuesta sirve simplemente como cortina de humo para ocultar la afirmación falsa crucial que se está planteando. Son el equivalente lógico de las estafas piramidales, donde cada venta sustenta la siguiente sin una base sólida. Quienes caen en la trampa son simplemente malos lógicos, si no descaradamente deshonestos. Luego, con pretensión, tejen teorías enormes y complejas en torno a este fenómeno, sin inmutarse por el error o la mentira inicial en su discurso.

A quienes argumentan que un solo cabello o una piedrecita apenas influyen, les sugiero que realicen el siguiente experimento físico sencillo: tomen una balanza precisa con el mismo peso en ambos lados, añadan un solo cabello o una piedrecita a un lado y ¡observen cómo se inclina! A quienes aún no se convenzan, porque creen dogmáticamente que la lógica es una cuestión de fantasía y convención, les sugiero (en tono de broma) que coloquen, debajo de la balanza más pesada, un detonador conectado a un artefacto explosivo sujeto a su nariz, ¡y que vean cómo la realidad les explota en la cara! Creo que ese argumento podría convencerlos finalmente, si sobreviven.

En cuanto a Eubúlides, cabe mencionar de paso sus otras paradojas. La más significativa es, sin duda, la paradoja del mentiroso, que, como he explicado detalladamente en otro lugar, [78]resulta excepcionalmente poderosa debido a la variedad de dificultades que plantea (aunque sigue siendo perfectamente resoluble). Otras tres paradojas [79]tratan sobre ambigüedades en el término «saber», concretamente sobre la incapacidad de reconocer inmediatamente a alguien a quien normalmente se reconoce al instante (como un pariente cercano o un viejo amigo), cuando este último lleva máscara, ha estado ausente demasiado tiempo o no se le observa con suficiente atención. Otra, la paradoja de los cuernos, afirma que lo que no se ha perdido debe estar en posesión; por lo tanto, si no se han perdido los cuernos, se deben tener. Esta apódosis implica una premisa mayor falsa, puesto que algo que uno «nunca ha tenido» puede caracterizarse igualmente (así como algo que «todavía tiene») como «no perdido»; por consiguiente, el

---

78 Véase mi libro *A Fortiori Logic* , Apéndice 7.4.

79 Es decir, las paradojas del Hombre Enmascarado, Elektra y el Hombre Ignorado.

consecuente no se deriva del antecedente; por lo tanto, si uno no ha perdido los cuernos, no se puede asumir que los tenga (ya que bien podría no haberlos tenido nunca).

De esta breve lista, se puede observar que las preguntas de Eubúlides suscitan reflexiones sobre la lógica: reflexiones sobre términos vagos, definiciones convencionales, ambigüedades, negaciones, autorreferencias y otras dificultades propias del discurso humano. Sería erróneo, a mi parecer, suponer que el motivo de tales preguntas fuera la provocación o la ofuscación (aunque, ciertamente, escépticos posteriores utilizaron estos dilemas con mala intención, como ya se ha mencionado). Más bien, diría que sirvieron como punto de partida para reflexiones y debates serios sobre la lógica, pues parece improbable que se formularan sin intentar resolver los problemas que planteaban. En algunos casos, sin duda se propusieron explicaciones o soluciones válidas (aunque tal vez no hayan llegado hasta nosotros), mientras que en otros las dificultades pudieron parecer insuperables. En otras palabras, dudo que Eubúlides fuera simplemente un sofista; lo clasificaría más bien como un lógico serio.

13. Protágoras contra Euatlo

De Temas (publicados solo en línea hasta el momento).

1. Una antigua paradoja

Una de las muchas supuestas paradojas que nos han llegado de los griegos es la disputa entre Protágoras (de Abdera, ca. 480-410 a. C.) y su alumno Euatlus (de quien no se sabe nada más). La historia, narrada por Aulo Gelio (romano, ca. 125-180 d. C.) [80], cuenta que Protágoras, un famoso sofista y un maestro costoso, acordó con Euatlus instruirlo en retórica, una disciplina esencial en aquella época para argumentar en los tribunales. El acuerdo consistía en que Euatlus no tendría que pagar a Protágoras la tarifa estipulada (o la parte impaga, según algunas versiones) hasta que estuviera completamente instruido y hubiera defendido su primer caso y lo hubiera ganado [81].

Se dice que, tras completar su curso, Euatlus no optó (por la razón que fuera) por utilizar sus nuevas habilidades ante ningún tribunal, por lo que nunca ganó ni perdió ningún caso y, en consecuencia, no estaba obligado contractualmente a pagar nada a Protágoras. Sin embargo, Protágoras, con motivos que analizaremos más adelante, lo demandó (ante el Areópago de Atenas). Euatlus decidió

---

80 Unos 550-600 años después, en sus *Noches del Ático* (ca. 150 C.E.).

81 Es de suponer que el contrato en cuestión era verbal, y no escrito, en aquella época; pero podemos asumir que ambas partes estuvieron de acuerdo con la cláusula estipulada.

defenderse personalmente. Según se cuenta, ante el tribunal se presentaron los siguientes argumentos:

a. **Protágoras** argumentó que seguramente se le debía pagar la tarifa, porque (i) si el tribunal fallaba a su favor, podría exigir el pago sobre esa base; y (ii) si el tribunal fallaba a favor del demandado, entonces este último habría ganado su primer caso y, por lo tanto, estaría obligado contractualmente a pagar la tarifa de todos modos.
b. **Euathlus respondió que seguramente** *no* debía pagar la tarifa, porque (i) si el tribunal fallaba a favor del demandante, entonces Euathlus habría perdido su primer caso y, por lo tanto, no estaría obligado contractualmente a pagar la tarifa; y (ii) si el tribunal fallaba a favor del demandado, entonces, sobre esa base, estaría exento de pago de todos modos.

Así, mientras el demandante argumentaba, aparentemente de forma convincente, que tenía derecho a una indemnización independientemente del fallo del tribunal, el demandado, a su vez, argumentaba, aparentemente con igual convicción, que estaba exento de pago en cualquier caso. Por esta razón, este caso se considera paradójico. Se dice que el tribunal quedó tan desconcertado por estos argumentos y contraargumentos que optó por aplazar *la audiencia indefinidamente* (o, según algunos, durante 100 años) para evitar dictar sentencia.

La importancia de esta disputa legal para la lógica y la filosofía radica en que da la impresión de que dos personas pueden argumentar de forma dilemática y paradójica en direcciones opuestas y ambas tener razón. Los enemigos de la razón humana se regocijan con este tipo de conclusión, ya que pondría en duda la fiabilidad y la eficacia de la razón. Pero, como demostraremos a continuación, dicha impresión es muy superficial. En mi opinión, existen al menos dos posibles soluciones a esta

llamada Paradoja de la Corte. De hecho, aunque pensé en la segunda antes que en la primera, esta última precede lógicamente a la primera. Posteriormente, gracias al estudio de la bibliografía sobre el tema realizado por Peter Suber, supe [82]que la primera solución ya había sido propuesta mucho antes por Aulo Gelio; la segunda solución parece ser novedosa.

2. Primera resolución

La solución más sencilla a este problema es suponer que el astuto Protágoras, al ver que Euatlus se demoraba en ponerse a trabajar, decidió acelerar el proceso. Protágoras atrapó a su discípulo con el argumento anterior, sabiendo perfectamente que perdería el primer juicio, pero ganaría el segundo. Sabía que perdería el primer juicio porque el acuerdo entre las partes solo obligaba a Euatlus a pagar los honorarios una vez que ganara su primer caso; no lo obligaba a ejercer la abogacía en un futuro próximo, ni siquiera nunca. Euatlus, ingenuamente, cayó en la trampa y defendió personalmente el caso ante el tribunal. Si el tribunal no hubiera suspendido *la sesión indefinidamente* , lógicamente le habría dado la victoria, lo que le habría permitido ganar su primer caso. Posteriormente, suponiendo que un segundo juicio fuera legalmente permisible —y los argumentos de ambas partes parten de esta premisa—, Protágoras habría sido el vencedor final. Por supuesto, no sería necesario un segundo juicio si Euathlus admitiera que, habiendo ganado el primero,

---

[82] En *La paradoja de la auto-reparación* . La sección sobre la paradoja de Protágoras vs. Euathlus se puede leer en línea; vale la pena leerla, incluyendo las notas, y no es muy larga. Véase en: https://dash.harvard.edu/bitstream/handle/1/10288413/Peter%20Suber%2c%20Paradox%20of%20Self-Amendment%2c%20Section%2020.html?sequence=1#A .

estaba seguro de perder el segundo, y saldara la cuenta de inmediato.

En otras palabras, el primer argumento de Protágoras (i) era mero camuflaje; en realidad, se basaba en su segundo argumento (ii). Euatlus, cegado por la vanidad, formuló dos ingeniosos contraargumentos, pensando en superar a su maestro. Pero Protágoras era más astuto. La única manera en que Euatlus podía evitar la derrota era contratando a un abogado [83]. Si el abogado ganaba el primer juicio, como era lógico esperar, ya que no se cumplía la única condición para la derrota, Euatlus no sería considerado ganador de su primer caso ante el tribunal, y Protágoras no podría ganar un segundo juicio. Por lo tanto, el maestro era, en efecto, superior al alumno. No existía una verdadera paradoja, puesto que había una solución real a la aparente paradoja.

Por supuesto, cabría añadir que Euathlus fue, en la práctica, el vencedor, puesto que con su contraargumento logró desconcertar tanto al tribunal que este desistió de juzgar el asunto, y no se vio obligado a pagar. Quizás lo esperaba y tuvo suerte. Pero desde un punto de vista teórico y lógico, desde la perspectiva actual, demostró no ser muy inteligente. No solo cometió la imprudencia de no contratar a un abogado para que lo defendiera, sino que además asumió erróneamente que su argumentación sería eficaz para refutar la de Protágoras. Se engañó a sí mismo pensando que, si ganaba el primer juicio, no tendría que enfrentarse a un segundo. Debería haber examinado con más detenimiento la argumentación de su maestro.

---

83 Suponiendo que esto fuera posible en Atenas en aquella época. Esta es una suposición razonable (aunque algunos comentaristas la nieguen), ya que, al parecer, Euathlus fue instruido por Protágoras para ser abogado.

Analicemos ahora los argumentos en términos más formales para clarificarlos. Introduciré los siguientes símbolos: sea P = Protágoras, Q = Euathlus; A = la condición acordada de que Q gane su primer caso para ser responsable, y C = el resultado acordado de que Q deba pagar honorarios a P. Sabemos desde el principio que si A entonces C, y si no A entonces no C: estos son los términos del acuerdo. Los argumentos son los siguientes (con mi comentario crítico en *cursiva*):

a. Según P: (i) si el tribunal dictamina que P gana, entonces C es verdadero (*pero objetivamente el tribunal no puede dictaminar a favor de P, ya que A aún no es verdadero, por lo que esto no tiene fundamento*); mientras que si (ii) el tribunal dictamina que Q gana (*como lógicamente debe ser*), entonces no-C es verdadero (*al concluir un primer juicio*); pero si Q gana, entonces A se vuelve verdadero y C debe seguirse (*en un juicio posterior*), en cuyo caso P finalmente gana.
b. Según Q: (i) si el tribunal dictamina que P gana, entonces no-A es verdadero (*al concluir un primer juicio*); pero si no-A es verdadero, entonces no-C debe seguirse (*en un juicio posterior*), por lo que Q finalmente gana (*pero objetivamente el tribunal no puede dictaminar a favor de P, ya que A aún no es verdadero, por lo que esto no tiene fundamento*); mientras que si (ii) el tribunal dictamina que Q gana (*como lógicamente debe ser*), entonces no-C es verdadero (*pero aquí Q no menciona las consecuencias posteriores que P señaló correctamente*).

En conclusión, lógicamente P es el ganador; los argumentos de Q son, de hecho, insuficientes para impedir este resultado. P pretende ganar de inmediato, pero en realidad su objetivo es la victoria a largo plazo (en la segunda ronda). Q imagina que podría perder la

primera ronda pero ganar la segunda, o que podría ganar la primera sin tener que enfrentarse a una segunda; pero todo esto son fantasías. Resulta difícil comprender por qué el tribunal consideró este caso demasiado confuso; el juez (o los jueces) no debían de ser muy brillantes.

3. Segunda resolución

Una solución más compleja al problema es la siguiente. Es posible que Protágoras demandara a Euatlus apelando a *una cláusula tácita* del acuerdo. El acuerdo contenía *solo una cláusula explícita* , a saber, que Euatlus tendría que pagar a Protágoras los honorarios si y solo si ganaba su primer caso. Si eso fuera así, Protágoras no tendría fundamento para solicitar un juicio, ya que esa condición obviamente no se había cumplido. Pero como demandó, debió haber pensado y argumentado que el acuerdo incluía una *cláusula tácita. (o quizás implícito)* Protágoras *comprendió* que Euathlus ejercería la abogacía en un plazo razonable [84], tras el cual sus nuevas habilidades serían puestas a prueba en un tribunal y se esperaría que pagara si ganaba. Protágoras no podía imaginar que el juez permitiría que el juicio siguiera adelante, y mucho menos que fallaría a su favor, sin una buena razón [85].

Evidentemente, lo que el juez debía decidir en este juicio era (y, desde la perspectiva actual, solo podía ser) si la alegación de Protágoras sobre la existencia de una cláusula tácita en el acuerdo estaba justificada. Podría

---

[84] Protágoras también alegaría, por supuesto, que el plazo razonable había expirado. Si el tribunal coincidiera en la existencia de una cláusula tácita pero no aceptara que se hubiera cumplido, simplemente aplazaría el caso durante un tiempo determinado (según su estimación).

[85] Como es lógico, el presente análisis se realiza exclusivamente desde un punto de vista lógico, aunque un tribunal de justicia podría razonar de manera diferente o incluso no razonar en absoluto (por ejemplo, soborno, favoritismo, presión).

haberla justificado considerando que, si a Euatlus se le hubiera permitido no ejercer la abogacía o ejercerla hasta el final de su vida, el acuerdo seguramente habría especificado dicha condición. Igualmente, podría haberla rechazado considerando que Protágoras daba por sentado algo sobre lo que debería haber obtenido un acuerdo explícito. Por lo tanto, el caso gira en torno a una cuestión tácita, más que exclusivamente a la cláusula explícita del acuerdo; es decir, la historia era más compleja de lo que se cuenta.

Además, de los argumentos presentados por ambas partes en este juicio se desprende que cada una previó la posibilidad de un segundo juicio en el que se pudiera revocar la sentencia del primero. Esto se deduce lógicamente del segundo argumento de Protágoras y del primero de Euatlus. En ambos casos, sería necesario un segundo juicio para finalizar la sentencia, a menos que, por supuesto, la parte perdedora aceptara libremente el resultado inevitable y saldara la cuenta por adelantado. El juez del primer juicio no podía fallar a favor de una parte y luego en su contra simultáneamente. Un primer veredicto de victoria o derrota debía establecerse antes de que se pudiera dictar un segundo, en vista de los términos del contrato. Por lo tanto, existe un factor temporal que debe tenerse en cuenta al analizar estos argumentos.

Cabe señalar que, mientras que el primer juicio tiene un resultado incierto, puesto que depende de la decisión (el criterio) del juez respecto a una supuesta cláusula tácita del acuerdo, el segundo juicio tiene un resultado previsible, puesto que depende únicamente de la cláusula explícita del acuerdo.

Claramente, como demostraremos a continuación, si incorporamos las consideraciones anteriores a los argumentos, la apariencia paradójica se disuelve fácilmente. Utilizaré aquí los siguientes símbolos: P =

Protágoras; Q = Euatlo; A = la condición explícita de que Q gane su primer caso; B = la supuesta condición implícita de que Q debía ejercer en un plazo razonable; y C = los honorarios que Q debe pagar a P. Los argumentos son los siguientes:

a. Según P: (i) si B entonces C ; y (ii) si no B entonces no C, *pero si no C entonces (más adelante) si A entonces C* ; por lo tanto, C de todos modos.
b. Según Q: (i) si B, entonces C, *pero si C entonces (más adelante) si no-A entonces no-C* ; y (ii) si no-B, entonces no-C; por lo tanto, no-C de todos modos.

De esto se desprende que los argumentos de ambas partes son muy similares, pero cada una ha omitido astutamente parte de las consecuencias (mostradas en *cursiva*). P ha truncado las consecuencias de (i) que Q señala, y Q ha truncado las consecuencias de (ii) que P señala. Por esta razón, parecen compensarse mutuamente. Pero si tomamos en cuenta *todos* los eventos subsiguientes (en un posible segundo juicio), obtenemos el siguiente argumento conjunto más objetivo (c), que claramente no es paradójico:

(i) si B entonces C, pero si C entonces (más adelante) si no-A entonces no-C ; y
(ii) si no-B entonces no-C, pero si no-C entonces (más adelante) si A entonces C .

Así pues, en (i) la conclusión final no es C, mientras que en (ii) sí lo es, lo que significa que no hay paradoja. Esto también demuestra que, si bien no se puede predecir objetivamente si P o Q ganarán el primer juicio, se puede afirmar que (i) si P gana el primer juicio, perderá el segundo, y si Q gana el primero, perderá el segundo. Obviamente, P no puede argumentar que tiene derecho a un segundo juicio (si pierde el primero) pero Q no; del mismo modo, Q no puede argumentar que tiene derecho a un segundo juicio (si pierde el primero) pero P no. Por lo

tanto, debemos tener en cuenta todos los acontecimientos posteriores para conciliar lógicamente todos los argumentos. Cabe señalar que si por alguna razón no hay un segundo juicio, tampoco hay paradoja, puesto que la conclusión será (i) C o (ii) no C.

Pase lo que pase, no hay paradoja alguna, porque ninguna de las partes puede, de hecho, contrariamente a las apariencias iniciales, reclamar una victoria inevitable; la victoria no llega de ambas maneras para ambas partes, sino solo de una manera para cada una.

Evidentemente, en este caso, ambas partes recurrían al truco sofístico habitual de ocultar al tribunal una parte inconveniente de su argumento imparcial. Euathlus era un buen discípulo de Protágoras, puesto que su contraargumento reflejaba con exactitud el de este último. Es decir, existía un elemento de deshonestidad en ambos argumentos; ambos eran, en el fondo, impostores intelectuales que, a sabiendas, exponían medias verdades. Por lo tanto, esta falsa paradoja no supone un desafío profundo para la lógica, contrariamente a lo que afirman los relativistas. En particular, la afirmación general de Protágoras de que «hay dos caras de la moneda» (*duo logoi*) se demuestra falsa en este contexto.

Conviene tener siempre presente que algunas personas dedicadas a la filosofía y la lógica, como en la vida en general, a veces se dejan llevar por impulsos perversos; de hecho, algunas mucho más que otras. Pueden mentir conscientemente o distorsionar inconscientemente hechos y argumentos por una gran variedad de motivos. Generalmente, la sed de poder, fama y fortuna influye en ello. Un académico puede desear la admiración de sus colegas o estudiantes; un esposo puede querer impresionar a su esposa; un desempleado puede aspirar a encontrar trabajo; y así sucesivamente. Es erróneo considerar todas las afirmaciones filosóficas como

imparciales. No todos los filósofos y lógicos son científicos puros ni santos.

Cabe añadir que el secreto del éxito al encontrar soluciones a problemas filosóficos y lógicos, y en particular a las paradojas, reside en el sincero deseo de hacerlo. Muchos filósofos y lógicos abordan los problemas con una actitud negativa, sin el verdadero propósito de resolverlos, sino más bien buscando racionalizar su aversión a la razón humana a través de ellos. El investigador honesto se guía por sus mejores instintos; desea sinceramente confirmar la eficacia de las facultades cognitivas humanas, y precisamente por eso tiene éxito.

4. Resoluciones inadecuadas

Como ya se mencionó, según la revisión de la literatura realizada por Suber, la primera resolución debe atribuirse a Aulo Gelio, pero la segunda parece ser original. El análisis de Suber muestra que la paradoja de la corte se ha tratado en diversas obras a lo largo del tiempo, pero aparentemente con mayor frecuencia desde una perspectiva jurídica que lógica. Las cuestiones legales involucradas son múltiples, pero la mayoría no necesitan ni deben considerarse desde una perspectiva puramente lógica. ¿Por qué? Porque el propósito del lógico en este caso no es decidir el caso, es decir, quién debe ganar o perder, sino simplemente explicar y eliminar la apariencia de paradoja.

Esta observación puede ilustrarse con la resolución (no mencionada por Suber) propuesta siglos atrás por Lorenzo Valla (Italia, ca. 1406-57). Este intento de resolución resulta insuficiente, ya que se basa en una distinción irreflexiva entre el pago derivado del veredicto judicial y

el pago derivado del acuerdo. Cito un artículo de la SEP que lo describe [86]:

> Si Euathlus pierde el caso, deberá pagar el resto de los honorarios, conforme al veredicto de los jueces; pero si gana, también deberá pagar, esta vez en virtud de su acuerdo con Protágoras. Sin embargo, Euathlus manipula hábilmente el argumento: en ningún caso tendrá que pagar, ni por la decisión del tribunal (si gana) ni por el acuerdo con Protágoras (si pierde). En resumen, Valla afirma que Euathlus no puede tenerlo todo y debe elegir una de las dos alternativas: cumplir con su acuerdo con Protágoras o con el veredicto de los jueces. Si fallan en contra de Protágoras, podrá intentar recuperar su dinero en un segundo juicio.

Analicemos la resolución propuesta por Valla utilizando los siguientes símbolos: sea P = Protágoras; Q = Euathlus; y C = la tarifa que Q debe pagar a P.

a. Según P: Si Q pierde (por cualquier motivo), entonces C (por veredicto); pero si Q gana, entonces C (por acuerdo).
b. Según Q: Si Q gana (por cualquier motivo), entonces no-C (por veredicto); pero si Q pierde, entonces no-C (por acuerdo).

---

[86] Véase *Stanford Enc. of Phil.* , disponible en línea en https://plato.stanford.edu/entries/lorenzo-valla/. Al parecer, este tema se trata en *Repastinatio dialéctica et philosophie.* El autor del artículo es aparentemente Lodi Nauta (2013).

La conclusión de Valla, tal como se presenta aquí [87], resulta ambigua. Aparentemente, le impone a Euatlus, en particular, la responsabilidad de «elegir entre una u otra alternativa» y acatar el acuerdo o el veredicto. Esto no nos aclara nada, puesto que el argumento de Euatlus demuestra su disposición a cumplir con cualquiera de las dos, salvo que las presenta como favorables a él. Valla añade que si Protágoras pierde un primer pleito, podría ganar un segundo. Pero, una vez más, esto no resuelve la paradoja, sino que solo repite una parte de ella.

De esto se desprende que Valla no ha analizado a fondo la cuestión: no considera en qué circunstancias Q podría perder o ganar «por el veredicto del tribunal»; y no se da cuenta de que el segundo argumento de cada parte, donde Q podría ganar o perder alternativamente «por el acuerdo», en realidad se refiere a un segundo litigio, puesto que el acuerdo evidentemente no se implementará voluntariamente (la mención que hace Valla de un segundo litigio se sitúa más allá de las cuatro proposiciones condicionales, que no enfatizan las secuencias temporales implicadas). En definitiva, Valla no llega a una solución fácil de la paradoja.

El problema con el enfoque de Valla es que, en la práctica, considera arbitraria la decisión inicial del tribunal, es decir, ajena al acuerdo entre las partes. No analiza en qué se basa el tribunal para juzgar que Euathlus

---

87 No he leído la obra de Valla. Debo suponer que el autor del artículo de la *SEP* sobre Valla presentó correctamente su razonamiento. Sospecho que no lo hizo, ya que los detalles que se ofrecen en el artículo son bastante vagos e inconclusos; no aclaran con exactitud qué resolución de la paradoja tenía Valla en mente. No todo el mundo domina la lógica. (Cabe mencionar que escribí al autor pidiéndole que verificara el resumen de la postura de Valla que aparece en el artículo de la SEP, pero no obtuve respuesta. El hecho de que el autor no se dignara a responder confirmó mi impresión de que no se trata de una fuente muy fiable: las personas inteligentes suelen recibir con agrado las consultas razonables).

puede perder la primera ronda. Esto puede ser aceptable legalmente, pero no lo es lógicamente, y aquí nos interesa la lógica. Lógicamente, la idea de que Euathlus pueda perder la primera ronda es insostenible, si nos atenemos a la cláusula explícita del acuerdo. Solo podría perder la primera ronda si el tribunal acepta la suposición de que existía una cláusula tácita en el acuerdo, como por ejemplo que debía ejercer la abogacía próximamente.

En cualquier caso, el veredicto de la primera ronda no puede ser arbitrario —es decir, irracional o injustificado—, sino que debe estar directamente relacionado con el acuerdo. Lo mismo se aplica a la segunda ronda. Por lo tanto, no se permite lógicamente ninguna discrepancia entre el veredicto y el acuerdo.

Curiosamente, mi primera reacción ante la paradoja hace unos años [88]fue muy similar a la de Valla. Pero tan pronto como me dispuse a considerar seriamente los temas para el presente ensayo, me di cuenta de su debilidad e inutilidad. La paradoja parece neutralizarse si introducimos una distinción entre "pago tras el veredicto judicial" y "pago según los términos del contrato" y asignamos símbolos diferentes a estas dos consecuencias, digamos C1 y C2. Estos dos términos pueden o no ser cuantitativamente idénticos; pero en cualquier caso se refieren a eventos distintos. Esta medida parece anular la paradoja, porque las consecuencias de las dos primeras proposiciones condicionales serían diferentes para cada parte. Por lo tanto:

a. Según P: Si P gana, entonces C1; pero si P pierde, entonces C2.
b. Según Q: Si Q gana, entonces no-C1; pero si Q pierde, entonces no-C2.

De ello se deduce que si P gana y Q pierde, entonces C1 y no-C2 son verdaderas; y si P pierde y Q gana, entonces

88 En 2009, en un correo electrónico a alguien.

no-C1 y C2 son verdaderas. Dado que C1 y C2 son términos distintos, incluso si se refieren a la misma cantidad monetaria, las consecuencias en el argumento de cada parte no son formalmente contradictorias (ya que el motivo definitorio es diferente), por lo que no hay paradoja. Pero, reiteramos, este enfoque no explica por qué el tribunal emitiría un veredicto incompatible con el acuerdo, por lo que resulta artificial. Para una resolución verdaderamente lógica, debemos centrarnos en el acuerdo y no admitir veredictos arbitrarios. Los veredictos arbitrarios solo enturbian las aguas.

Otras soluciones que he visto también son defectuosas, ya sea porque confunden cuestiones legales y lógicas, o porque no consideran todas las consecuencias de todas las suposiciones. En cualquier caso, cualquiera de mis dos soluciones propuestas basta para resolver la paradoja; pero ambas juntas abarcan todas las eventualidades. Cabe destacar que, en la primera solución, no solo resolvemos la paradoja, sino que también, incidentalmente, decidimos el caso (a favor de Protágoras); mientras que en la segunda, como lógicos, nos limitamos a resolver la paradoja, dejando la decisión legal al juez.

En conclusión: la paradoja del tribunal se debe a varios factores que deben desentrañar y tenerse en cuenta para su resolución. (a) Los términos «ganar/perder» no pueden referirse a juicios arbitrarios del tribunal; de ser así, la paradoja podría perpetuarse. (b) Debe entenderse que los argumentos presentados por ambas partes implican que el proceso de resolución tiene dos fases: un primer juicio seguido de un segundo juicio (o una concesión y un acuerdo inmediatos sin necesidad de un segundo juicio); la apariencia de paradoja se debe en parte a que se pasa por alto este factor temporal. (c) Existe cierta ambigüedad en el significado de «el acuerdo», y deben considerarse todas las interpretaciones posibles: si se entiende que el acuerdo se refiere exclusivamente a la cláusula explícita,

entonces el primer juicio se centra únicamente en ella y se decide fácilmente; pero si se alega que el acuerdo implica un entendimiento tácito, entonces el primer juicio apunta a una decisión sobre esa cláusula tácita y la segunda ronda trata sobre la cláusula explícita.

14. discurso antinómico budista

De Lógica Budista 1 y 2.

1. El tetralema

El pensamiento filosófico y científico occidental se basa en la lógica aristotélica, cuyos principios fundamentales son las tres «Leyes del Pensamiento». Estas se pueden resumir como «A es A» (Identidad), «Nada es A y no-A a la vez» (No contradicción) y «Nada es ni A ni no-A» (Exclusión del término medio). Cabe destacar que estas leyes no se presentan como meras hipótesis, sino como premisas incontrovertibles de todo pensamiento racional humano [89].

Religiones como el judaísmo, el cristianismo y el islam, aun cuando se adhieren a estas leyes en gran parte de su discurso y las respetan superficialmente, en sus intentos por interpretar sus textos sagrados y hacer que sus doctrinas parezcan razonables para sus conversos, a menudo las han ignorado. Esto es especialmente cierto en el caso de las corrientes místicas dentro de estas religiones, pero se podrían citar muchos ejemplos en textos convencionales. Lo mismo puede decirse de algunos aspectos de la filosofía budista.

---

89 Consulte mi *libro Lógica del futuro* , capítulos 2 y 20, y ensayos posteriores sobre el tema (publicados en mi sitio web www.thelogician.net).

El *tetralema* [90]es una derivación de las leyes del pensamiento, con referencia a dos términos o proposiciones cualesquiera, denominados A y B, y sus opuestos no-A y no-B. Son concebibles cuatro combinaciones de estos cuatro términos: «A y B» (ambos), «no-A y no-B» (ninguno), «A y no-B» y «no-A y B» (uno u otro). Según la lógica aristotélica, estas cuatro proposiciones son incompatibles entre sí (*solo una de ellas puede ser verdadera* , ya que si se afirmaran dos o más, entonces «A y no-A» o «B y no-B» o ambas serían verdaderas, y estas últimas implicaciones son contradictorias) y exhaustivas (*al menos una de ellas debe ser verdadera* , ya que si se negaran todas, entonces «no A y no no-A» o «no B y no no-B» o ambas serían verdaderas, y estas últimas implicaciones contradicen el principio del tercero excluido).

Ahora bien, lo que hace Nagarjuna es sustituir B por el término A (es decir, considera el caso B = A) y afirmar que las cuatro posibilidades lógicas de combinación mencionadas anteriormente se aplican en ese caso particular, de modo que «A y A (=B)», «no-A y no-A (=no-B)», «A y no-A (=no-B)» y «no-A y A (=B)» parecen lógicamente aceptables. A continuación, argumenta que existen cuatro posibilidades existenciales: **afirmación de A (A + A = A), negación de A (no-A + no-A = no-A), afirmación y negación de A (A y no-A) y ni afirmación ni negación de A (ni A ni no-A)** . De este modo, aparentemente utiliza los principios y la terminología de la lógica común para llegar a un resultado completamente opuesto. Esto le da a él y a los lectores la impresión de que es perfectamente razonable afirmar y negar, o no afirmar ni negar.

---

[90] Véase Cheng, págs. 36-38, sobre este tema. Allí se refiere a la declaración inicial del Texto Masorético, así como a XVII:12a y XXIII:1a. Etimología: Gk. *tetra* = cuatro, *lemma* = alternativas. Término acuñado en contraste con el dilema «A o no-A».

En la lógica aristotélica, las dos últimas alternativas quedan excluidas desde el principio: «tanto A como no-A» por el principio de no contradicción y «ni A ni no-A» por el principio del tercero excluido. Las únicas posibilidades lógicas que quedan son «A» o «no-A». La postura antiaristotélica puede considerarse, desde una perspectiva positiva, como una postura antinominalista, que nos recuerda que las cosas nunca son exactamente lo que parecen o que no pueden clasificarse ni etiquetarse con precisión. Sin embargo, en última instancia, pretenden la muerte de la lógica; pues sin las leyes del pensamiento, ¿cómo distinguir entre juicios verdaderos y falsos?

La ley de identidad «A es A» es la convicción de que las cosas poseen alguna identidad (sea cual sea) en lugar de otra, o ninguna identidad en absoluto. Es una afirmación de que el conocimiento es, en última instancia, posible, y un rechazo al relativismo o al oscurantismo puros. El objetivo de Nagarjuna es negar la identidad.

Cabe señalar que Aristóteles es muy preciso en su formulación de la ley de la contradicción, afirmando en su *Metafísica :* «El mismo atributo no puede pertenecer y no pertenecer *al mismo sujeto en el mismo sentido* » (cursivas mías). Así, una formulación alternativa de las leyes del pensamiento sería el «trilema» (llamémoslo así): « *o totalmente A, o totalmente no-A, o en parte A y en parte no-A* », lo que excluye la cuarta alternativa: «tanto totalmente A como totalmente no-A». La crítica budista a las leyes del pensamiento se fundamenta en parte en el hecho de que las personas se refieren subconscientemente a este «trilema», pensando superficialmente que, en efecto, pueden ocurrir cosas opuestas en el mismo lugar en diferentes momentos, o al mismo tiempo en diferentes lugares o en diversos aspectos, sin que ello genere una dificultad lógica irresoluble. Sin embargo, debe quedar

claro que la postura budista es mucho más radical, pues acepta la antinomia absoluta.

De igual modo, con respecto a la ley del tercero excluido, que afirma que la situación «ni A ni no-A» es imposible *en la práctica* , la posibilidad de incertidumbre *en el conocimiento* , respecto a si A o no-A es el caso en realidad, lleva a creer erróneamente que esta ley del pensamiento es debatible. Sin embargo, es importante comprender que el fundamento de esta regla lógica es inductivo, no deductivo; es decir, afirma que *al final* del proceso de adquisición del conocimiento, se obtendrá «A» o «no-A», y no cabe esperar una tercera alternativa. Esto no excluye que, *mientras tanto* , pueda surgir una situación de incertidumbre. La postura de Nagarjuna se aprovecha de esta confusión.

Nagarjuna interpreta la limitación implícita en el dilema «A o no-A» como un «dualismo» arbitrario por parte de los pensadores comunes [91]. Esto solo demuestra que no

---

[91] Es engañoso llamar a esto una «dualidad» o «dicotomía», como suelen hacer los budistas, porque sugiere que una cosa unitaria fue arbitrariamente dividida en dos, e incidentalmente, que bien podría haber sido dividida en cuatro. Pero, a nivel perceptivo, no hay elección involucrada, ni «fragmentación» de nada. Un fenómeno que aparece es *una sola* cosa, llamémosla « **un** » (un nombre propio, o un indicativo «esto»), y no una disyunción. El problema de la «dicotomía» surge solo a nivel conceptual. *La negación* es un acto racional, es decir, solo podemos hablar de «no **-un** », de lo que no aparece, trayendo primero a la mente algo « **un** », que apareció previamente (en la sensación o la imaginación). En *la conceptualización inicial* , dos fenómenos se comparan y contrastan, entre sí y con otras cosas, en algún(os) aspecto(s); La cuestión es, entonces, si son lo suficientemente similares entre sí y lo suficientemente diferentes de otras cosas como para ser juzgadas como «iguales» y etiquetadas con un término general (por ejemplo, «A»), o si deben ser juzgadas como «diferentes» o existe incertidumbre. En *la etapa posterior del reconocimiento* , debemos decidir si un tercer fenómeno encaja en la clase formada para los dos anteriores (es decir, se incluye en «A») o no encaja (es decir, se

comprende la formalización (o finge no comprenderla, en un intento de confundir a los lectores crédulos). Cuando los lógicos usan una variable como «B» y admiten que «no-A y B» y «A y no-B» son posibles en principio, no pretenden que sea una generalidad aplicable a *todos* los valores de B (como «A»), sino solo una afirmación genérica aplicable a *cualquier* valor consistente de B. En el caso específico donde B = A, dichas dos combinaciones deben eliminarse porque son ilegales (es decir, infringen dos de las leyes del pensamiento).

La propiedad de los símbolos antes mencionada, es decir, su aplicabilidad solo condicional dentro de las restricciones de la consistencia, es evidente en toda la ciencia de la lógica formal, y aquí Nagarjuna la ignora por completo. Su motivación, por supuesto, era verbalizar y racionalizar la doctrina de Buda de que la verdad última está más allá de *nama* y *rupa* , nombre y forma (es decir, discriminación y discurso), cognoscible solo por una conciencia trascendental (la doctrina de la Doble Verdad). Más precisamente, como subraya Cheng, la intención de Nagarjuna era demostrar que la lógica es inherentemente inconsistente y, por lo tanto, que la razón es una locura confusa que debe rechazarse. Es decir, en última instancia no intentaba (aquí y en todo momento) defender un tetralema con B igual a A —ni siquiera afirmar que las cosas son A y no-A, o ni A ni no-A—, sino que deseaba que miráramos más allá de las distinciones de la conceptualización y los juicios de la lógica.

Pero, como se ha demostrado anteriormente, no tiene éxito en esta búsqueda. Su crítica se basa en una tergiversación de la ciencia lógica. Afirma demostrar que la lógica es confusa y contradictoria, pero en realidad lo

---

incluye en «no-A») o permanece en duda. En este último caso, nos preguntamos si es «A» o «no-A», y advertimos que no puede ser ambas cosas ni ninguna de las dos.

que presenta como la tesis de la ciencia lógica no es lo que esta afirma, sino precisamente lo que prohíbe explícitamente. Además, suponiendo que la teoría lógica condujera a contradicciones, como él afirma, este hecho no nos llevaría a rechazarla si no existiera también una apelación tácita a nuestra preferencia por lo lógico en la práctica. Si la lógica fuera falsa, las contradicciones serían aceptables. Así, curiosamente, Nagarjuna apela a nuestro hábito lógico en su propia recomendación de ignorar la lógica. En resumen, aunque da la ilusión de que es razonable abandonar la razón, es fácil ver que su conclusión es inevitable y sus métodos son erróneos.

2. Ni real ni irreal

Pero Nagarjuna también concibe la realidad última ("vacuidad" [92]) como un "camino intermedio". [93]—de modo que el mundo de la experiencia no debe considerarse ni real ni irreal («no hay nada, ni mental ni no mental, que sea real» y «no puede concebirse como irreal», según Cheng). En este contexto, Nagarjuna se basa claramente en una de las disyunciones lógicamente imposibles mencionadas anteriormente, a saber, «ni A ni no-A» (dicho de paso). Ahora quiero demostrar por qué la afirmación de Nagarjuna parece superficialmente razonable y verdadera.

Como he aclarado y explicado en repetidas ocasiones [94], el conocimiento debe considerarse o abordarse desde una

---

92 Más allá de la conciencia de “Shunyata” existe una percepción más vívida llamada “Mahamudra”, según Chögyam Trungpa en *El juego de la ilusión* . Pero tales sutilezas no nos conciernen aquí.

93 Véase Cheng, págs. 38-39, sobre este tema. Allí se refiere a MT XIII:9a y XVIII:7.

94 Véase mi *libro Lógica del futuro* , capítulos 60-62, y ensayos posteriores sobre el tema.

perspectiva fenomenológica (esta es la única tesis epistemológica coherente). Debemos comenzar reconociendo y observando *las apariencias* como tales, sin juzgarlas inicialmente como realidad o ilusión. A primera vista, todas las apariencias parecen bastante *reales* . Pero, después de un tiempo, debemos reconocer que algunas apariencias entran en conflicto con otras y juzgarlas (es decir, una o más de las que entran en conflicto) como *ilusorias* . Dado que en nuestro «mundo» no hay nada más que apariencias, todas las demás apariencias que no se consideran ilusiones (es decir, siempre que no sean invalidadas lógicamente por conflictos con otras apariencias) mantienen su estatus inicial de realidades.

Es decir, la distinción entre apariencias como realidades o ilusiones surge dentro del mundo de las apariencias mismo, simplemente clasificando algunas de una manera y otras de otra. No tenemos un concepto de realidad o ilusión más allá de la apariencia. Usar la categoría de realidad para referirnos a algo *más allá de* la apariencia es un robo de concepto, un mal uso del concepto, una extrapolación que ignora su génesis real en el contexto de la apariencia. Aplicar el concepto de ilusión a *todas* las apariencias, partiendo de la base de que algunas son ilusiones, es una generalización injustificada que ignora cómo surge este concepto en relación con un evento específico (a saber, la inconsistencia entre ciertas apariencias y la consiguiente disminución de su credibilidad intrínseca). Además, afirmar que ninguna apariencia es real o que todas son ilusiones es contraproducente, ya que dicha afirmación lógicamente se enmarca dentro de la categoría de apariencia.

Lo ilusorio existe aunque no sea la realidad; existe como apariencia. Lo real también es aparente, al menos en parte. Por lo tanto, la apariencia en sí misma no debe entenderse como realidad (ya que algunas apariencias son

ilusorias), ni puede equipararse a la ilusión (ya que no todas las apariencias han sido ni pueden ser ilusorias). La apariencia es, pues, el *terreno común* de realidades e ilusiones, su característica común, la síntesis dialéctica de esas tesis y antítesis. Es un género; son especies mutuamente excluyentes de ella. (La diferencia entre apariencia y existencia es otro tema que he tratado en otro lugar; en resumen, la existencia es un género de apariencia y no apariencia, siendo estos últimos conceptos relativos al de la conciencia, mientras que la primera se presupone independiente).

Ninguna de estas ideas permite concluir que las apariencias no son «ni reales ni irreales» (siempre que se entienda «irreal» como «no real»). Lo único que podemos afirmar es que algunas apariencias son reales y otras irreales. Formalmente, la relación lógica correcta entre los tres conceptos es la siguiente. *Deductivamente* , la apariencia está implícita en la realidad y la ilusión, pero no las implica; pues la realidad y la ilusión son contradictorias, de modo que no pueden ser ambas verdaderas ni ambas falsas. Además, *inductivamente* , la apariencia implica la realidad, a menos que se la considere una ilusión (debido al descubrimiento de alguna inconsistencia).

Más precisamente, todas las apariencias se clasifican inicialmente como reales. Cualquier apariencia que se encuentre contradictoria es (deductivamente) ilusoria, y su contradicción es, por consiguiente, evidente y (deductivamente) real. Todas las demás apariencias permanecen clasificadas como reales, siempre que no sean cuestionadas. Aquellas que son cuestionadas deben evaluarse dinámicamente. Cuando una apariencia es desmentida por otra, ambas se ponen en duda debido al conflicto entre ellas, y por lo tanto ambas se vuelven inicialmente *problemáticas* . Posteriormente, sus credibilidades relativas deben sopesarse tentativamente en

el contexto general del conocimiento empírico y racional disponible, y reevaluarse repetidamente a medida que dicho contexto se desarrolla y evoluciona. Sobre esta base, una de estas apariencias puede considerarse más creíble que la otra, de modo que la primera se etiqueta como *probable* (cercana a la realidad) y la segunda como relativamente *improbable* (cercana a la ilusoria). En el límite, pueden caracterizarse respectivamente como efectivamente (inductivamente) reales o ilusorias. Así pues, la realidad y la ilusión son los extremos (100 % y 0 %, respectivamente) en un amplio rango de probabilidades con muchos grados intermedios (incluida la problemática en el punto medio).

Para ser aún más precisos, *las percepciones puras* (es decir, las apariencias concretas, los fenómenos) nunca son ilusorias. El juicio de valor de «ilusorio» se refiere propiamente a conceptos (es decir, apariencias abstractas, «universales»). Cuando decimos que una percepción fue ilusoria, simplemente queremos decir que la interpretamos erróneamente. Es decir, lo que inicialmente consideramos una percepción pura, en realidad contenía *un componente conceptual* que, como se demostró, era erróneo. Por ejemplo, veo ciertas formas y colores a lo lejos y pienso «ahí viene una chica en bicicleta», pero al acercarme me doy cuenta de que solo vi un montón de basura al borde de la carretera. La percepción pura son las formas y los colores que veo; la interpretación errónea es «chica en bicicleta», la interpretación más acertada es «montón de basura». La percepción inicial no ha cambiado, pero mi mayor proximidad le ha añadido detalles perceptuales. Mi primera impresión era correcta, solo que mi juicio inicial era erróneo. Reviso este último concepto, no a través de algún medio superior de conocimiento, sino simplemente mediante *una mayor percepción y concepción* .

En rigor, la percepción nunca está en cuestión; lo que evaluamos son nuestras concepciones. En la práctica, ciertamente, suele ser muy difícil aislar una percepción de su interpretación, es decir, de sus apéndices conceptuales. Nuestra percepción de las cosas es, en gran medida, «eidética». Sin embargo, este hecho no tiene por qué llevarnos a rechazar ninguna percepción (como hacen rápidamente muchos filósofos occidentales, así como budistas), ni siquiera ninguna concepción. Las «impurezas» conceptuales en las percepciones no son necesariamente erróneas. Sabemos que lo son cuando descubrimos una causa específica de queja: una contradicción lógica o experiencial. Mientras no encontremos tal fallo específico, pueden considerarse correctas. Esto simplemente significa que debemos aplicar las reglas de aducción [95]a nuestras interpretaciones inmediatas de percepciones individuales, del mismo modo que lo hacemos con teorías complejas relativas a conjuntos de percepciones. Estas reglas son universales: ningún juicio está exento de la necesidad de un examen y una reevaluación cuidadosos.

Ahora bien, a juzgar por el relato de Cheng y ciertas citas de Nagarjuna que contiene, podríamos interpretar que este último intentaba expresar precisamente lo que yo he dicho. Por ejemplo, Cheng escribe [96]: «Lo que Nagarjuna quería negar es que los fenómenos empíricos… sean absolutamente reales… Sin embargo, [esto] no significa que no exista nada. *No anula nada en el mundo* » (cursivas mías). Interpreto esta no anulación como un reconocimiento de la apariencia como base mínima del conocimiento. Nagarjuna pudo haber tenido dificultades

---

[95] La aducción trata todo conocimiento conceptual como hipotético, para ser probado repetidamente, en competencia con todas las hipótesis alternativas concebibles, con referencia a toda la lógica y experiencia disponibles.

[96] Pág. 42.

para desarrollar una terminología adecuada (distinguiendo existencia, apariencia y realidad, como hago yo anteriormente), influenciado sin duda por su inclinación a las afirmaciones paradójicas que parecían expresar y confirmar la doctrina mística budista.

Pero si eso es lo que quiso decir, entonces no ha logrado llegar a un «camino medio» (una negación de la Ley del Tercero Excluido), sino solo a un «camino común» (un terreno común aceptado). En mi opinión, eso no es un logro insignificante: ¡el descubrimiento filosófico de la fenomenología! Pero para él eso sería trivial, si no contraproducente, pues lo que busca es negar la conciencia ordinaria y sus racionalidades limitantes, y así dar un salto a una conciencia diferente y superior, capaz de alcanzar la verdad trascendental o la realidad última.

Es interesante observar que la negación efectiva de la realidad en toda apariencia por parte de la escuela Madhyamika no fue aceptada por una escuela posterior de filosofía Mahayana, la Yogacara (siglos VII - VIII d. C.). Cheng describe la postura de esta última de la siguiente manera [97]: "Todo objeto, tanto mental como no mental, puede ser probado lógica o dialécticamente como ilusorio. Pero para que algo sea ilusorio, debe existir un pensamiento que padezca ilusión. *El hecho mismo de la ilusión prueba la existencia y la realidad de una cierta conciencia o mente* . Afirmar que todo lo mental y lo no mental es irreal es un suicidio intelectual. La realidad de algo debería, al menos, admitirse para que tenga sentido hablar de ilusión" (cursivas mías). Ese es el tenor del argumento fenomenológico que presento anteriormente, aunque mi conclusión final claramente no coincide con la de la Yogacara, que sostiene que todo es conciencia o mente (un tipo de idealismo), sino que deja abierta la

97 Pág. 25.

posibilidad de juzgar y clasificar las apariencias como materia o mente en función de diversas consideraciones.

El rechazo del «dualismo» en la escuela Madhyamika llega al extremo de implicar que la «vacuidad» no se encuentra en el nirvana, la antítesis del samsara (según la perspectiva budista anterior), sino en «ni el samsara ni el nirvana». En realidad, afirmaciones similares se pueden encontrar en el Canon Pali, es decir, en las escuelas Theravada mucho más antiguas, por lo que no se trata de una construcción exclusivamente Mahayana. La diferencia radica en el énfasis: tales afirmaciones, relativamente raras en el período anterior, son la norma y se repiten con frecuencia en el período posterior. Un ejemplo se puede encontrar en el *Dhammapada* , un sutra que data del siglo III [a] . C. [98], es decir, cuatrocientos o quinientos años antes de Nagarjuna. Aquí, el samsara se compara con un arroyo o su orilla, y el nirvana con la otra orilla; y se nos insta a trascender ambos.

> *Cuando hayas cruzado el río del Samsara, alcanzarás el Nirvana… Ha llegado a la otra orilla, entonces alcanza la visión suprema y todas sus ataduras se rompen. Aquel para quien no existe ni esta orilla ni la otra, ni ambas…*

Esta fórmula es legítima si se interpreta como una advertencia de que *la búsqueda* del nirvana (iluminación

---

98 Supuestamente, esta es la fecha de composición, aunque el traductor, Juan Mascaro, en su introducción, afirma que fue «compilado» en ese momento, lo que parece implicar una composición anterior. En ese comentario no queda claro cuándo se estima que se escribió el sutra por primera vez. Y si fue mucho más tarde, digamos durante el período de cristalización del pensamiento Mahayana, entre el 100 a. C. y el 100 d. C., este último pudo haber influido en los monjes que lo transcribieron. Véase el capítulo 26 (383-5) para la cita.

y liberación) constituye un obstáculo para alcanzarlo, una mera forma sutil de samsara (ignorancia y apego). No hay contradicción alguna en afirmar que *la idea del* nirvana como meta de acción nos mantiene en el samsara; se trata de una proposición causal común. La fórmula también es lógicamente aceptable si se considera como un recordatorio de que ninguna palabra ni concepto —ni siquiera «samsara» o «nirvana»— puede capturar o transmitir la totalidad de los significados que se pretenden transmitir (es decir, «no» aquí debería expresarse con mayor precisión como «no del todo»). Tampoco hay contradicción alguna en afirmar que quien ha alcanzado el nirvana no necesita abandonar el mundo de aquellos atrapados en el samsara, sino que puede seguir existiendo y actuando en él, aunque de una manera distintiva y libre de apego.

Pero sería una contradicción hablar de «vacuidad» como «ni samsara ni nirvana», dado que el concepto de nirvana se define originalmente como no-samsara; la verdad no puede ser una tercera alternativa. En el mejor de los casos, se podría decir que la vacuidad es un nivel superior de nirvana (en un sentido amplio), que no debe confundirse con el nivel inferior al que se refiere el término original, nirvana, ni, por supuesto, con samsara. En ese caso, nirvana (en el sentido genérico del término, que significa literalmente no-samsara) incluye tanto una especie superior como una inferior; y la afirmación «ni samsara ni nirvana inferior» es entonces compatible con la afirmación «nirvana superior». Hay una gran diferencia entre el lenguaje crudo, poético y dramático, y su interpretación literal.

15. Más discurso antinómico budista

De Rumiaciones 5 y Reflexiones lógicas y espirituales 3.10.

1. La tesis de "No tengo tesis"

a. El [99]filósofo budista Nagarjuna (India, *c* . 150-250 d. C.) atacó toda tesis que consideraba racional por todos los medios que consideraba lógicos, y declaró que su propio discurso era inmune al escrutinio y la crítica, al decir (según una traducción):

> " *Si tuviera una tesis, estaría culpable; puesto que solo yo no tengo tesis, solo yo estoy libre de culpa* " (VV 29) [100].

---

[99] Huelga decir que los siguientes comentarios no constituyen un ataque al budismo, sino a la retórica de Nagarjuna. Este tipo de juegos no benefician al budismo. Pienso en Nagarjuna cada vez que leo el versículo 306 del *Dhammapada* : «Quien dice lo que no es… y quien dice no haber hecho lo que sabe que ha hecho… peca contra la verdad». Para mí, es simplemente un filósofo como cualquier otro; su interés por el budismo es secundario (al igual que su estatus de santo a ojos de muchos).

[100] Nagarjuna en *Vigraha Vyavartani* (*Evitando las discusiones*), versículo 29. La traducción utilizada aquí es la de 'Namdrol' en el foro de budismo E-Sangha (http://www.lioncity.net/buddhism/index.php?s=d8946a5bcb1f56f3e9e21a108125823f&showtopic=5604&st=100&#entry82577). Sin embargo, cabe señalar que la palabra "solo" en esta traducción podría

El primer aspecto de la declaración de Nagarjuna es una descarada **afirmación de no tener ninguna afirmación.** Esto, por supuesto, es contradictorio. Toda proposición que pretende ser significativa y verdadera (ya sea sobre alguna experiencia o sobre una abstracción, sea positiva o negativa) es una aserción, una afirmación. Fingir no hacer ninguna afirmación, aun cuando claramente se hace una, es una violación del principio de identidad: es negar que un hecho sea un hecho.

No existe una manera lógica de negar o criticar las tesis o metodologías de otros sin someter el propio discurso a evaluación. Toda negación o crítica es discurso, y todo discurso está sujeto a revisión lógica. Pretender la posibilidad lógica de una dispensa es deshonesto (y si tal pretensión implícitamente ya es bastante mala, lo es aún más si se hace explícitamente).

El discurso de Nagarjuna era, de hecho (como demuestro en **Ilógica budista**), un espectáculo repleto de falacias, una mera parodia de la lógica disfrazada de lógica. Pero sabía que quienes carecían de formación en lógica caerían en la trampa, y selló su destino intelectual con esa afirmación engañosa. Para neutralizar cualquier debate posterior, los indujo a creer que simplemente había demostrado el absurdo lógico de la lógica y de todas las doctrinas basadas en ella, sin haber propuesto él mismo ninguna metodología ni doctrina propia.

Su supuesta refutación de la razón no solo estaba plagada de errores de razonamiento, sino que su afirmación final de «no hacer ninguna afirmación» era una burla a la lógica y la sinceridad. Él, por supuesto, simplemente dice «no hago ninguna afirmación», y niega persistentemente que esta declaración constituya una afirmación. A eso lo

---

no estar en el original, a juzgar por otras traducciones que he visto, aunque parece ser la intención de Nagarjuna.

llamo manipulación psicológica descarada, motivada por el afán de superioridad. Se aprovecha cínicamente de la credulidad de algunas personas para dominarlas intelectualmente.

El segundo aspecto de la afirmación de Nagarjuna puede considerarse una versión atenuada de la paradoja del mentiroso, ya que nos dice: « **Todos menos yo están equivocados** ». Si bien esta afirmación no es inconsistente en sí misma (Dios podría decirla con toda sinceridad), lógicamente genera dudas debido a su carácter *autoexcluyente* .

En esencia, esto significa: «Soy el único ser humano que posee conocimiento; sé que todos los demás son incapaces de alcanzar el verdadero conocimiento». Solo un necio se deja engañar por una pretensión de privilegio tan infundada. La razón considera que todas las personas, con el esfuerzo suficiente, tienen acceso al conocimiento, aunque no todas desarrollen su potencial por igual. La razón exige que el discurso sea razonado y justo, es decir, que se base en *normas generales comunes* sobre cómo determinar la verdad y la falsedad.

Si Nagarjuna basara su crítica a los medios humanos ordinarios para adquirir conocimiento en la afirmación de haber alcanzado un *«nivel superior» de conciencia* (es decir, la iluminación budista o la profecía bíblica), no podríamos refutarlo de manera convincente (al ser incapaces de probar o refutar tales afirmaciones basadas en la experiencia). Pero él no utiliza tal afirmación como base; intenta desacreditar la razón mediante el discurso lógico *ordinario* . En ese caso, es un blanco legítimo para el análisis lógico.

La declaración de infalibilidad se interpreta entonces como una arrogancia manifiesta, una falta de respeto hacia otros pensadores. Al afirmar «solo yo estoy exento de toda crítica», el autor se concede agresivamente una

dispensa especial: solo él posee el camino al conocimiento; todos los demás son idiotas o personas deshonestas. Es un discurso totalitario y dictatorial.

Comparemos esta afirmación desdeñosa de «ustedes no saben nada» con la declaración inclusiva «yo (o nosotros) no sabemos nada». Esta última, aunque implica «sé que no sé nada» y, por lo tanto, es contradictoria, es al menos modesta; tanto es así que tal admisión se considera ampliamente una señal de sabiduría (y es encomiable, en su forma modificada, es decir, como «no sé *casi* nada, muy poco»).

La autoexclusión es una forma encubierta de incoherencia, pues recurre a *un doble rasero* . Quien hace tal afirmación presenta argumentos superficialmente racionales contra la experiencia y la lógica humanas, pero no se pregunta ni nos explica cómo él (un ser humano común) logró (utilizando precisamente los medios cognitivos que rechaza) alcanzar ese supuesto conocimiento verdadero. El autor critica a los demás, pero no se critica a sí mismo con la misma eficacia.

Este es un modo de pensamiento falaz que se encuentra frecuentemente entre quienes pretenden ser filósofos escépticos. Se manifiesta de muchas maneras sutiles. Es prudente estar siempre alerta ante tales prácticas, aplicando la prueba reflexiva que aquí se demuestra.

b. Analizando con más detalle la declaración anterior de Nagarjuna, se puede agregar lo siguiente.

Para empezar, ¿qué se entiende aquí por “ ***tener una tesis*** ”? Se refiere a cualquier creencia explícita o incluso implícita, cualquier opinión clara o incluso vaga que se sostenga (considerada conocimiento), cualquier proposición que se defienda o que se apruebe implícitamente desde un punto de vista lógico. El tema que Nagarjuna aborda aquí es cualquier resultado de la cognición racional humana, cualquier creencia, opinión o

doctrina a la que se pueda llegar, correcta o incorrectamente, mediante la conciencia ordinaria, es decir, a través de la experiencia, la negación, la abstracción, la formulación de hipótesis o el argumento inductivo o deductivo.

¿Y qué se entiende aquí por " *estar equivocado* "? Se refiere a cometer un error en el curso de la observación o el razonamiento, de modo que alguna tesis a la que uno se ha adherido es en realidad una ilusión en lugar de una realidad, falsa en lugar de verdadera, errónea en lugar de correcta.

¿Cómo determinamos el estatus apropiado para una tesis? Lo determinamos (sugiero) *mediante la aplicación holística de toda la ciencia de la lógica a la totalidad de los datos de la experiencia* . Nuestros conceptos de lo correcto o incorrecto cognitivo se construyen, en sí mismos, mediante la lógica y la experiencia, sin recurrir a ninguna justificación externa extraordinaria (como la revelación profética o la realización mística, o simplemente la autoridad de algún personaje importante o de un documento o institución religiosa).

Ahora bien, Nagarjuna es evidentemente consciente de todo esto, pero su intención es anular la fiabilidad independiente de la experiencia y la razón ordinarias. Su estrategia y táctica para lograr este fin, en todo su discurso, como he demostrado a lo largo de mi *obra La ilógica budista* , consiste en dar la impresión (por paradójica que parezca) de que la lógica puede invalidarse mediante la lógica misma. Y esta doble frase suya, «Si tuviera una tesis, estaría equivocado; puesto que solo yo no tengo tesis, solo yo estoy libre de culpa» *,* encaja a la perfección en su destructivo programa filosófico.

A primera vista, esta frase podría interpretarse como un único argumento:

Si X (se propone una proposición), entonces Y (se comete un error).

pero no X (ninguna proposición)

Por lo tanto, no es Y (sin error)

Aunque la apódosis anterior es lógicamente inválida, puesto que niega el antecedente para negar el consecuente, Nagarjuna no duda en dejarla pasar sin comentarios, sabiendo que bastará para convencer a algunos, si bien es consciente de que quienes tienen formación lógica la detectarán y objetarán. Pero para este último público, reserva una forma de manipulación más sutil.

Es evidente que el propósito del famoso verso 29 del discurso de Nagarjuna es *aparentar coherencia lógica* . Con él, pretende dar la impresión de que su discurso antiracional es justificable, que cuenta con el respaldo de la lógica. Si bien ataca la lógica, al mismo tiempo debe fingir que la utiliza, pues sabe que esta estrategia es necesaria para convencer a la gente. Para la mayoría, una apariencia de lógica (mera retórica) basta para tranquilizar su capacidad crítica. Veremos ahora cómo lo logra.

La primera parte de la afirmación de Nagarjuna, a saber, «Si tuviera una tesis, estaría equivocado», no pretende (como algunos han supuesto) justificar su discurso en su conjunto. No se incluye aquí como un argumento con fuerza intrínseca que respalde o demuestre directamente su filosofía. Ciertamente, no es un principio lógico obvio, ni un axioma, en el que todos estarían de acuerdo sin objeción, del cual se pueda inferir o incluso generalizar su discurso. No, es en sí misma una inferencia y aplicación de la tesis principal de Nagarjuna, a saber, la afirmación de que «Todo conocimiento humano basado en la experiencia y la razón ordinarias es necesariamente erróneo».

Esta última afirmación subyacente es su premisa principal en una educación productiva (aquí tácita), es decir, una que deduce una proposición hipotética particular a partir de una categórica más general [101]. Este argumento es formalmente válido y se desarrolla de la siguiente manera:

Todas las X (opiniones) son necesariamente Y (erróneas);

por lo tanto,

Si esto es X (se propone una proposición), entonces esto es Y (se comete un error).

De esta forma, la primera parte de la afirmación de Nagarjuna se presenta como algo inferido, en lugar de una aseveración arbitraria. Se presenta astutamente como una aplicación de información ya admitida, en vez de una afirmación aislada. Si se acepta la premisa, la conclusión se deduce lógicamente (esta es la apariencia de lógica), pero ¿se ha aceptado ya la premisa? No. Cabe señalar también que, una vez que la conclusión parece estar dada, puede utilizarse por generalización para reforzar la premisa; aunque esto constituye una circularidad, funciona psicológicamente.

Además, mediante este argumento productivo implícito, Nagarjuna logra aparentar coherencia consigo mismo: en efecto, nos dice: «¡Miren, no solo ataco el conocimiento ajeno, sino que estoy dispuesto a aplicar la misma crítica rigurosa al mío!». Esta declaración virtuosa, por supuesto, es una ilusión, pues no está poniendo en duda el principio general, sino simplemente reafirmándolo. En ningún momento ha demostrado que «Todas las proposiciones son falsas». Su postura es pseudológica.

---

101 Véase *Lógica del futuro* , capítulo 29.3.

Como aclara la siguiente parte de su declaración, no considera que su discurso se ajuste a la regla crítica que él mismo ha formulado. La proposición «Si tuviera una tesis, estaría equivocado» es una hipótesis contrafactual; su propio discurso nunca se convierte en un tema abierto a debate. Parece tener una mentalidad abierta, pero es un juicio preestablecido. Su intención es «evitar toda discusión» y situarse desde el principio fuera de la contienda. Aparentemente, primero admite y luego niega vehementemente que su propio discurso sea producto de la conciencia ordinaria. Esta enrevesada evasión de la responsabilidad cognitiva ha engañado a muchos.

Pasando ahora a la segunda parte de la afirmación de Nagarjuna, a saber: «Como solo yo no tengo tesis, solo yo estoy libre de culpa». Como ya se ha señalado, esto puede considerarse la premisa menor y la conclusión de una apódosis inválida, en la que la primera parte de la afirmación es la premisa mayor. Pero también podríamos suponer, de forma más generosa, que Nagarjuna pretendía una apódosis válida, utilizando como premisa mayor tácita la proposición obvia: «Si uno no tiene tesis, no puede equivocarse».

Se puede argumentar, con razón, que esta premisa se dejó tácita simplemente porque es obvia y fácilmente aceptada por todos. Es cierto que si uno no se atreve a pronunciar palabra, pensamiento ni siquiera intención, si no tiene opinión, no pretende tener conocimiento, si permanece en silencio, tanto interior como exteriormente, nunca cometerá errores. *Porque el estatus de verdad o falsedad solo se aplica a afirmaciones con sentido.*

Una piedra nunca se equivoca, porque no tiene pensamientos. Del mismo modo, una persona irreflexiva puede, por su ignorancia, ceguera o estupidez, cometer muchos errores en la vida, pero no comete ningún error en el sentido lógico de haber propuesto una proposición

inapropiada. Todo lo que es tan obvio (y vacío) no requiere debate. La siguiente apódosis está, por lo tanto, implícita en la declaración de Nagarjuna:

Si no X (no se propone ninguna proposición), entonces Y (no se comete ningún error).

pero no X (ninguna proposición)

Por lo tanto, no es Y (sin error)

Este argumento posee una premisa mayor genuina, además de una forma válida. Esto le confiere a su discurso una apariencia de lógica, ayudándole a persuadir a más víctimas. Sin embargo, su premisa menor sigue siendo cuestionable y, sin duda, negable. (En consecuencia, su conclusión también es, por supuesto, cuestionable). Da por sentado que «no tiene tesis», pero esta afirmación dista mucho de ser cierta. La premisa mayor tácita actúa como una cortina de humo para la premisa menor.

Además, cabe señalar que, si bien «tener razón» implica «no estar equivocado», lo contrario no es necesario. Nagarjuna sugiere que su supuesta impecabilidad implica la corrección de su postura, ¡pero esto no se deduce! Solo si su crítica a todas las tesis opuestas fuera correcta (lo cual es completamente falso), y su tesis no fuera susceptible de críticas similares y, por lo tanto, fuera la única posibilidad lógica restante, se podría llegar a tal conclusión.

Nagarjuna sí tiene una tesis. Su tesis principal, el objetivo de todo su discurso filosófico, es, como ya se mencionó, la afirmación de que «todo conocimiento humano basado en la experiencia y la razón ordinarias es necesariamente erróneo». Esto, para empezar, se puede considerar una tesis; si bien es una tesis sumamente escéptica, cargada de

implicaciones negativas. Es un principio lógico que *negar una tesis equivale a afirmar la tesis opuesta.* Su afirmación de que su doctrina no es una tesis, en la premisa menor que se presenta aquí, es una mera aseveración arbitraria.

Además, él «tiene una tesis» cada vez que hace una afirmación específica de cualquier tipo, incluida la afirmación que aquí se analiza: «Si tuviera una tesis, estaría equivocado; como solo yo no tengo tesis, solo yo estoy libre de culpa». Cabe señalar que Nagarjuna piensa que hacer una afirmación negativa es, de alguna manera, «no tener una tesis», pero la polaridad de una afirmación no disminuye la necesidad de justificación; de hecho, se podría argumentar que, por el contrario, las afirmaciones negativas son más difíciles de fundamentar que las positivas.

Y debemos incluir como «tesis» suyas no solo las afirmaciones explícitas, sino también todas las suposiciones y sugerencias implícitas en su discurso (como la premisa mayor implícita y la apódosis resultante que acabamos de destacar). Da igual si estos elementos explícitos, o no declarados ni admitidos, constituyen información, método lógico, contenido o proceso.

Todos estos elementos del discurso, ya sean verbales o implícitos, se ajustan, con toda justicia, a nuestra comprensión y definición comunes de lo que significa «tener una tesis». Ninguna de estas proposiciones categóricas o hipotéticas (salvo quizás «si el silencio no implica error») es evidente por sí misma. No surgieron *de la nada* en la mente de Nagarjuna, prefabricadas y justificadas por sí mismas.

Todas ellas son productos complejos de la cognición humana ordinaria, basados en la experiencia y generados por la razón (aunque, en el caso de Nagarjuna, la mente implicada esté perturbada). Indudablemente, en conjunto

conforman una filosofía específica, una teoría de la lógica, una epistemología y una ontología. El mero hecho de que podamos (como se hace aquí) considerarlas y debatirlas demuestra que son «tesis».

Debe mantenerse la ley de identidad (A es A): los hechos son hechos y de nada sirve pretender lo contrario. Nagarjuna podrá rechazar eternamente el predicado de «tener una tesis», pero nosotros insistimos en él con seguridad. Sus argumentos no han logrado en absoluto eludir este juicio justo y verdadero. En consecuencia, su doctrina es contradictoria. No solo «tiene una tesis», sino que, puesto que su tesis es que «tener una tesis es estar equivocado», debe ser reconocido (según sus propios términos) como alguien que está equivocado.

En resumen: la afirmación de Nagarjuna «Si tuviera una tesis, sería culpable; puesto que no tengo tesis, soy el único que está libre de culpa» teje una compleja red de engaños. Induce a error mediante sutiles ambigüedades e imitaciones superficiales de la lógica. Una vez que se revela su deshonestidad, debe ser rechazada categóricamente.

El mero hecho histórico de que Nagarjuna sea famoso y admirado por muchos no justifica aferrarse a su doctrina *hasta la saciedad* , intentando a posteriori encontrar maneras de hacerla coherente con la lógica. La fama no es prueba de una verdad oculta; es vanidad. La mayoría de quienes lo hacen solo buscan una gloria reflejada. En cualquier caso, el apego a la autoridad es un argumento *ad hominem* . Los seguidores religiosos y académicos que le otorgaron y perpetúan su autoridad no son lógicamente competentes, por muchos que sean. Es como si un ciego guiara a otro ciego.

c. Nagarjuna defiende su idea de la «no-tesis» en el siguiente verso (VV 30), describiéndola como « ***una no-aprehensión de las no-cosas*** » (según una traducción

[102]). Ahora bien, esta es una frase muy curiosa. Para el impresionable, suena muy profunda, cargada de significado. Parece sugerir que este hombre posee una forma privilegiada de conocimiento superior que trasciende la experiencia y el razonamiento ordinarios.

Pero, en realidad, tomado literalmente, todos somos perfectamente capaces de «no aprehender las no-cosas» y lo practicamos a diario, ¡por la sencilla razón de que las no-cosas no pueden ser aprehendidas! Lógicamente, esto es todo lo que significa esta frase; ténganlo bien presente. ¿Qué estará tramando entonces el viejo zorro?

Nagarjuna intenta proyectar su postura de "no tener tesis" lo más lejos posible lógicamente de nuestra concepción común de "tener tesis", es decir, de la conciencia ordinaria, que consiste en "la aprehensión de las cosas". Lógicamente, solo tiene tres alternativas para elegir:

- la 'falta de aprehensión de las cosas' (inconsciencia);
- la 'aprehensión de las no-cosas' (una conciencia de otro mundo);
- o la 'no aprehensión de las no-cosas'.

Con su marcada inclinación por la rivalidad y los extremos dramáticos, Nagarjuna, por supuesto, eligió el

---

[102] Por Frederick J. Streng. El texto completo de su traducción parece ser el publicado en Internet en: http://www.orientalia.org/article491.html . Cabe señalar que Plamen Gradinarov considera que la frase «no aprehensión de las no-cosas» es una traducción incorrecta. Sin embargo, si bien acepto la objeción de este último, no estoy de acuerdo en que la traducción más libre de Streng sea completamente inadmisible. En mi opinión, puede que no sea literalmente precisa, pero capta el espíritu y la intención paradójicos de Nagarjuna. Véase nuestra discusión sobre este tema en http://nyaya.darsana.org/topic3.html . En cualquier caso, incluso si no es recomendable basarse en la frase «no aprehensión de las no-cosas», las críticas a Nagarjuna en esta sección aún pueden plantearse sobre otras bases.

tercero de estos términos como vehículo para su obra. Si bien el sentido obvio de esta frase es pueril, posee amplitud y atractivo poético. Parece implicar simultáneamente «conocimiento sin consciencia» y «conciencia de lo incognoscible».

Así pues, su «no-aprehensión» es una mezcla de aprehensión y no-aprehensión, o algo completamente distinto. Del mismo modo, sus «no-cosas» son cosas de algún tipo, además de no-cosas, o quizás algo totalmente diferente.

En otras palabras, los términos negativos en la frase "no aprehensión de las no-cosas" no son entendidos por Nagarjuna ni recibidos por sus discípulos y estudiantes como meras negaciones de los términos positivos correspondientes, sino como **términos paradójicos** , que pueden (de acuerdo con el esquema del tetralema) ser a la vez *positivos o negativos o ambos o ninguno* .

Es (y no es) 'aprehensión/no aprehensión de cosas/no cosas' todo en uno.

Nagarjuna destaca en la historia de la filosofía mundial como el opositor más acérrimo de las leyes del pensamiento. No solo utiliza libremente proposiciones contradictorias o que incluyen un término medio, sino que incluso emplea términos cargados de contradicción e inclusión de un término medio.

Ahora bien, algunos podrían decir: "¿Qué tiene de malo eso?". Argumentarán: "El mundo real es extremadamente sutil, y solo podemos aspirar a expresarlo en el pensamiento de forma muy aproximada; Nagarjuna simplemente intenta tener en cuenta esta incertidumbre en su discurso; las leyes del pensamiento son solo exigencias arbitrarias que nos obligan a forzar nuestros pensamientos a usar camisas de fuerza prejuiciosas".

Pero la laxitud lógica no es la actitud adecuada ante un mundo real extremadamente complejo y difícil de

expresar. Precisamente por la gran dificultad de la tarea cognitiva en cuestión, se exige claridad y cautela. Evitar controles y equilibrios en nuestros juicios no aumenta su eficacia, sino que los hace menos fiables.

En el caso que nos ocupa, si Nagarjuna posee realmente una forma privilegiada de conciencia trascendental, puede simplemente decirlo. Las leyes del pensamiento no le impiden en absoluto plantear tal afirmación. No necesita andarse con rodeos ni pretender tener algo inefable e incuestionable. Puede y debe ser franco y defender su postura con equidad, como cualquier otra persona.

Si considera que los términos «aprehensión» y «cosas» tienen algún defecto lógico intrínseco, puede argumentar su postura abiertamente; no necesita recurrir a alusiones, sugerencias ni falacias. La mayoría de los pensadores somos de mente abierta y estamos dispuestos a corregir nuestros errores: si estos términos son erróneos, no nos aferramos a ellos; somos flexibles y estamos listos para modificarlos o reemplazarlos según sea necesario a la luz de nuevas evidencias y razonamientos.

Pero Nagarjuna es como un acusado que, al verse obligado a comparecer ante el tribunal, se niega a admitir su identidad, a reconocer la autoridad de la ley y de los jueces, a declararse culpable o inocente, o a presentar su defensa. Peor aún, en un flagrante desacato al tribunal, ni siquiera admite que su negativa sea tal: la califica de «no-tesis». ¿Acaso eso detiene el proceso judicial o hace que el tribunal lo declare inocente? Sin duda, no.

Nagarjuna malinterpreta la naturaleza de la negación. Piensa que si una persona dice «X» y otra dice «no X», la carga de la prueba recae más en la primera que en la segunda. Considera que hacer una afirmación es lógicamente más exigente que hacer una negación. *En su confusión, imagina que decir «no» equivale a no decir nada* . La mayoría de los lógicos discreparían con él y

argumentarían que cualquier tesis planteada (aunque sea por insinuación) requiere igualmente ser probada, independientemente de su polaridad.

Yo iría más allá y diría que, por el contrario, una afirmación negativa es más difícil de demostrar que una positiva. Se puede probar una afirmación positiva con relativa facilidad si se presentan pruebas suficientes a su favor. Pero, ¿cómo se prueba una afirmación negativa? Es mucho más difícil, ya que las negaciones no se experimentan directamente, sino únicamente por la ausencia de afirmaciones positivas. En última instancia, una negación solo puede probarse indirectamente, por la imposibilidad de probar cualquier afirmación positiva contraria.

Así pues, la supuesta autolimitación de Nagarjuna a las negaciones no solo no lo exime de la necesidad de pruebas, sino que, por el contrario, aumenta la carga lógica que recae sobre él. Tiene razón al considerar que las negaciones son significativamente diferentes de las afirmaciones, pero no se da cuenta de que esta diferencia le perjudica. Afirma carecer de fundamento epistemológico u ontológico, y aun así pretende rechazar sin más todas las teorías del conocimiento y la realidad. ¡Una afirmación tan grandilocuente y fantasiosa requiere, sin duda, mucha más justificación que cualquier otra!

Cabe destacar, de paso, que la «no aprehensión de las no-cosas» de Nagarjuna no debe interpretarse (como algunos hacen) como una defensa de la experiencia meditativa no verbal o la intuición. Ese no es el objetivo principal de su filosofía antiracional, aunque su declarada afiliación budista pueda llevar a suponerlo.

Si Nagarjuna fuera un hombre profundamente inmerso en la meditación, no estaría escribiendo filosofía. Si su intención fuera promover la meditación, simplemente enseñaría métodos de meditación y no incitaría a disputas

verbales. No, este hombre tiene ambiciones filosóficas. Supuestamente, estas obras pretenden plasmar en palabras parte del «razonamiento» que, según él, siguió el Buda antes de alcanzar la iluminación. Nagarjuna parte de la premisa de que este «razonamiento» es necesariamente antilógico, un rechazo de la razón.

Pero debemos reconocer que esta suposición no es más que un prejuicio de su mente distorsionada. Fue un revolucionario filosófico, alguien que creía que la razón debía ser subvertida, trascendida. Sin embargo, resulta más creíble adoptar una perspectiva evolucionista y considerar la meditación como una forma de seguir avanzando, más allá de los límites del pensamiento discursivo, sin necesidad de negar dicho pensamiento dentro de sus límites aplicables.

Defender el respeto por la lógica no implica fomentar la palabrería sin fin, sino exigir que todo pensamiento que surja sea sometido a una evaluación cognitiva responsable. La lógica es posible sin palabras, mediante intenciones silenciosas. Incluso en la meditación profunda, se produce una especie de «verificación de la realidad» a través de la lógica, sin necesidad de palabras. Sin duda, solo así un aspirante a Buda podrá alejarse de las ilusiones comunes y las fantasías insensatas, encaminándose hacia la plena realización.

Contrariamente a la creencia de Nagarjuna, la racionalidad y la espiritualidad no están necesariamente en conflicto. La razón y la meditación pueden, hasta cierto punto, ser mutuamente beneficiosas. No es el pensamiento en sí mismo, ni mucho menos la lógica, sino el exceso de pensamiento, especialmente la charla irrelevante, lo que obstaculiza la concentración y la contemplación meditativas. A menudo se requiere un cierto grado de pensamiento adecuado para preparar la mente para la meditación.

d. De hecho, como demostraré a continuación, la oración «Si tuviera una tesis, estaría equivocado» [103]es una **imposibilidad formal** . Anteriormente la interpreté y simbolicé como «Si X (se propone una proposición), entonces Y (se comete un error)», asignando al antecedente y al consecuente dos símbolos separados, X e Y. Pero ahora analicemos estos constituyentes con mayor detenimiento.

¿Qué significa aquí "cometer un error"? No se trata de un predicado común. El consecuente Y no se refiere simplemente a un error en general, sino específicamente a un error *en el antecedente* X. Y nos dice que X *es incorrecto* . Por lo tanto, Y implica formalmente la negación de X, es decir, ¡no X! Partiendo de esta premisa, la oración de Nagarjuna queda así: " **Si X, entonces no X** ", es decir, "Si X es verdadero, entonces X es falso"; una proposición hipotética paradójica, cuya conclusión sería la categórica "X es falso" (como se sugirió anteriormente).

Sin embargo, la cuestión no termina ahí. Si ahora consideramos el significado de X —a saber, «se propone una proposición»—, podemos suponer razonablemente que se refiere a *cualquier proposición* . En ese caso, la proposición en cuestión podría incluso ser la negación de X; de modo que podemos sustituir X por noX en toda la hipótesis. Así, obtenemos «Si noX, entonces no noX», es decir, « **Si no X, entonces X** », o dicho de otro modo, «Si X es falso, entonces X es verdadero». Esta es también, por supuesto, una proposición paradójica, cuya conclusión formal es «X es verdadero».

---

[103] Otras dos traducciones de esta frase confirman y amplían esta interpretación. «Si hiciera alguna proposición, incurriría en un error lógico» (Streng). «Si hubiera planteado alguna tesis, el error lógico habría sido mío» (Gradinarov).

Así pues, mediante una interpretación universal de «tener una tesis», que incluye también «no tenerla», nos encontramos ahora no solo ante una paradoja, ¡sino ante una *doble paradoja* ! Es decir, nuestra conclusión no es solo que X es falso, sino que X es a la vez verdadero y falso. Esta última conclusión, por supuesto, contradice el principio de no contradicción, como en el caso de la paradoja del mentiroso.

Esto significa que la afirmación de Nagarjuna es una imposibilidad formal: una contradicción en sí misma; no solo es falsa, sino carente de sentido. No constituye un discurso legítimo en absoluto, y mucho menos una postura o teoría filosófica sostenible. Las palabras o símbolos que utiliza son lógicamente inconcebibles, por lo que es como si no dijera nada. Parece decir algo inteligible, pero es una ilusión.

Ahora bien, podría objetarse que Y no implica necesariamente que X *sea* erróneo, sino simplemente que X *podría serlo* . Es decir, «cometer un error» podría significar que X es incierto en lugar de estar definitivamente refutado. En ese caso, tendríamos las siguientes dos hipótesis: «Si X, posiblemente no X» y «Si no X, posiblemente X»; o, en una sola frase: «Se proponga X o no X, el resultado es incierto». De hecho, esta postura más modal y ambigua bien podría considerarse la intención exacta de Nagarjuna (que algunos han interpretado como una «ilocución» evasiva).

A primera vista, debido al uso de palabras vagas o símbolos, esta objeción puede parecer creíble y las conclusiones contradictorias involucradas aparentemente se disipan. Pero, tras una reflexión, subyace un conflicto: afirmar X, o negarlo, es contrario a una postura que ni afirma ni niega X. Una afirmación categórica (que afirma o niega X) es *incompatible con* una afirmación problemática (que dice que X puede ser verdadero o

falso). No se puede afirmar a la vez *tener* conocimiento (de X, o de no X) y afirmar *carecer* de él (considerando abierta la cuestión de la verdad o falsedad). Esto es tan contradictorio como afirmar que una misma cosa (X) es verdadera y falsa.

Alguien que desconozca la lógica de las proposiciones hipotéticas podría objetar que X, o no X, se propone solo hipotéticamente en el antecedente, y por lo tanto podría ser problemático en el consecuente. Pero esta es una objeción lógicamente insostenible, debido al proceso de adición (descrito en el capítulo sobre lógica formal); es decir, debido a que "Si X, entonces Y" implica "Si X, entonces (X e Y)". En el presente caso, esto significa: "Si X se afirma, entonces X es a la vez afirmado e incierto". Basta con que la contradicción ocurra condicionalmente, como aquí, para que la condición se refute; por lo tanto, nuestra conclusión *es bastante formal* : "X no puede afirmarse". QED.

Alguien podría objetar, finalmente, que la certeza en el antecedente y la incertidumbre en el consecuente no necesariamente son simultáneas y, por lo tanto, no generan un conflicto lógico. Tal objeción sería válida, siempre que existiera un proceso de pensamiento que separara el principio y el final de la proposición hipotética. Sin embargo, en el caso que nos ocupa, Nagarjuna afirma claramente que, al mismo tiempo que se propone algo, se comete un error; es decir, el error no es otro que la propia proposición. Por consiguiente, no se puede argumentar de manera creíble una separación temporal, y la tesis de Nagarjuna sigue siendo ilógica.

Cabe señalar que toda la presente discusión se ha centrado únicamente en la primera parte del versículo 29, es decir, la premisa mayor: «Si tuviera una tesis, estaría equivocado». Hemos constatado que esta proposición hipotética es lógicamente errónea, independientemente de

que Nagarjuna admita o no que «tiene una tesis». Así pues, reconsideremos ahora su premisa menor y su conclusión de que «no está equivocado».

Aquí hemos introducido un nuevo giro en el análisis, al darnos cuenta de que "Si X, entonces Y" (entendido como "Si X, entonces no X") implica "Si no X, entonces Y" (ya que esto último se deduce de "Si no X, entonces X", que a su vez se deduce de lo anterior al reemplazar X por no X). Así pues, ahora tenemos una nueva premisa mayor para Nagarjuna, a saber, "Si no X, entonces Y", que significa: "Si no tengo una tesis, seré culpable".

Si tomamos esta premisa mayor implícita junto con la premisa menor de Nagarjuna, a saber, «No tengo tesis», la conclusión es «Estoy equivocado». Esta conclusión es, nótese, lo opuesto a la suya («No estoy equivocado»). Por lo tanto, aunque Nagarjuna se jacta de que su pensamiento es impecable, ¡es demostrablemente erróneo!

En pocas palabras, dejando de lado toda su retórica, lo único que dice es: «ninguna tesis es verdadera»; no es más que otra versión de la paradoja del mentiroso. Y su intento de atenuar su afirmación, con la aclaración «excepto mi tesis», es lógicamente una mera aclaración: un caso particular que se ajusta perfectamente a la regla general. Además, antes de que se pueda aplicar una excepción, la regla misma debe poder formularse de manera coherente, y esta claramente (como se acaba de demostrar) no lo es.

Por último, cabe señalar que ninguna de estas refutaciones implica que el silencio sea imposible o carezca de valor. Si (como sostienen algunos comentaristas) el propósito de Nagarjuna era promover el cese del discurso, sin duda lo hizo de forma equivocada. No necesitaba desarrollar una filosofía polémica y antilógica. Le habría bastado con postular, como un hecho psicológico, que el silencio

(interno y externo) es propicio para la meditación profunda.

2. Llamar a las cosas por su nombre

El budismo, sin duda desde sus inicios, presenta una mezcla de lógica e ilógica en su discurso. Al analizar sus cuatro principales escuelas filosóficas —Abhidharma, Prajnaparamita, Madhyamika y Yogacara—, la más propensa a descartar las tres leyes del pensamiento (es decir, la identidad, la no contradicción y la exclusión del término medio) fue Madhyamika [104]. Sin embargo, esta tendencia se inició en la anterior Prajnaparamita, como demuestran los ejemplos del *Sutra del Diamante* [105].

En este sutra encontramos ejemplos de argumentos lógicos válidos. Por ejemplo, en la Sección 12 hay un argumento a fortiori bien formulado [106]: «dondequiera que

---

104 Véase mi obra *«Ilógica budista»* sobre este tema, así como los comentarios sobre el discurso de Nagarjuna en mis *Reflexiones* , Parte I, capítulo 5. Debo recalcar que mi intención, tanto en las críticas anteriores como en la presente, no es rechazar el budismo en sí mismo, sino demostrar que puede armonizarse con la razón. Considero innecesaria y contraproducente la actitud de muchos filósofos budistas que, aparentemente, creen imposible alcanzar la Realización (es decir, la iluminación, la liberación y la sabiduría) sin rechazar la lógica. Mi principio rector es que ambas son perfectamente compatibles, y que, de hecho, la razón es un medio esencial (junto con la moral y la meditación) para lograr ese fin deseable.

105 A juzgar por su idioma sánscrito, la centralidad del ideal del bodhisattva y otros aspectos destacados, este sutra es un texto Mahayana. Se cree que fue compuesto y escrito en la India alrededor del año 350 d. C., aunque al menos una autoridad sugiere una fecha tan temprana como el año 150 d. C. A modo de comparación, Nagarjuna, el fundador de la filosofía Madhyamika, estuvo activo aproximadamente entre los años 150 y 200 d. C.; por lo tanto, este texto Prajnaparamita fue escrito durante el mismo período, si no mucho después.

106 Mu Soeng, pág. 111.

se predique este sutra, o incluso cuatro de sus versos, ese lugar será respetado por todos los seres… ¡Cuánto más [digno de respeto] será quien pueda memorizar y recitar este sutra…!». Pero también encontramos claras antinomias, como «el Dharma… no es ni comprensible ni elusivo» (dicho a pesar de que «no comprensible» significa elusivo, y «no elusivo» significa comprensible).

Pero el *Sutra del Diamante* utiliza repetidamente un tipo de argumento que, como lógico, yo clasificaría como una variante más de la ilógica budista. Este argumento afirma: « **Lo que se llama X no es en realidad X; por lo tanto, se llama X** » (o a veces: « **Lo que se llama X no es realmente X; es simplemente un nombre, por eso se llama X** »).

En el mencionado sutra se encuentran más de veinte ejemplos de este argumento. He aquí uno: «Lo que el Tathagata ha llamado Prajnaparamita, la sabiduría trascendental suprema, no es, de hecho, la Prajnaparamita y por eso se la llama Prajnaparamita». He aquí otro: «… lo que se denomina seres no son realmente seres. Tal es simplemente un nombre. Por eso el Tathagata se refirió a ellos como seres».[107]

Lo que cuestiono o discuto aquí respecto a este tipo de discurso es únicamente la conjunción « [108]*por lo tanto* » o « *por eso »* . No niego que se pueda usar un nombre

---

[107] En Mu Soeng: págs. 145 y 151, respectivamente. Encontré un argumento similar en otro texto Mahayana: "Y es porque para ellos [los bodhisattvas] el entrenamiento consiste en no entrenar que se dice que están entrenando" (mi traducción de una traducción francesa), que se encuentra en el capítulo 2, v. 33 del "Sutra de las palabras del Buda sobre la Sabiduría Suprema" (véase Eracle, pág. 61).

[108] Suponiendo, por supuesto, que la traducción de esta edición sea correcta (y parece bastante respetable; véase la página ix del prefacio). Mi argumento es que ningún lógico ha validado formalmente tal argumento; de hecho, es formalmente inválido, puesto que la conclusión contradice una premisa.

inapropiado para algo, ni siquiera que las palabras nunca puedan expresar con exactitud lo que uno realmente quiere decir. Pero afirmar que se le da el nombre X a algo *porque* no es X es, sin duda, absurdo e insostenible.

Esto no es simplemente «no llamar a las cosas por su nombre», sino llamar a algo por su nombre aun creyendo que no lo es. Esto, al menos en apariencia, es contrario a la lógica. Si la etiqueta no es aplicable, ¿por qué usarla? Es más, ¿por qué alardear de esta inversión inconcebible, diciendo «por lo tanto»?

Decir que algo "no es de hecho o verdaderamente X" implica que la palabra X tiene un significado que no se ajusta a lo que se está considerando; en tal caso, ¿por qué llamar a esa misma cosa "X" en contra de toda lógica? ¿Por qué no llamarla simplemente "no X" (o darle otro nombre más específico) y evitar la paradoja?

Discursos como «tal cosa es solo un nombre» resultan contraproducentes, ya que, de hecho, utilizan nombres que sí transmiten algún significado. La frase sugiere que ninguna palabra tiene una referencia válida, pero se basa en la efectividad de las palabras que emplea para comunicar sus diversas intenciones. Es una afirmación que intenta eximirse de las críticas que formula a *todas* las afirmaciones de este tipo.

En los ejemplos anteriores, el argumento se basa en nuestra comprensión de términos como «Prajnaparamita» (es decir, perfección de la sabiduría) o «seres», y al mismo tiempo intenta invalidar dicha comprensión. Por lo tanto, no se puede decir que comunique nada inteligible.

Sin duda, no podemos expresar adecuadamente la realidad última (o a Dios) con palabras. Pero sigue siendo cierto que podemos expresar verbalmente la inefabilidad (como se hizo en la frase anterior). No es necesario devaluar las palabras en sí mismas para admitir que tienen sus límites.

Además, es muy dudoso que tales afirmaciones paradójicas (como «nombra esto X *porque* no es X») sean psicológicamente útiles para alcanzar la iluminación; simplemente confunden cognitivamente e incapacitan la mente racional. En lugar de acallar la mente inquisitiva, lo único que hacen es estimularla con preguntas sin respuesta a nivel subconsciente. Estas afirmaciones sin sentido son producto de una desafortunada moda que se desarrolló en el budismo en cierta época [109].

Esa clase de perversidad intelectual llegó a parecer profunda, como lo parece para algunos pensadores posmodernos en Occidente hoy en día, precisamente porque una antinomia lógica no implica nada, y esa vacuidad de significado se equipara (erróneamente) con el Vacío subyacente a todos los fenómenos. El enorme vacío de conocimiento que deja la antinomia da la ilusión de estar repleto de significado, cuando en realidad no es más que evidencia de ignorancia. Tengan esto muy presente.

Cabe añadir que, en efecto, existe una especie de paradoja estructural en el acto meditativo, pero la costumbre del *Sutra del Diamante* de «llamar a las cosas por su nombre» no es la misma. La paradoja reside en que, si buscamos la iluminación mediante la meditación, no podemos aspirar a alcanzarla, pues entonces se alimenta nuestro ego (que se aferra a este valor trascendental como a un objeto mundano); sin embargo, la meditación es el mejor camino hacia la iluminación. Por lo tanto, debemos simplemente practicarla: sentarnos y dejar que nuestra iluminación innata (nuestra «naturaleza búdica») brille con el tiempo.

Cabe recordar también que el budismo se originó a partir de un profundo realismo. Es, esencialmente, una

---

[109] Aunque no están del todo ausentes en la literatura Abhidharma anterior ni en la literatura Yogacara posterior, no son infrecuentes en cierta literatura Prajnaparamita (incluido el *Sutra del Diamante*) y son bastante comunes en la literatura Madhyamika.

búsqueda de la Realidad. Desde esta perspectiva, la noción budista de «talidad» puede considerarse un compromiso con la Ley de la Identidad. El hombre iluminado es aquel que percibe las cosas, en particular y en general, *tal como son en realidad* .

Esto se pone de manifiesto, por ejemplo, en el siguiente intercambio zen. Un monje le preguntó a Li-shan: "¿Cuál es la razón [de que Bodhidharma viniera de Occidente, es decir, de la India a China]?", a lo que el maestro zen respondió: "Simplemente porque las cosas son como son", y en el comentario de DT Suzuki se indica que esto se refiere a la "talidad" [110].

110 *La doctrina zen de la no-mente* , pág. 93.

Referencias principales

Cheng, Hsueh-li. *Lógica vacía: budismo madhyamika de fuentes chinas* . Delhi: Motilal Banarsidass, 1991.

Dawkins, Richard. *El mayor espectáculo del mundo: la evidencia de la evolución.* Nueva York: Free Press, 2009.

El Dhammapada . Juan Mascaro, trad. Harmondsworth, Mx.: Penguin, 1973.

Éracle, Jean. *Enseignements du Bouddha* . París: Librio, 2005.

Kneale, William y Martha. *El desarrollo de la lógica* . Oxford, Londres: Clarendon, 1962.

Maggiori, Roberto. *La filosofía del día.* París: Flammarion, 1994.

Mu Soeng. *El Sutra del Diamante: transformando nuestra percepción del mundo* . Boston: Wisdom, 2000.

Rescher, Nicholas. *Paradojas: sus raíces, alcance y resolución.*

Suzuki, DT *La doctrina zen de la no-mente.* Ed. Christmas Humphreys. York Beach, ME.: Weiser, 1972.

Trungpa, Chögyam. *El juego de la ilusión.* Shambhala: Boston, 1994.

Obras de Avi Sion

Avi Sion es autor de varias obras sobre lógica y filosofía, entre ellas: *Lógica del futuro* (1990); *Lógica judía* (1995); *Ilógica budista* (2002); *Fenomenología* (2003); *La lógica de la causalidad* (1999, 2003, 2010); *Voluntad y conceptos causales afines* (2004); *Reflexiones* (2005); *Meditaciones* (2006); *Reflexiones lógicas y espirituales* (2008-9); *A Fortiori Logic* (2013).

Estas obras pueden estudiarse gratuitamente en: www.TheLogician.net y www.AviSionBooks.com.

Las paradojas y sus resoluciones pueden leerse en línea en: www.avisionparadoxes.bravesites.com/

Puede consultar la lista completa de sitios web de Avi Sion en: www.avisionlinks.bravesites.com.

Para cualquier pregunta o comentario, puede contactar con Avi Sion en la siguiente dirección de correo electrónico: avi-sion@thelogician.net

Referencias principales