

# **La construcción del espacio de posibilidad**

Agustín Rayo

Cátedra Gaos 2013

Instituto de Investigaciones Filosóficas, UNAM

*A mi abuela Carmen*

## Prefacio

---

Este libro es una versión editada y re trabajada de una serie de seis conferencias que impartí en el Instituto de Investigaciones Filosóficas de la UNAM en enero del 2013, como parte de la Cátedra Gaos.

Mucho del material filosófico de las conferencias proviene de mi libro *The Construcion of Logical Space* (Oxford University Press, 2013). Los textos recorren terrenos similares, pero creo que los recorridos son suficientemente diferentes para meritar publicación independiente.

El proyecto original fue un esfuerzo de siete años, y no estuvo libre de tropiezos. Cada capítulo es el resultado de una larga secuencia de ediciones y reediciones, y algunas de las cicatrices de ese forcejeo están todavía visibles en el texto. Cuando di la Cátedra, en cambio, tenía el material bien digerido desde el principio. Usé las conferencias para retomar las ideas desde una perspectiva más fresca, y para desarrollarlas sin los errores y los contratiempos que plagaron el proyecto original. Fue como visitar un país distante por segunda vez, con un buen mapa del territorio y con amigos locales.

El material de la Cátedra tuvo la ventaja adicional de haber sido moldeada por la comunidad filosófica del Instituto. Ese era el plan desde el principio, y fue por eso que decidí utilizar un formato libre para dar las conferencias. En lugar de presentar un texto prefabricado, o trabajar sobre la base de una colección fija de transparencias, decidí que las conferencias estarían guiadas por los intereses del público. Llegué a cada sesión con una idea general del tema a discutir, y con una guía resaltando puntos importantes, pero no mucho más. El rumbo de la discusión estuvo a cargo de la audiencia. Reconozco que fue una estrategia arriesgada, pero tuve la suerte de contar con una audiencia estupenda, y nuestras discusiones enriquecieron el material mucho más de lo que hubiera podido imaginar.

Cada uno de los seis textos que presento a continuación está basado en la grabación sonora de una de las conferencias. Francisco Ortega produjo excelentes transcripciones de las grabaciones, y yo trabajé sobre las transcripciones. Pulí la prosa, reorganicé el texto, corregí errores y reparé omisiones, pero hice siempre lo posible por respetar el espíritu de las grabaciones. El texto resultante tiene una estilo menos formal y una estructura más libre de la que tendría si sus orígenes no fueran orales, pero creo que retiene algo del flujo y la sencillez de una conversación entre amigos.

En algunos casos incorporé las preguntas del público al texto principal, junto con mis respuestas; en otros casos me pareció más iluminador retener el formato original del intercambio. En un esfuerzo por evitar que se interrumpiera el flujo de la exposición, simplifiqué la formulación de algunas de las preguntas, y cambié la ubicación de algunos de los intercambios, pero espero haber preservado la sustancia de la discusión original.

Tomé la decisión de no identificar a los autores de preguntas específicas. No quería que nadie se sintiera incómodo con mi reformulación de su pregunta, y no me parecía correcto atribuir preguntas sin darle a sus autores la oportunidad de hacer correcciones. Quisiera, sin embargo, agradecer a Atocha Aliseda, Axel Barceló, Maite Ezcurdia, Miguel Ángel Fernández, Claudia Lorena García, Eduardo García Ramírez, Ekain Garmendia, Edgar González, Guillermo Hurtado, Mariana Lojo, Raymundo Morado, Gustavo Ortiz, Elias Okon, Carlos Pereda, Pedro Ramos, Ricardo Salles, Miguel Ángel Sebastián, Pedro Stepanenko, Moisés Vaca, Margarita Valdés, y a todos aquellos quienes hicieron preguntas sin que sus nombres quedaran registrados en la grabación. Me gustaría mencionar especialmente a Víctor Cantero, Javier García Salcedo, Cristian Gutiérrez, Moisés Macías, Felipe Melahuac, y Carlos Romero, quienes organizaron un grupo de lectura paralelo a las conferencias del que me beneficié enormemente.

También quisiera agradecer a Manola Rius, y a su excelente equipo editorial; si más gente fuera como ellos, el mundo editorial sería un lugar mejor. Me gustaría mencionar, en particular, a Alejandro Magallanes por su finísimo trabajo gráfico en el diseño de los pósters que se utilizaron para difundir las conferencias.

Le debo agradecimientos especiales a Margarita Valdés y a todos aquellos quienes apoyaron mi nominación para la Cátedra; también a los miembros del Consejo Interno del Instituto, por aprobar la propuesta, y a Pedro Stepanenko y Eduardo García Ramírez por su generosidad y sabiduría durante el proceso entero. Haber recibido la Cátedra es un privilegio que nunca olvidaré.

Más todavía, quisiera agradecer a mi esposa Came, por su cariño y por su apoyo.

Cambridge, Massachusetts  
Junio del 2013

## Sesión 1: Introducción

*¿Qué es el espacio de posibilidad, y en qué sentido lo construimos?*

---

En la sesión de hoy me gustaría contestar a las preguntas “¿qué es el espacio de posibilidad?” y “¿en qué sentido lo construimos?”. También me gustaría discutir algunas de las herramientas filosóficas con las que vamos a estar trabajando en las sesiones siguientes.

¿Qué es el espacio de posibilidad? La manera breve de decirlo es que el espacio de posibilidad es el conjunto de alternativas con las que trabajamos cuando nos preguntamos cómo es el mundo.

Imaginen por ejemplo, que nos preguntamos si la nieve es blanca. Eso presupone una distinción entre dos maneras diferentes en las que puede ser el mundo: puede ser tal que la nieve es blanca o tal que la nieve no es blanca. Tenemos, pues, dos posibilidades diferentes: la posibilidad de que la nieve sea blanca y la posibilidad de que no sea el caso que la nieve es blanca. Cuando *aseveramos* ‘la nieve es blanca’, hacemos dos cosas: distinguimos entre estas dos posibilidades, y nos comprometemos con un lado de la distinción; específicamente, nos comprometemos con que el mundo corresponda a la posibilidad de que la nieve sea blanca, y no a la posibilidad de que no sea el caso que la nieve es blanca.<sup>1</sup>

Sobra decir que ‘la nieve es blanca’ no es la única aseveración que uno podría hacer. Uno podría aseverar también ‘el pasto es verde’, o ‘el Instituto es azul’, y cada una de estas aseveraciones presupone una distinción entre maneras de ser del mundo. El espacio de posibilidad es simplemente el conjunto de todas esas distinciones.

Identificamos una *región* de el espacio de posibilidad cuando identificamos ciertas distinciones y nos comprometemos con un lado de cada una de ellas. Aquí hay un ejemplo. Consideremos, por un lado, la distinción entre maneras de ser del mundo donde la nieve es blanca y maneras de ser del mundo donde no es el caso que sea blanca, y, por otro, la distinción entre maneras de ser del mundo donde el Instituto es azul y maneras de ser del mundo donde no es el caso que el Instituto es azul. Ahora tomemos partido con respecto a cada una de esas distinciones. Con respecto a la primera distinción, comprometámonos con la idea de que la nieve es blanca, y con respecto a esta distinción, comprometámonos con la idea de que el instituto es azul. El resultado es que identificamos una región del espacio de posibilidad. ¿Qué región? La región en donde la nieve es blanca y el Instituto es azul.

---

<sup>1</sup> Muchas de estas ideas vienen de Robert Stalnaker. Ver, por ejemplo, Stalnaker, Robert C. (1979). ‘Assertion’. *Syntax and Semantics*, 9, 315–322, reimpresso en Stalnaker, Robert C. (1999). *Context and Content*. Oxford University Press, Oxford. Ver también Stalnaker, Robert C. (1984). *Inquiry*. MIT Press, Cambridge, MA.

Identificamos un *punto* en el espacio de posibilidad cuando tomamos partido con respecto a todas las distinciones que constituyen el espacio de posibilidad. Los puntos son, por tanto, las regiones mínimas no vacías del espacio de posibilidad.

**Pregunta:** Los puntos del espacio de posibilidad, ¿son simplemente mundos posibles?

**Respuesta:** Sí. Lo que yo llamo ‘puntos en el espacio de posibilidad’ mucha gente en la literatura llama simplemente ‘mundos posibles’. La razón por la que uso terminología diferente aquí es que cuando la gente habla de mundos posibles suele traer bagaje filosófico que yo quiero evitar. Aquí quiero que empecemos de cero.<sup>2</sup>

\* \* \*

Tenemos, pues, una caracterización del espacio de posibilidad. Pasemos a la segunda pregunta. ¿En qué sentido *construimos* el espacio de posibilidad? Es, después de todo, natural pensar que el espacio de posibilidad está dado, y que en nada depende de nosotros.

La situación se complica, sin embargo, porque hay veces que no es obvio si lo que tenemos ante nosotros es una distinción o dos. Consideren, por ejemplo, las oraciones:

(1) Hay agua

(2) Hay H<sub>2</sub>O

(1) corresponde a la distinción entre aquellas maneras de ser del mundo de acuerdo con las cuales hay agua y aquellas de acuerdo con las cuales no es el caso que haya agua. (2) corresponde a la distinción entre aquellas maneras de ser del mundo de acuerdo con las cuales hay H<sub>2</sub>O y aquellas de acuerdo con las cuales no es el caso que hay H<sub>2</sub>O.

Es claro que ‘hay agua’ y ‘hay H<sub>2</sub>O’ son oraciones diferentes. Lo que no es obvio es si estas oraciones hacen la misma distinción entre maneras de ser del mundo, o si hacen distinciones diferentes. No es obvio si la región del espacio de posibilidad que identifican es la misma o no.

Dado lo que sabemos hoy día sobre química, es plausible suponer que (1) y (2) hacen la misma distinción. En otras palabras: que haya agua *es simplemente* que haya H<sub>2</sub>O; no hay *ninguna diferencia* entre un mundo con agua y un mundo con H<sub>2</sub>O; o si lo quieren pensar de una manera más colorida: cuando Dios creó al mundo y se aseguró

---

<sup>2</sup> La discusión canónica sobre posibilidad metafísica es Kripke, Saul A. (1980). *Naming and necessity*. Harvard University Press, Cambridge, MA. Muchas de los artículos seminales sobre mundos posibles pueden encontrarse en Loux, M. (ed.) (1979). *The Possible and the Actual*. Cornell University Press, Ithaca, NY.

de que hubiera agua no había nada *adicional* que tuviera que hacer o dejar de hacer para que hubiera H<sub>2</sub>O; y también al revés: cuando Dios se aseguró de que hubiera H<sub>2</sub>O no había nada *adicional* que tuviera que hacer para que hubiera agua.

(1) y (2) tienen *significados* diferentes, y hay por lo tanto un sentido en el que corresponden a distinciones diferentes. Pero eso no es el sentido de ‘distinción’ que estamos usando aquí. Aquí estamos preocupados no por maneras de *representar* el mundo, sino por maneras de *ser* del mundo. Lo que queremos hacer aquí es identificar regiones del espacio de posibilidad. Diremos, por tanto, que dos enunciados hacen la misma distinción cuando correspondan a la misma región del espacio de posibilidad.

Si *p* y *q* son enunciados diferentes, ¿bajo qué condiciones corresponden a la misma región del espacio de posibilidad? Según la propuesta que quiero defender aquí, corresponderán a la misma región del espacio de posibilidad siempre y cuando sea verdadero lo siguiente:

Que *p* sea el caso es simplemente que *q* sea el caso.

Llamaremos enunciados de este tipo ‘enunciados de identidad’. Sabemos, por ejemplo, que el siguiente enunciado de identidad es verdadero:

**[Agua]**

Que haya agua es simplemente que haya H<sub>2</sub>O.

Es por eso que digo que el enunciado ‘hay agua’ corresponde a la misma región del espacio de posibilidad que el enunciado ‘hay H<sub>2</sub>O’. Es por eso que los enunciados no hacen dos distinciones, sino una: hacen la misma distinción usando palabras diferentes.

Esto nos pone en posición de entender en qué sentido construimos el espacio de posibilidad. Para saber con qué espacio de posibilidad trabajar necesitamos determinar qué *enunciados de identidad* aceptar, pero hemos visto que no siempre es obvio qué enunciados de identidad son verdaderos. (No es obvio, por ejemplo, si debemos aceptar [Agua].) ‘Construimos’ el espacio de posibilidad cuando decidimos qué enunciados de identidad aceptar.

**Pregunta:** A mí no me queda claro que lo que estemos construyendo sea el espacio de posibilidad. ¿No es más bien que estamos construyendo una *concepción* del espacio de posibilidad?

**Respuesta:** De acuerdo. Lo único que he argumentado hasta el momento es que cuando aceptamos un conjunto de enunciados de identidad construimos nuestra *concepción* del espacio de posibilidad.

Una posición que no hemos excluido, y que bien podría ser la correcta, es que hay tal cosa como una concepción objetivamente correcta del espacio de

posibilidad. En ese caso, sería natural pensar que, a diferencia de nuestras *concepciones* del espacio de posibilidad, el espacio de posibilidad objetivamente correcto está dado. No depende de lo que hagamos nosotros, y, en particular, no es algo que construyamos.

Mi posición, sin embargo -- y eso es algo que vamos a discutir en la tercera sesión -- es que no hay tal cosa como un espacio de posibilidad objetivamente correcto. Podemos decir que un conjunto de distinciones es más o menos útil para describir el mundo, pero no tiene sentido hablar de distinciones 'objetivamente correctas'. De acuerdo con esta posición, cuando hablamos del espacio de posibilidad lo que estamos haciendo es, en realidad, hablar de la concepción del espacio de posibilidad que *nosotros* aceptamos. En ese sentido, no hay diferencia entre construir el espacio de posibilidad y construir nuestra concepción del espacio de posibilidad. Este es un tema delicado, que vamos a tener que discutir con cuidado más adelante.

**Pregunta:** Cuando aceptamos [Agua], ¿estamos simplemente diciendo que 'hay agua' implica a 'hay H<sub>2</sub>O', y viceversa?

**Respuesta:** Esa es una manera un poco peligrosa de poner el punto porque filósofos diferentes usan la palabra 'implica' de maneras diferentes. Una manera de interpretar 'implica' es como 'implica lógicamente', y eso no es lo que queremos en este contexto porque 'hay agua' y 'hay H<sub>2</sub>O' no son lógicamente equivalentes.<sup>3</sup> Otra interpretación de 'implica' es 'implica analíticamente', donde *p* implica analíticamente a *q* si *q* se sigue lógicamente de la conjunción de *p* y todas las verdades analíticas.<sup>4</sup> Eso tampoco funcionaría aquí. Lo que necesitamos una noción de 'implica' que corresponda a algo así como '*p* implica a *q* si todo mundo posible en el que *p* es verdadero es también un mundo posible en el que *q* es verdadero'. Podemos llamar a eso 'implicación metafísica'. Con esa terminología, lo que estoy diciendo es que para saber si (1) y (2) trazan la misma distinción lo que tenemos que hacer es determinar si son metafísicamente equivalentes.

**Pregunta:** ¿Vamos a incluir situaciones imposibles en el espacio de posibilidad?

---

<sup>3</sup> ¿Qué es consecuencia lógica? El texto canónico es Tarski, A., 1983c. "On the Concept of Logical Consequence", en Tarski, Alfred, 1956, *Logic, Semantics, Metamathematics: papers from 1923 to 1938* (traducción de J. H. Woodger), Oxford: Oxford University Press. Ver también Etchemendy, John, 1990, *The Concept of Logical Consequence*, Cambridge, MA: Harvard University Press, y Gomez-Torrente, Mario, 1996, "Tarski on Logical Consequence," *Notre Dame Journal of Formal Logic*, 37: 125–151.

<sup>4</sup> En la filosofía analítica, la discusión clásica sobre analiticidad es Quine, W. V. O. (1953), 'Two dogmas of empiricism', en Quine, W. V. O. *From a Logical Point of View*. Harper and Row, New York. Para una discusión más contemporánea de estos temas, ver Chalmers, David *Constructing the World*, Oxford University Press, Oxford, 2012.



**Respuesta:** No. Solo incluimos situaciones posibles. Parte de lo que estamos tratando de hacer aquí, sin embargo, es identificar la línea divisoria entre lo posible y lo imposible. Los enunciados de identidad que aceptemos van a determinar los límites de nuestro espacio de posibilidad. El enunciado de identidad 'que haya agua *es simplemente* que haya H<sub>2</sub>O' impone un límite a nuestro espacio de posibilidad porque nos dice que no hay tal cosa como una posibilidad en donde hay agua pero no H<sub>2</sub>O, o viceversa. Nos dice que una situación en la que hay agua pero no H<sub>2</sub>O queda fuera del espacio de posibilidad.

Para decidir si aceptar un enunciado de identidad como [Agua] parte de lo que tenemos que hacer es ir al mundo. No podemos conocer su verdad *a priori*. Esto significa que la línea que divide a la posibilidad de la imposibilidad es una que tampoco podemos conocer *a priori*. (De esto vamos a hablar mucho más en la tercera sesión.)

\* \* \*

Me gustaría darles algunos ejemplos adicionales de enunciados de identidad. Los ejemplos se van a ir poniendo más y más controversiales conforme vayamos avanzando, pero tomen en cuenta que en este momento no los estoy tratando de convencer de que acepten los enunciados. Por ahora lo único que quiero hacer es familiarizarlos con el operador 'es simplemente'. Más adelante voy a argumentar que algunos de estos enunciados son más plausibles de lo que uno podría pensar, pero por ahora no se preocupen de si son verdaderos o falsos.

Empiezo con un ejemplo poco controversial:

**[Hermano]**

Que Julián sea un hermano es simplemente que comparta un padre con alguna otra persona.

Este enunciado es interesante porque uno podría pensar que se trata de un caso que puede decidirse *a priori*. Más específicamente, uno podría pensar que para determinar si los enunciados 'Julián es un hermano' y 'Julián comparte un padre' corresponden a la misma distinción del espacio de posibilidad basta con que reflexionemos a la luz de nuestro dominio del lenguaje.

Segundo ejemplo:

**[Venus]**

Que Susana viaje a Héspero es simplemente que viaje a Fósforo

¿Por qué aceptar este enunciado? Todo depende de si uno cree que Héspero es Fósforo. Si uno cree que Héspero es Fósforo, cree que no hay manera de ir a Héspero sin ir a Fósforo, y que no hay manera de ir a Fósforo sin ir a Héspero. No solo eso, si

uno cree que Héspero es Fósforo, uno cree que si Dios crea un mundo en donde Susana va a Héspero, no tiene que hacer nada *adicional* para asegurarse que Susana vaya a Fósforo; ya había hecho todo lo que tenía que hacer. También al revés: uno cree que si Dios crea un mundo en donde Susana va a Fósforo, no tiene que hacer nada *adicional* para asegurarse que Susana vaya a Héspero; ya había hecho todo lo que tenía que hacer. Si, en cambio, uno cree que Héspero no es Fósforo, uno debe rechazar [Venus].

Una identidad de primer orden como 'Héspero = Fósforo' puede pensarse como un caso particular de lo que yo aquí estoy llamando 'enunciados de identidad'. Podemos pensar en 'Héspero = Fósforo' como equivalente a 'ser idéntico a Héspero *es simplemente* ser idéntico a Fósforo'.<sup>5</sup> Esto significa que los enunciados de identidad nos permiten meter en la misma bolsa a ciertas analiticidades, como en el caso de [Hermano], y a ciertas identidades de primer orden, como en el caso de [Venus].

Siguiente ejemplo:

### **[Calor]**

Que algo esté caliente es simplemente que su energía cinética promedio sea alta.

Me aseguran mis amigos que saben de estas cosas que de acuerdo con la ciencia contemporánea [Calor] no es estrictamente verdadero. Resulta que no es más que un ejemplo de juguete, que usan los filósofos. Desafortunadamente, yo no sé cuál es el principio verdadero, así que yo también voy a usar el ejemplo de juguete.

[Calor] nos da otra manera de ilustrar el hecho de que nuestra aceptación de un enunciado de identidad depende de las teorías científicas que aceptamos. Hubo un momento en el que se creía que estar caliente consistía en tener altas cantidades de una cierta sustancia llamada 'calórico'. Una persona que crea eso no va a aceptar [Calor]. Quien cree que estar caliente consiste en tener suficiente calórico, cree que, en principio, podría haber un objeto que esté hecho de partículas con energía cinética bien alta, pero que tenga poco calórico, y que, por lo tanto, no esté caliente. A alguien que acepta [Calor], en cambio, le parecería absurdo pensar que podría haber un objeto hecho de partículas con energía cinética bien alta que no esté caliente. Una situación como esa está fuera del espacio de posibilidad.

En general, uno de los efectos de aceptar un enunciado de identidad es que excluimos situaciones de nuestro espacio de posibilidad. Esta exclusión tiene ventajas y desventajas. La ventaja es que hay ciertas preguntas que no tenemos que contestar. Por ejemplo, alguien que acepta [Calor] es alguien que no tiene que contestar a la pregunta siguiente: 'Estoy de acuerdo en que las partículas de este objeto tienen energía cinética bien alta, pero ¿por qué está caliente?' Uno puede simplemente

---

<sup>5</sup> Aquí presupongo que lo que tenemos en mente es el sentido débil de identidad, de acuerdo con el cual 'Héspero = Fósforo' no implica 'Héspero existe'. Cuando 'Héspero = Fósforo' se entiende en sentido fuerte, es equivalente a la conjunción de 'ser idéntico a Héspero *es simplemente* ser idéntico a Fósforo' y 'Héspero existe'.

rechazar la pregunta diciendo ‘que el objeto esté caliente *es simplemente* que estén en movimiento sus partículas’.<sup>6</sup> La desventaja de aceptar un enunciado de identidad es que uno pierde recursos teóricos, porque uno se queda con menos posibilidades con las cuales trabajar. Como hemos visto, quien acepta [Calor] no puede apelar a situaciones en las que un objeto está caliente sin tener energía cinética alta. En algunas las sesiones siguientes voy a argumentar que la manera de decidir qué enunciados de identidad aceptar es sopesando estas ventajas y desventajas.

**Pregunta:** ¿Qué quiere decir que haya una situación que está fuera del espacio de posibilidad?

**Respuesta:** Lo único que limita al espacio de posibilidad es el absurdo. Una situación que está fuera del espacio de posibilidad es una situación absurda: por ejemplo, una situación en la que Susana viaja a Héspero (es decir, Venus) pero no a Fósforo (es decir, Venus).

Otro ejemplo. Supongamos que uno acepta [Calor], y cree que no hay diferencia entre estar caliente y tener energía cinética promedio alta. Uno piensa, por tanto, que lo siguiente es un absurdo: una situación en la que hay una porción de agua cuyas sus partículas tienen energía cinética es bien alta, pero que no está caliente. Estamos ante una situación en la que la porción tiene energía cinética promedio alta (es decir, está caliente) pero, por otro lado, no está caliente, y eso es un absurdo.

Algo que es importante subrayar es que los enunciados que describen una situación como la que acabamos de discutir son metafísicamente inconsistentes, pero no son *lógicamente* inconsistentes o *analíticamente* inconsistentes. La inconsistencia metafísica no siempre puede reconocerse *a priori*.

**Pregunta:** ¿Podríamos entonces tener enunciados coherentes que no correspondan a ninguna región del espacio de posibilidad?

**Respuesta:** Sí, en el sentido siguiente: podemos tener representaciones bien formadas y significativas que no representen una situación metafísicamente consistente. El enunciado ‘este vaso contiene agua pero no H<sub>2</sub>O’ está bien formado y es perfectamente significativo, pero es metafísicamente inconsistente. En ese sentido, los límites del significado van más allá de los límites de la posibilidad metafísica.

\* \* \*

---

<sup>6</sup> Hay una magnífica discusión sobre este tema en Block, Ned and Stalnaker, Robert (1999). ‘Conceptual analysis, dualism, and the explanatory gap.’ *Philosophical Review*, **108**, 1–46.

El siguiente ejemplo en nuestra lista de enunciados de identidad es un caso particular de la posición que suele llamarse ‘fiscalismo’. A saber:

**[Rojo]**

Que María experimente la sensación de ver rojo es simplemente que María esté en el estado cerebral R.

Quien acepta este enunciado identifica una propiedad fenoménica (la propiedad de experimentar la sensación de ver rojo) con una propiedad cerebral (la propiedad estar en el estado cerebral R).

Hay quien ha argumentado que [Rojo] falso. El ejemplo canónico viene de Frank Jackson.<sup>7</sup> Imaginemos que María creció desde pequeña en un mundo blanco y negro. (Tal vez creció con guantes blancos para nunca verse las manos, y sin espejos para nunca verse la cara.) María tiene acceso, sin embargo, a una enorme colección de libros blancos escritos con tinta negra, y leyendo esos libros se convirtió en una gran científica del color. De hecho, aprendió todo lo que la ciencia del color podría enseñar.

Ahora imaginemos que se le informa a María lo siguiente: ‘A las doce del día te vamos a mostrar un tomate maduro. Vas a experimentar la sensación de ver rojo, y vas a estar en el estado cerebral R, y va a ser maravilloso’. Es tentador pensar que cuando María finalmente mira el tomate, aprende algo; específicamente, aprende *qué es lo que se siente ver rojo*.

Algunas personas han concluido a partir de este experimento mental que [Rojo] tiene que ser falso. Sabemos, por un lado, que María poseía desde el principio la información de que a las doce iba a estar en el estado cerebral R, y, por otro, que María aprende algo cuando, a las doce, experimenta la sensación de ver rojo. Uno podría pensar que de esto se sigue que rechazar la identidad entre estar en el estado cerebral R y experimentar la sensación de ver rojo.

Esto nos permite ilustrar el hecho de que aceptar un enunciado de identidad tiene ventajas y desventajas. ¿Qué desventajas tendría la aceptación de [Rojo]? La desventaja inmediata es que quien acepta [Rojo] se enfrenta a un problema cuando trata de dar cuenta de la situación cognitiva de María. Si experimentar la sensación de ver rojo es simplemente estar en un cierto estado cerebral, y si María ya sabía que iba a estar en ese estado cerebral, ¿cómo es posible que haya aprendido algo cuando le enseñaron el tomate?

No digo que no se pueda contestar a esta pregunta. (De hecho, mi opinión es que debemos aceptar [Rojo], y en la quinta sesión voy a argumentar que un amigo de [Rojo] está en posición de modelar el logro cognitivo que experimenta María cuando mira el tomate.) Concedo, sin embargo, que no es nada obvio que esto se pueda hacer.

---

<sup>7</sup> Ver, por ejemplo, Jackson, Frank (1982). ‘Epiphenomenal qualia’. *Philosophical Quarterly*, 32, 127–136. Hay una buena lista de referencias bibliográficas sobre el tema en Byrne, Alex (2006). ‘Review of *There’s Something about Mary*’. *Notre Dame Philosophical Reviews*, 2006.01.20. Disponible en <<http://ndpr.nd.edu/review.cfm?id=5561>>.

Es más natural pensar que lo que tenemos que hacer para modelar el logro cognitivo de María es rechazar [Rojo], porque esto nos pondría en posición de decir que hay situaciones metafísicamente consistentes en la que María está en el estado cerebral R sin experimentar la sensación de ver rojo. Dado que situaciones de este tipo son metafísicamente consistentes, corresponden a una cierta región del espacio de posibilidad. Esto nos permite decir que lo que aprende María cuando mira al tomate es que la región relevante del espacio de posibilidad no está actualizada. La región actualizada corresponde a posibilidades en las que María está en el estado cerebral R, y sí experimenta la sensación de ver rojo. Tenemos, por tanto, un método inmediato para modelar del logro cognitivo de María.<sup>8</sup>

La conclusión de todo esto es que si rechazamos [Rojo] incluimos posibilidades adicionales en nuestro espacio de posibilidad, y esto nos da recursos teóricos que podrían ser útiles para modelar el logro cognitivo de María. Desafortunadamente, hay otro lado de la moneda. Quien rechaza [Rojo] está obligado a contestar preguntas como: ¿cómo es que el estado cerebral R corresponde a la fenomenología que de hecho experimentamos cuando vemos al tomate, y no a alguna otra? En cambio, quien acepta [Rojo] está en posición de rechazar la legitimidad de esa pregunta. Podría decir que se trata de una pregunta tan ilegítima como la pregunta ‘¿cómo es que las cosas hechas de H<sub>2</sub>O están hechas de agua?’.

¿Estamos ante un caso en el que las ventajas de aceptar un enunciado de identidad son suficientes para compensar las desventajas? La respuesta no es nada obvia. De hecho, una de las discusiones principales en la filosofía de la mente contemporánea es si uno debe aceptar enunciados como [Rojo]. Nos gustaría poder aceptar [Rojo], para no tener que contestar a la pregunta de por qué no está invertida nuestra fenomenología del color (o a la pregunta de por qué no somos zombies).<sup>9</sup> Pero para tener el lujo de rechazar esas preguntas tenemos que explicar en qué consiste el logro cognitivo de María, y no es nada obvio que quien acepta [Rojo] lo pueda hacer.

La conclusión es que no solo es imposible separar el proyecto de generar una concepción del espacio de posibilidad de nuestro quehacer científico, también es imposible separarlo de nuestro quehacer filosófico. Para decidir con qué posibilidades trabajar tenemos que hacer ciencia, pero también tenemos que hacer filosofía.

\* \* \*

Me gustaría que consideráramos otros dos enunciados de identidad, y estos sí que son bien controversiales. Los escogí como ejemplos de que nuestra concepción del espacio de posibilidad puede estar entrelazada con nuestras teorías metafísicas. Aquí tienen el primer ejemplo:

---

<sup>8</sup> Hay buenas razones, sin embargo, para pensar que este modelo tiene que estar equivocado. Ver Lewis, David (1988). ‘What experience teaches’. *Proceedings of the Russellian Society*, 13, 29–57. Reimpreso en Lewis, David (1999). *Papers in Metaphysics and Epistemology*. Cambridge University Press, Cambridge.

<sup>9</sup> Famosamente, David Chalmers rechaza enunciados como [Rojo]. Ver, por ejemplo, Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press, New York.

### **[Muerte]**

Que ocurra la muerte de Sócrates *es simplemente* que Sócrates muera

Supongamos que Dios está creando el mundo y se asegura de que exista Sócrates y de que muera. Alguien podría pensar que eso no garantiza que exista la muerte de Sócrates, porque para eso haría falta que Dios abriera 'el cajón de los eventos', escogiera uno de ellos -- una muerte -- y lo agregara al mundo. Quien adopte esta posición tendrá que rechazar [Muerte], porque lo que dice [Muerte] es que no hay diferencia entre que muera Sócrates y que exista la muerte de Sócrates. Alguien que acepta [Muerte] cree que cuando Dios se aseguró de que Sócrates muriera, ya había hecho todo lo que tenía que hacer para que fuera el caso que existe la muerte de Sócrates. También al revés: quien acepta [Muerte] cree que cuando Dios se aseguró de que existiera la muerte de Sócrates, ya había hecho todo lo que tenía que hacer para que Sócrates muriera.

Aquí está el segundo ejemplo:

### **[Mesa]**

Que haya una mesa es simplemente que haya partículas organizadas mesísticamente

Una vez más, alguien que rechaza [Mesa] diría que cuando Dios se asegura de que el mundo contenga partículas organizadas mesísticamente, esto no fue suficiente para que hubiera una mesa. Para eso haría falta que Dios le agregara al mundo un objeto adicional, un objeto que esté 'constituido' por las partículas organizadas mesísticamente. Alguien que acepte el [Mesa], en cambio, diría que no hay diferencia entre que haya partículas organizadas de la manera correcta y que haya una mesa.

Hay discusiones en metafísica que *presuponen* que rechazamos enunciados de identidad como [Mesa]. Famosamente Peter van Inwagen ha defendido una posición de acuerdo con la cual existen partículas ordenadas mesísticamente, pero no mesas.<sup>10</sup> Otros han estado de acuerdo con van Inwagen en que la existencia de partículas organizadas mesísticamente no basta para que exista una mesa, pero, a diferencia de él, creen que sí hay mesas. Creo que un defecto de esta discusión es que todo el mundo rechaza [Mesa], y no se dice mucho acerca de por qué. Mi sospecha es que, de hecho, hay buenas razones para aceptar [Mesa], y que la discusión de si aceptar la existencia de mesas además de partículas organizadas mesísticamente es un pseudo-problema.

\* \* \*

Hay un último ejemplo en mi lista de enunciados de identidad:

### **[Dinosaurios]**

---

<sup>10</sup> van Inwagen, Peter (1990). *Material Beings*. Cornell University Press, Ithaca, NY.

Que el número de los dinosaurios sea cero es simplemente que no haya dinosaurios.

Alguien que rechaza [Dinosaurios] es alguien que cree que cuando Dios hizo al mundo y se aseguró de que no hubiera dinosaurios, eso no bastó para que el número de los dinosaurios fuera cero. Para eso hacía falta que Dios, además, abriera el cajón de los números, seleccionara al cero, y se lo agregara al mundo. Alguien que acepta [Dinosaurios], en cambio, cree que cuando Dios hizo al mundo y se aseguró de que no hubiera dinosaurios, ya había hecho todo lo que tenía que hacer para que fuera el caso que el número de los dinosaurios es cero.

Según Quine, los enunciados 'no hay dinosaurios' y 'el número de los dinosaurios es cero' tienen compromisos ontológicos muy diferentes, porque solo el segundo está comprometido con la existencia de números.<sup>11</sup> Yo lo que creo es que, a pesar de eso, debemos aceptar [Dinosaurios]. De hecho, creo que debemos aceptar un principio más general:

### [Números]

Que el número de las Fs sea  $n$  es simplemente que haya  $n$  Fs.<sup>12</sup>

Defender [Números] es uno de los objetivos principales de estas conferencias.

Quien acepta [Números] acepta una posición de acuerdo con la cual la verdad de los enunciados matemáticos no requiere nada del mundo; en otras palabras, quien acepta [Números] acepta una posición de acuerdo con la cual la negación de una verdad aritmética describe una situación *absurda*, o metafísicamente inconsistente.

De hecho, es muy fácil demostrar que quien acepta [Números] debe creer que un mundo sin números sería absurdo:

*Prueba:* supongamos, para reducción al absurdo, que no hay números. Dado [Números], que el número de los números sea cero es simplemente que no haya números. Por tanto, el número de los números es cero. Por tanto, el cero existe. Por tanto, hay números. Contradicción.

**Pregunta:** Pareciera ser el caso que los enunciados ' $2+2=4$ ' y ' $500+500=1000$ ' son verdaderos en exactamente los mismos puntos del espacio de posibilidad. ¿Tenemos que aceptar 'que sea el caso que  $2+2=4$  es simplemente que sea el caso que  $500+500=1000$ '?

**Respuesta:** Yo pienso que sí. La razón por la que no hay diferencia entre una manera de ser del mundo de acuerdo con la cual  $2+2=4$  y una manera de ser

---

<sup>11</sup> Quine, Willard Van Orman (1948). 'On What There Is', *Review of Metaphysics*, 2. Reprinted in Quine, Willard Van Orman (1953) *From a Logical Point of View*. Harvard University Press, Cambridge, MA.

<sup>12</sup> Aquí estoy pensando en [hay  $n$  Fs] como una abreviación de  $[\exists!_n x (Fx)]$ .

del mundo de acuerdo con la cual  $500+500=1000$  es que ni las condiciones de verdad de ' $2+2=4$ ' ni las condiciones de verdad de ' $500+500=1000$ ' le imponen restricciones no-triviales al mundo; en ambos casos, las condiciones de verdad quedarían satisfechas por cualquier manera no absurda en la que el mundo podría ser.

Sería un error concluir de esto que todas las verdades matemáticas *significan* lo mismo. Esa es una tesis diferente, y la diferencia es crucial. Sería insostenible una filosofía de la matemática de acuerdo con la cual ' $2+2=4$ ' y ' $500+500=1000$ ' jueguen el mismo papel en nuestras prácticas comunicativas, o una filosofía de la matemática de acuerdo con la cual 'Susana cree que  $2+2=4$ ' implica 'Susana cree que  $500+500=1000$ '.

En la quinta sesión vamos a discutir este punto. Voy a tratar articular una filosofía de las matemáticas que acepte [Números], y que, al mismo tiempo, nos permita hacer justicia a las diferencias entre ' $2+2=4$ ' y ' $500+500=1000$ '.

\* \* \*

He estado argumentando que nuestra concepción del espacio de posibilidad no es independiente de nuestras mejores hipótesis acerca de cómo es el mundo. Esto es porque, por un lado, nuestra concepción del espacio de posibilidad depende de los enunciados de identidad que aceptemos, y, por otro, nuestras decisiones acerca de qué enunciados de identidad aceptar no pueden separarse del resto de nuestra investigación acerca de cómo es el mundo.

Voy a articular esta idea con mucho más detalle más adelante, pero por ahora me gustaría contrastar la posición que estoy defendiendo aquí con otras posiciones filosóficas que ustedes conocen. Con eso termino la carla de hoy.

Las posiciones filosóficas con las que me gustaría contrastar la mía son las de Carnap y Quine.<sup>13</sup> De hecho, posiciones reales son mucho más sutiles de lo que vamos a discutir aquí, pero los detalles no importan para nuestros propósitos actuales así que vamos a discutir versiones simplificadas de sus posiciones.

Carnap -- o, mejor, la versión simplificada de Carnap -- creía que tenemos una comprensión robusta de la noción de analiticidad. Eso le permitía pensar en una región del espacio de posibilidad como una descripción del mundo 'analíticamente consistente', es decir, una descripción del mundo que no contradice ninguna verdad analítica. De acuerdo con esta posición, basta con que seamos usuarios competentes de nuestro lenguaje -- y que, en virtud de nuestra competencia lingüística, conozcamos la verdad de las verdades analíticas -- para que sepamos en qué consiste el espacio de posibilidad. Esto significa que somos capaces de conocer el espacio de

---

<sup>13</sup> Aquí los textos canónicos son Carnap, Rudolf *La Construcción Lógica del Mundo*, IIF-UNAM, y Quine, W. V. O. (1953), 'Two dogmas of empiricism', en Quine, W. V. O. *From a Logical Point of View*. Harper and Row, New York.



posibilidad *a priori*. El trabajo empírico viene después, cuando vamos al mundo y tratamos de determinar qué posibilidades están actualizadas, y cuales pueden ser descartadas como falsas.

Sabemos, por ejemplo, que es analíticamente consistente que la nieve sea azul y que es analíticamente consistente que la nieve sea blanca. Nuestro Carnap simplificado diría, por tanto, que hay regiones correspondientes en el espacio de posibilidad. El paso siguiente es que vayamos al mundo, y llevemos a cabo una investigación empírica decidir cuál de estas posibilidades está actualizada.

El resultado de todo esto es que, de acuerdo con Carnap simplificado, nuestro quehacer científico se puede dividir limpiamente en dos; tenemos, por un lado, un componente *a priori*, que consiste en fijar una concepción del espacio de posibilidad, y, por otro, un componente *a posteriori* que consiste en decidir, de las posibilidades antes identificadas, cuáles son compatibles con la manera en la que de hecho es el mundo y cuáles no.

Consideremos ahora la crítica de Quine. La idea, en su versión simplificada, es que no tenemos una noción de analiticidad suficientemente robusta como para hacer el trabajo teórico que Carnap necesita, y que, por tanto, no estamos en posición de definir el espacio de posibilidad a la manera de Carnap.

Yo lo que estoy proponiendo aquí es una especie de carnapeanismo post-quineneano. Quiero retener la idea carnapiana de que trabajamos con un espacio de posibilidad, pero en lugar de pensar en las regiones del espacio de posibilidad como correspondiendo a descripciones del mundo *analíticamente* consistentes, quiero pensar en ellas como correspondiendo a descripciones del mundo *metafísicamente* consistentes. Mientras que Carnap quería construir el espacio de posibilidad a partir del conjunto de verdades analíticas, yo lo quiero construir a partir del conjunto de enunciados de identidad verdaderos.

Con esta manera de proceder, hay algo que ganamos y algo que perdemos. Lo que ganamos es que hacemos justicia al punto de Quine, porque nuestra propuesta no presupone la noción de analiticidad. Lo que perdemos es que la posición de Carnap nos permite determinar *a priori* cuál es el espacio de posibilidad; nos permite, en otras palabras, establecer *a priori* los límites del absurdo. Para nosotros no es así. Dado que nuestro espacio de posibilidad está construido a partir de los enunciados de identidad verdaderos -- y dado que no podemos, en general, decidir si aceptar un enunciado de identidad independientemente de nuestro quehacer científico y filosófico -- no podemos decidir *a priori* cómo construir el espacio de probabilidad. De acuerdo con la propuesta que estoy defendiendo aquí, todo está mezclado. Cuando teorizamos acerca del mundo desarrollamos una concepción del espacio de posibilidad y, al mismo tiempo, tratamos de determinar qué regiones del espacio de posibilidad son compatibles con la manera en la que realmente es el mundo.

**Pregunta:** Nuestro lenguaje está sujeto a variación contextual. ¿Qué efecto tiene esto en la propuesta?

**Respuesta:** La complica. En un mundo sin variación contextual, un enunciado determina de una vez por todas una división del espacio de posibilidad; pero lo que sucede en vista del fenómeno de la variación contextual es que en contextos diferentes el mismo enunciado divide al espacio de posibilidad de maneras diferentes.

**Pregunta:** ¿Cómo están relacionadas las regiones de tu espacio de posibilidad con las posibilidades metafísicas de las que hablaba Kripke?

**Rayo:** Al final van a coincidir. Hay discusiones en las que se apela, sin más, a 'intuiciones filosóficas' para decidir qué es metafísicamente posible y qué no. Aquí, en cambio, vamos a basar nuestros juicios sobre posibilidad en juicios sobre qué enunciados de identidad aceptar, y la decisión de qué enunciados de identidad aceptar va a estar basada en nuestro quehacer científico (y en nuestro quehacer filosófico no específicamente modal).

En la última sesión vamos a ver que uno puede probar un teorema que muestra que una vez que uno decide qué enunciados de identidad aceptar, eso puede utilizarse para fijar los valores de verdad de todos los enunciados de un lenguaje modal de primer orden. Eso significa que hay un sentido muy preciso en el que los enunciados de identidad nos dicen todo lo que necesitamos saber para fijar los límites de posibilidad metafísica.

Consideremos un enunciado modal como 'es esencial a Agustín tener a Julieta como madre'. Hubiéramos podido expresar la misma idea utilizando un enunciado diferente: 'parte de lo que es ser Agustín es tener a Julieta como madre'. Lo interesante es que el operador 'parte de lo que es' se puede definir sobre la base del operador 'es simplemente'. En general 'parte de lo que es ser F es ser G' se puede definir como 'ser F es simplemente (ser G y ser F)'. El resultado es que podemos expresar 'parte de lo que es ser Agustín es tener a Julieta como madre' usando solamente enunciados de identidad, y eso basta para capturar el hecho de que tener a Julieta como madre es esencial a quien soy yo. Es gracias a esta observación que el teorema funciona.

**Pregunta:** No entiendo cómo es que los límites del espacio de posibilidad pueden fijarse utilizando solo enunciados de identidad. 'Los felinos son mamíferos' no es un enunciado de identidad, pero podríamos necesitarlo para fijar los límites del espacio de posibilidad.

**Respuesta:** Habría que utilizar el enunciado 'parte de lo que es ser felino es ser mamífero', que es, por definición, equivalente a 'ser un felino es simplemente

(ser un felino y ser un mamífero)'. En otras palabras, cuando Dios hace el mundo y se asegura de que Félix sea un felino ya ha hecho todo lo que tiene que hacer para que Félix sea un mamífero. Vamos a hablar mucho más de estos temas en la última sesión.

## Sesión 2: Lenguaje y mundo

¿Qué tipo de correspondencia existe entre nuestro lenguaje y la realidad que representa?

---

Quisiera empezar recordándoles algunos ejemplos de lo que la vez pasada llamamos 'enunciados de identidad'. Aquí está el primero:

### [Hermano]

Que Julián sea un hermano *es simplemente* que Julián comparta un padre con alguna otra persona

¿Qué quiere decir esto? Que no hay *ninguna diferencia* entre lo que se requiere del mundo para que sea verdad 'Julián es un hermano' y lo que se requiere del mundo para que sea verdad 'Julián comparte un padre'. Una forma colorida de poner el punto es diciendo que cuando Dios hizo al mundo y se aseguró de que Julián fuera un hermano, ya había hecho todo lo que tenía que hacer para que fuera el caso que Julián compartiera un padre. No hay nada *adicional* que Dios tuviera que hacer, o dejar de hacer. Y lo mismo al revés: si Dios se asegura de que Julián comparta un padre, no hay más por hacer con respecto al proyecto de asegurarse de que Julián sea un hermano.

[Hermano] es un enunciado de identidad en el que, plausiblemente, se describe una conexión analítica entre los enunciados involucrados, pero yo estoy pensando en el operador 'es simplemente' como algo más general. Creo, por ejemplo, que es verdad lo siguiente:

### [Agua]

Que el vaso esté lleno de agua es simplemente que el vaso esté lleno de  $H_2O$

En otras palabras: lo que se requiere del mundo para que sea verdad que 'el vaso está lleno de agua', es exactamente lo mismo que lo que se requiere del mundo para que sea verdad 'el vaso está lleno de  $H_2O$ '. Esta no es, sin embargo, una conexión analítica, y no es una conexión que uno esté en posición de establecer *a priori*. Para decidir si aceptar [Agua], uno tiene que ir al mundo y llevar a cabo una investigación empírica.

No solo eso. Tal y como estoy utilizando el operador 'es simplemente' aquí, puede ser verdadero [que  $p$  sea el caso *es simplemente* que  $q$  sea el caso] aunque  $p$  contenga

términos que refieren a ciertos objetos y  $q$  no tenga términos que refieran a esos objetos.<sup>14</sup> Aquí hay un ejemplo:

**[Muerte]**

Que ocurra la muerte de Sócrates es simplemente que Sócrates muera

Alguien que acepta [Muerte] cree que lo que se requiere del mundo para que ‘Ocurre la muerte de Sócrates’ sea verdadero es exactamente lo que se requiere para que ‘Sócrates muere’ sea verdadero. No es que cuando Dios se haya asegurado de que Sócrates muera falte algo por hacer para que ocurra la muerte de Sócrates; no es que, habiéndose asegurado de que Sócrates muera, Dios tuviera que abrir el cajón de sucesos y elegir uno de ellos -- la muerte de Sócrates -- para agregarlo al mundo.

Alguien que acepta [Muerte] es alguien que cree que dos enunciados pueden describir el mismo hecho a pesar de diferir en sus compromisos ontológicos, en el sentido de Quine.<sup>15</sup> El enunciado ‘ocurre la muerte de Sócrates’ incluye un término – ‘la muerte de Sócrates’ – que refiere a un suceso: la muerte de Sócrates. Por lo tanto, si preguntamos por sus compromisos ontológicos, la respuesta incluiría sucesos. En cambio, el enunciado ‘Sócrates muere’ no parece incluir ninguna expresión que refiera a sucesos. Por lo tanto, si preguntamos por sus compromisos ontológicos, no esperaríamos que la respuesta mencionara sucesos.

¿Cómo es esto posible? ¿Cómo es posible que dos enunciados describan el mismo hecho a pesar de diferir en sus compromisos ontológicos quineanos?

Creo que aquí podemos utilizar una idea de Frege: la idea de que un mismo hecho puede conceptualizarse de maneras diferentes.<sup>16</sup> Tomemos el hecho de que Sócrates muere. Una manera de conceptualizar este hecho es dividiéndolo en el objeto Sócrates y la propiedad de morir. Lo dividimos de esta manera cuando lo describimos utilizando el enunciado ‘Sócrates muere’. Pero otra manera de conceptualizar ese mismo hecho es dividiéndolo en el suceso de la muerte de Sócrates y la propiedad de ocurrir. Así es como lo dividimos cuando lo describimos utilizando el enunciado ‘Ocurre la muerte de Sócrates’.

El objetivo de la sesión de hoy y, de mucho del material en las sesiones siguientes, es defender este tipo de posición.

---

<sup>14</sup> Aquí las esquinas ‘[’ y ‘]’ se utilizan para expresar lo que Quine llamó ‘cuasi-encomillado’ (‘quasi-quotation’). Si ‘ $x$ ’ es una variable que varía sobre expresiones lingüísticas, la expresión ‘[ $x$  es tal-y-tal]’ refiere al enunciado que resulta de concatenar el valor de ‘ $x$ ’ con la expresión ‘es tal-y-tal’. Esto es útil porque permite el equivalente a cuantificadores que ligan variables que están dentro de contextos encomillados. Uno puede decir, por ejemplo, ‘existe un nombre  $c$  tal que el enunciado [ $c$  es rojo] es verdadero’. Ver Quine, W. V. [1940]. *Mathematical Logic*. Cambridge, MA: Harvard University Press.

<sup>15</sup> Quine, Willard Van Orman (1948). ‘On What There Is’, *Review of Metaphysics*, 2. Reprinted in Quine, Willard Van Orman (1953) *From a Logical Point of View*. Harvard University Press, Cambridge, MA.

<sup>16</sup> Ver, Frege, Gottlob (1884). *Los principios fundamentales de la aritmética*, sección 64, y Frege, Gottlob (1892). ‘Sobre Concepto y Objeto’. *Vierteljahrsschrift für wissenschaftliche Philosophie*, 16, 192–205.

**Pregunta:** Me parece que hay una lectura de ‘que ocurra la muerte de Sócrates *es simplemente* que Sócrates muera’ que presupone que ‘Sócrates muere’ es, en algún sentido, más fundamental que ‘Ocurre la muerte de Sócrates’. ¿No es así?

**Respuesta:** Absolutamente. Hay un uso de ‘es simplemente’ que presupone que cuando tenemos dos enunciados vinculados con ‘es simplemente’ el de la derecha ‘explica’ al de la izquierda, o es en algún sentido ‘más fundamental’. Ese no es, sin embargo, el uso con el que yo quiero trabajar aquí. Aquí me quiero mantener totalmente neutral acerca de si uno de los enunciados explica al otro, o es más fundamental.

Una manera de forzar la lectura simétrica con la que yo quiero trabajar es utilizando ‘no hay diferencia’ en vez de ‘es simplemente’. Uno podría decir, por ejemplo ‘*no hay diferencia* entre que muera Sócrates y que ocurra su muerte’. Sería, por lo tanto, menos riesgoso utilizar ‘es simplemente’, pero a mí me parece más natural decir ‘es simplemente’, así que es esa la locución que voy a utilizar aquí.

Si alguien que quisiera trabajar con la lectura asimétrica de ‘es simplemente’, yo lo recomendaría es hacer la asimetría explícita, para evitar confusión. Si uno quisiera, por ejemplo, trabajar con la lectura fundamentalista de [Muerte] yo lo que recomendaría es decir algo así como ‘que ocurra la muerte de Sócrates es, dicho en términos más fundamentales, que Sócrates muera’.

**Pregunta:** Cuando hablamos de enunciados de identidad, ¿estamos hablando de identidad a secas?

**Respuesta:** El predicado de identidad ‘=’ que utilizamos en los cursos de lógica de la licenciatura toma términos singulares como argumentos, y yo no quiero presuponer aquí que ‘que Sócrates muera’ es un término singular. No quiero comprometerme con que un enunciado como [Muerte] pueda parafrasearse como:

que ocurra la muerte de Sócrates = que Sócrates muera

porque no quiero comprometerme con que ese enunciado sea gramatical.

Si uno quisiera formalizar el tipo de enunciados que estamos considerando aquí, creo que lo más prudente sería introducir un nuevo símbolo; por ejemplo, ‘≡’. A diferencia del símbolo de identidad usual, el símbolo ‘≡’ no es un predicado, sino un operador. Podemos, por tanto, formalizar [Muerte] como:

Ocurre la muerte de Sócrates ≡ Sócrates muere

Una ventaja de esta notación es que puede generalizarse. Nos permite expresar, por ejemplo, 'estar hecho de agua es simplemente estar hecho de H<sub>2</sub>O' como:

$x$  está hecho de agua  $\equiv_x$   $x$  está hecho de H<sub>2</sub>O

Para entender qué papel juega 'x' en ' $\equiv_x$ ' vale la pena contrastar el enunciado anterior con este:

$x$  está hecho de agua  $\equiv$   $x$  está hecho de H<sub>2</sub>O

Este segundo enunciado se leería 'que *él* esté hecho de agua es simplemente que *él* esté hecho de H<sub>2</sub>O', en donde 'él' es un pronombre al que no se le ha asignado referencia. (Si alguien preguntara, '¿quién es él?', no habría una buena respuesta: 'él' está funcionando como variable libre.)

Cuando usamos ' $\equiv_x$ ', en cambio, todas las ocurrencias de la variable 'x' quedan ligadas. El resultado es que en lugar expresar un pensamiento incompleto (que *él* esté hecho de agua es *simplemente* que *él* esté hecho de H<sub>2</sub>O), expresamos un pensamiento completo (estar hecho de agua es *simplemente* estar hecho de H<sub>2</sub>O).

De hecho, podemos utilizar el operador ' $\equiv$ ' para ligar más de una variable. Por ejemplo, el enunciad de identidad

$x$  es hermano de  $y \equiv_{x,y}$  ( $x \neq y$  &  $x$  comparte un padre con  $y$ )

puede leerse 'que individuos sean hermanos es simplemente que sean distintos y compartan un padre'.

Ahora bien, ¿en qué sentido es ' $\equiv$ ' un operador de *identidad*? Uno puede pensar en 'ocurre la muerte de Sócrates  $\equiv$  Sócrates muere' como expresando identidad entre *proposiciones*; es decir, como expresando identidad entre la proposición de que ocurre la muerte de Sócrates y la proposición de que Sócrates muere. De manera similar, uno puede pensar en ' $x$  está hecho de agua  $\equiv_x$   $x$  está hecho de H<sub>2</sub>O' como expresando una identidad entre *propiedades*; es decir, como expresando identidad entre la propiedad de estar hecho de agua y la propiedad de estar hecho de H<sub>2</sub>O.

Aquí podría surgir una objeción. Uno podría objetar que cuando aseveramos 'ocurre la muerte de Sócrates  $\equiv$  Sócrates muere' (o 'que ocurra la muerte de Sócrates es simplemente que Sócrates muera') no estamos hablando acerca de *proposiciones*. Y cuando aseveramos ' $x$  está hecho de agua  $\equiv_x$   $x$  está hecho de H<sub>2</sub>O' (o 'estar hecho de agua es simplemente estar hecho de H<sub>2</sub>O') no estamos hablando acerca de *propiedades*.

La razón por la que esta objeción no me preocupa es que yo quisiera aceptar los siguientes enunciados de identidad:

**[Proposiciones]**

Que sea verdadera la proposición de que Sócrates muere *es simplemente* que Sócrates muera; y que sea falsa la proposición de que Sócrates muere *es simplemente* que Sócrates no muera.

**[Propiedades]**

Que algo tenga la propiedad de estar hecho de agua *es simplemente* que esa cosa esté hecha de agua.

La moraleja de todo esto es que entre quienes estén dispuestos a aceptar [Proposiciones] y [Propiedades], no hay ningún obstáculo para pensar en enunciados de la forma ' $p \equiv q$ ' como expresando identidades entre proposiciones, y no hay ningún obstáculo para pensar en enunciados de la forma ' $p(x) \equiv_x q(x)$ ' como expresando identidades entre propiedades.

Hay quien no aceptaría [Proposiciones] y [Propiedades], sin embargo. Para ellos lo mejor es *no* pensar en ' $\equiv$ ' como expresando identidad; es mejor, por ejemplo, leer ' $p \equiv q$ ' como 'que sea el caso que  $p$  es simplemente que sea el caso que  $q$ ', y no presuponer que esto sea equivalente a 'la proposición de que  $p$  es idéntica a la proposición de que  $q$ '.

En cualquier caso, aquí continuaré utilizando la terminología 'enunciado de identidad' para hablar de enunciados de la forma ' $p \equiv q$ ' (o de la forma ' $p(x) \equiv_x q(x)$ ').

**Pregunta:** Antes decías que quien acepta [Muerte] cree que los enunciados 'ocurre la muerte de Sócrates' y 'Sócrates muere' describen el mismo hecho. ¿Estás presuponiendo que uno puede hablar de hechos sin comprometerse con una tesis metafísicamente pesada?

**Respuesta:** Absolutamente, y eso es muy importante. A mí me parece que uno debería aceptar lo siguiente:

**[Hechos]**

Que tenga lugar el hecho de que Sócrates muera es simplemente que Sócrates muera.



Cuando digo que ‘ocurre la muerte de Sócrates’ y ‘Sócrates muere’ describen el mismo hecho, esto es simplemente una manera colorida de decir que no hay diferencia entre que ocurra la muerte de Sócrates y que Sócrates muere.

\* \* \*

A nadie le preocupa que haya un individuo al que podamos referir usando términos singulares bien diferentes. A nadie le preocupa, por ejemplo, que ‘el lucero de la tarde’ y ‘el lucero de la mañana’ refieran al mismo individuo. Lo mismo sucede en el caso de las propiedades. A nadie le preocupa que el predicado ‘está hecho de agua’ y el predicado ‘está hecho de H<sub>2</sub>O’ expresen la misma propiedad.

Cuando hablamos de hechos, sin embargo, los filósofos suelen tener una actitud muy diferente. Es bien controversial pensar que uno pueda describir el mismo hecho utilizando enunciados tan diferentes como ‘hay una mesa’ y ‘hay partículas organizadas mesísticamente’, o enunciados tan diferentes como ‘el número de los dinosaurios es cero’ y ‘no hay dinosaurios’. Son, por lo tanto, bien controversiales los siguientes enunciados de identidad:

**[Mesa]**

Que haya una mesa *es simplemente* que haya partículas organizadas mesísticamente

**[Dinosaurios]**

Que el número de los dinosaurios sea cero *es simplemente* que no haya dinosaurios.

El objetivo principal de la sesión de hoy es argumentar que, así como podemos utilizar términos singulares distintos para referir al mismo individuo, o predicados distintos para expresar la misma propiedad, también podemos utilizar enunciados bien distintos para describir el mismo hecho.

\* \* \*

Me gustaría comenzar describiendo una posición diferente a la que quiero defender. Es una posición de acuerdo con la cual basta con que dos enunciados sean gramaticalmente diferentes para que no describan el mismo hecho. En otras palabras, es una posición de acuerdo con la cual hay obstáculos *gramaticales* para la verdad de ciertos enunciados de identidad.

Quiero describir esta tesis rival para contrastarla con la posición que a mí me gustaría defender, una posición de acuerdo con la cual no hay trabas gramatical para que sea verdad un enunciado de identidad como [Mesa], o como [Dinosaurios]. No se va a

seguir de lo que digo que [Mesa] o [Dinosaurios] sean verdaderos. Lo único que se va a seguir es que si no son verdaderos, *no es simplemente por razones gramaticales*.

Discutamos, pues, la posición rival: el metafisicalismo. El metafisicalismo consiste de dos tesis diferentes; la primera es una tesis metafísica, y la segunda es una tesis semántica.

La tesis metafísica es la siguiente:

El mundo está metafísicamente estructurado. Hay una manera metafísicamente privilegiada de dividir a los hechos del mundo en sus partes constitutivas.

Consideremos, por ejemplo, el hecho de que Sócrates muere. El metafisicalista pensaría que hay una conceptualización metafísicamente fundamental de ese hecho. Yo (que no soy metafisicalista) no tengo ni idea de cuál podría ser. Tal vez es la conceptualización que divide al hecho en: (1) Sócrates y (2) la propiedad de morir. Tal vez es la que divide al hecho en: (1) la muerte de Sócrates y (2) la propiedad de ocurrir. Lo que dice la primera de las dos tesis metafisicalistas es que a lo más una de estas conceptualizaciones puede ser metafísicamente fundamental.

La segunda tesis metafisicalista es una tesis lingüística. Nos dice que, para que un enunciado atómico sea verdadero, tiene que existir una *correlación* entre la estructura semántica del enunciado y la estructura metafísica del hecho correspondiente.

Consideremos, por ejemplo, el enunciado 'Susana corre', y asumamos que su estructura semántica corresponde a su forma gramatical superficial. Consideremos ahora el hecho de que Susana corre. ¿Cuál es la estructura metafísica de este hecho? ¿Cuál es su conceptualización metafísicamente fundamental? Tal vez tenemos suerte y la conceptualización metafísicamente fundamental es la que divide al hecho en Susana y la propiedad de correr. En ese caso el enunciado 'Susana corre' puede ser verdad, porque tenemos una correspondencia entre la estructura semántica del enunciado y la estructura metafísica del hecho relevante, y porque 'Susana' refiere a una de las entidades que resulta de dividir al hecho (Susana), y 'corre' refiera a la otra (la propiedad de correr). En cambio, si no tuviéramos ese tipo de correspondencia entre estructura semántica y estructura metafísica, el enunciado no podría ser verdadero.<sup>17</sup>

**Pregunta:** ¿Qué es la estructura semántica de un enunciado?

**Respuesta:** Es simplemente la estructura que utilizamos cuando describimos una asignación composicional de condiciones de verdad al enunciado.

---

<sup>17</sup> Bob Hale y Crispin Wright discuten una versión de la posición metafisicalista en Hale, B. & Wright, C. (2009), 'The metaontology of abstraction', en Chalmers, D., Manley, D. & Wasserman, R., eds (2009), *Metametaphysics*, Oxford University Press, New York.

**Pregunta:** ¿No podría decir el metafisicalista que tanto ‘Susana corre’ como ‘la propiedad de correr está instanciada por Susana’ son verdaderos, pero que refieren a hechos distintos, con estructuras metafísicas distintas?

**Rayo:** Absolutamente. Mientras los enunciados describan hechos diferentes, el metafisicalista puede decir que los dos son verdaderos.

Más específicamente, el metafisicalista podría argumentar que hay dos hechos relevantes. El primero es el hecho de que Susana corre. Su estructura metafísica refleja la estructura semántica de ‘Susana corre’, y las expresiones ‘Susana’ y ‘corre’ refieren a las entidades en las que este hecho está dividido. El segundo es el hecho de que la propiedad de correr está instanciada por Susana. Su estructura metafísica refleja la forma semántica de ‘la propiedad de correr está instanciada por Susana’, y las expresiones ‘la propiedad de correr’, ‘está instanciada por’, y ‘Susana’ refieren a las entidades en las que este hecho está dividido.

Vale la pena señalar que si ‘Susana corre’ y ‘la propiedad de estar instanciada por Susana’ describen hechos diferentes, el siguiente enunciado de identidad no puede ser verdadero.

**[Correr]**

Que la propiedad de correr esté instanciada por Susana es simplemente que Susana corra

Este es un ejemplo del hecho de que los metafisicalistas están comprometidos a rechazar ciertos enunciados de identidad por razones meramente gramaticales.

**Pregunta:** No siempre es verdad que la estructura semántica de un enunciado corresponde a su forma gramatical. Uno podría pensar, por ejemplo, que la forma semántica de ‘Susana corre’ es algo así como ‘Hay un suceso que es un correr y en el que Susana está involucrada como agente’.

**Respuesta:** Estoy absolutamente de acuerdo. He asumido tácitamente que la estructura semántica de un enunciado es básicamente su estructura gramatical, y esto no siempre es verdad. Si la estructura semántica de ‘Susana corre’ fuera como tú propones, el metafisicalista podría aceptar:

Que exista un suceso de correr en el que Susana esté involucrada como agente es simplemente que Susana corra.

Es importante tener claro, sin embargo, que incluso en ese caso el metafisicalismo sería incompatible con ciertos enunciados de identidad. El metafisicalista tendrá que rechazar [Correr], por ejemplo, a menos de que se

comprometa con la tesis de que un enunciado como 'la propiedad de correr está instanciada por Susana' tiene la misma estructura semántica de 'hay un suceso que es un correr y en el que Susana está involucrada como agente'.

Sigue siendo el caso, por tanto, que si uno es un metafisicalista hay barreras gramaticales – barreras de gramática profunda, pero barreras gramaticales al final del día -- para la aceptación de enunciados de identidad. De acuerdo con la posición que me gustaría defender aquí, en cambio, la gramática -- e incluso la gramática profunda -- no impone ninguna restricción de este tipo.

\* \* \*

Es importante distinguir entre dos tipos de metafisicalismo: el duro y el suave. De acuerdo con el suave, la restricción metafisicalista solo es relevante cuando estamos discutiendo metafísica. En otras palabras, para hablar de manera 'metafísica seria' tenemos que asegurarnos de que la estructura semántica de nuestros enunciados atómicos corresponda a la estructura metafísica de los hechos correspondientes. Pero cuando no estamos teniendo conversaciones metafísicamente serias, podemos ignorar la distinción. De acuerdo con el metafisicalismo duro, en cambio, la restricción siempre está presente. Aunque no estemos teniendo una conversación metafísica seria, un enunciado atómico solo puede ser verdad si su estructura semántica corresponde a la estructura metafísica del hecho correspondiente.<sup>18</sup>

El metafisicalismo suave no está peleado con la posición que voy a defender aquí. No es que yo sea defensor del metafisicalismo suave; a mí, de hecho, me parece una posición altamente problemática. Aquí no voy a presuponer, sin embargo, que la posición es falsa. Aquí solo voy a estar argumentando en contra del metafisicalismo duro.

\* \* \*

Margarita Valdez sugirió durante la pausa que el *Tractatus* de Wittgenstein podría interpretarse como una posición metafisicalista.<sup>19</sup> Creo que Margarita tiene razón. De hecho, yo antes utilizaba la palabra 'tractarianismo' para hablar de metafisicalismo. Decidí, sin embargo, cambiar de terminología porque no todos los lectores del *Tractatus* están de acuerdo en que el Wittgenstein del *Tractatus* deba interpretarse como metafisicalista. Hay gente como Warren Goldfarb que cree que la manera correcta de leer el *Tractatus* es otra.<sup>20</sup> (De hecho, la lectura alterna genera una posición parecida a la que yo quiero defender aquí.)

---

<sup>18</sup> La discusión canónica de metafisicalismo suave es Sider, Theodore (2012). *Writing the Book of the World*. Oxford University Press, Oxford and New York.

<sup>19</sup> Wittgenstein, Ludwig (1921). *Tractatus Logico-Philosophicus*.

<sup>20</sup> Goldfarb, Warren (1997). 'Metaphysics and Nonsense: On Cora Diamond's *The Realistic Spirit*'. *Journal of Philosophical Research*, 22, 57–73.

Hasta aquí sobre metafisicalismo. Lo que me gustaría hacer ahora es introducir la posición que quiero defender: el composicionalismo. Al igual que el metafisicalismo, el composicionalismo consiste de dos tesis. Pero en este caso no se trata de una tesis metafísica y una lingüística; se trata de dos tesis lingüísticas. Es por esa razón que el composicionalismo es neutral con respecto a ciertas discusiones metafísicas. El composicionalismo no toma partido con respecto a la tesis de que existe la estructura metafísica.

La primera tesis composicionalista es una tesis acerca de qué es lo que hace falta para que un símbolo cuente como un término singular legítimo. La segunda es una tesis acerca de qué hace falta para que un término singular legítimo sea referencial (es decir: no vacío).

De acuerdo con la primera tesis, hay tres condiciones que juntas son suficientes para que una símbolo cuente como término singular legítimo.

- La primera condición es que el nuevo símbolo debe tener el comportamiento *sintáctico* correcto.

Supongamos, por ejemplo, que estamos introduciendo la nueva palabra 'Bumberto'. Para que 'Bumberto' cuente como término singular legítima 'Bumberto corre' tiene que contar como gramatical. También tiene que contar como gramatical 'existe algo tal que eso es idéntico a Bumberto'. No voy a tratar de darles una lista completa de pruebas gramaticales porque no la tengo. Lo único que estoy tratando de hacer es darles una idea del tipo de pruebas gramaticales que tengo en mente.

- La segunda condición es que les hayamos asignado condiciones de verdad a todo enunciado con el nuevo término que queramos tener 'en activo' (es decir: todo aquel que queramos tener disponible para la comunicación).
- La tercera condición es que nuestra asignación de condiciones de verdad respete relaciones de inferencia lógica entre enunciados.

Supongamos que  $t$  es el símbolo que estamos introduciendo, y que admitimos un enunciado atómico ' $P(t)$ '. Supongamos además que tenemos algún enunciado de la forma ' $P(t) \vee A$ '. Dado que ' $P(t) \vee A$ ' es consecuencia lógica de ' $P(t)$ ', nuestra tercera condición requiere que las condiciones de verdad que le asignemos a ' $P(t)$ ' sean por lo menos tan fuertes como las condiciones de verdad que le asignamos a ' $P(t) \vee A$ '.

Voy a ilustrar esta primera tesis composicionalista con un ejemplo que tomo prestado de Frege.<sup>21</sup> Voy a suponer que estamos introduciendo un símbolo que se escribe y se pronuncia 'dirección', pero también voy suponer que se trata de un símbolo nuevo, un símbolo al que no se le había asignado significado previamente. (No presupondré, por lo tanto, que se trata de la palabra castellana 'dirección'.) Vamos a enfocarnos en construcciones de la forma 'la dirección de  $a$ ' donde ' $a$ ' es el nombre de una línea, y vamos a preguntarnos si secuencias de símbolos de este tipo cuentan como términos singulares legítimos, de acuerdo con la primera tesis composicionalista.

La primera tesis composicionalista consiste de tres condiciones. A la primera no le quiero dedicar tiempo. Voy simplemente a presuponer que 'la dirección de  $a$ ' tiene el comportamiento sintáctico deseado.

La segunda condición de la primera tesis exige que le asignemos condiciones de verdad a enunciados que incluyan expresiones de la forma 'la dirección de  $a$ '. No hace falta, sin embargo, que le asignemos condiciones de verdad a todo enunciado gramatical con estas expresiones. Solo tenemos que preocuparnos por los enunciados que queramos tener en activo -- aquellos que queramos tener disponibles para la comunicación. No es necesario, por ejemplo, asociar condiciones de verdad con un enunciado como 'la dirección de  $a$  es de color azul'.

Consideremos una manera particular de asignarle condiciones de verdad a los enunciados del lenguaje que estamos considerando. La idea básica es decir que al enunciado 'la dirección de  $a$  es igual a la dirección de  $b$ ' se le va a asignar la condición de verdad:

las líneas  $a$  y  $b$  son paralelas

Estamos, por lo tanto, estipulando que un enunciado que contiene el término 'la dirección de  $a$ ' va a tener condiciones de verdad que pueden especificarse hablando de líneas, pero sin hablar de direcciones.

¿Cómo podríamos generalizar esta idea? ¿Qué condiciones de verdad habría que asignarle, por ejemplo, al enunciado 'existe una  $x$  tal que  $x$  = la dirección de  $a$ '? Respuesta:

existe una línea  $z$  tal que  $z$  es paralela a la línea  $a$

Hemos visto cómo asignarle condiciones de verdad a identidades como 'la dirección de  $a$  es igual a la dirección de  $b$ '; también hemos visto cómo asignarle condiciones de verdad a enunciados que cuantifican sobre direcciones como 'existe una  $x$  tal que  $x$  =

---

<sup>21</sup> Frege, Gottlob (1884). *Los principios fundamentales de la aritmética*. Sección 64. Véase también Dummett, M. (1981), *Frege: Philosophy of Language*, segunda edición, Harvard, Cambridge, MA (p. 497), y Wright, C. (1983), *Frege's Conception of Numbers as Objects*, Aberdeen University Press, Aberdeen.

la dirección de  $a'$ . Resulta que es posible generalizar este tipo de propuesta. Podemos dar una definición recursiva de condiciones de verdad para un trozo bien interesante de un lenguaje de primer orden con términos de la forma 'la dirección de  $a'$ '.<sup>22</sup>

Nuestras asignación no está bien definida para todos los enunciados del lenguaje. No hemos, por ejemplo, fijado condiciones de verdad para 'la dirección de  $a$  es color azul, o 'la dirección de  $a = \text{Julio César}$ '. Afortunadamente, el composicionalismo permite que 'la dirección de  $a'$  cuente como un término singular legítimo aunque no le asignemos condiciones de verdad a enunciados que no queramos tener en activo.

**Pregunta:** ¿Por qué no decir que los enunciados que no están en activo son falsos, en lugar de decir que carecen de condiciones de verdad?

**Respuesta:** Uno podría decir eso, pero tendríamos que decir también que hay enunciados que son falsos y tienen negaciones falsas. Tomemos, por ejemplo, 'la dirección de  $a$  corre' y 'la dirección de  $a$  no corre'. Dado que ninguno de estos enunciados está en activo, tendríamos que decir que los dos son falsos. Eso me parece poco limpio, desde un punto de vista lógico; pero no creo que este sea un punto profundo: creo que podemos proceder de la manera que más nos convenga.

Una característica de la asignación de condiciones de verdad que estamos considerando es que respeta las relaciones de inferencia lógica entre enunciados. Siempre que un enunciado en activo sea consecuencia lógica de otro, la condición de verdad del enunciado que implica es por lo menos tan fuerte como la condición de verdad del enunciado implicado.<sup>23</sup> Esto significa que también se cumple la tercera condición de la primera tesis composicionalista.

Podemos concluir, por tanto, que se satisfacen las tres condiciones del composicionalista. Secuencias de la forma 'la dirección de  $a'$  cuentan como términos singulares legítimos.

\* \* \*

---

<sup>22</sup> Aquí está la propuesta completa, donde las condiciones de verdad de  $\varphi$  están dadas por  $[\varphi]^N$ :

- $[[\text{la dirección de } a = \text{la dirección de } b]]^N = [a \text{ es paralela a } b]$
- $[[x_i = \text{la dirección de } a]]^N = [z_i \text{ es paralela a } a]$
- $[[x_i = x_j]]^N = [z_i \text{ es paralela a } z_j]$
- $[[\exists x_i (\varphi)]]^N = [\exists z_i ([\varphi]^N)]$
- $[[\varphi \wedge \psi]]^N = \text{la conjunción de } [\varphi]^N \text{ y } [\psi]^N$
- $[[\neg \varphi]]^N = \text{la negación de } [\varphi]^N$

<sup>23</sup> Un detalle técnico: aquí presupongo una lógica libre negativa, para evitar el resultado de que los enunciados de la forma ' $\exists x (x = t)$ ' cuentan como verdades lógicas.

Hemos estado discutiendo la primera de las dos composicionalistas. La segunda tesis composicionalista especifica bajo qué condiciones un término singular legítimo contará como no vacío:

Asumamos que  $t$  es un término singular legítimo, es decir, un término que satisface las tres condiciones de la primera tesis composicionalista. Para que  $t$  tenga referencia, basta con que el mundo satisfaga las condiciones de verdad que le han sido asignadas al enunciado ' $\exists x (x = t)$ ', o a algún equivalente inferencial.

La idea es la siguiente. Supongamos que hemos escogido los enunciados que queremos tener en activo, y que entre ellos está ' $\exists x (x = t)$ ' (o algún equivalente inferencial). A estos enunciados les hemos asignado condiciones de verdad. La asignación no es totalmente libre porque tiene que respetar conexiones lógicas entre enunciados, pero este es el único constreñimiento que tenemos que respetar. Lo que diría el composicionalista es que mientras se respete esa condición, y mientras sea el caso que la condición de verdad que le hayamos asignado a ' $\exists x (x = t)$ ' (o a algún equivalente inferencial) esté satisfecha, el término  $t$  tendrá referencia. Será, en otras palabras, un término singular no vacío.

Volvamos al ejemplo de las direcciones. ¿Cuáles son las condiciones de verdad del enunciado ' $\exists x (x = \text{la dirección de } a)$ '? Que exista algo que sea paralelo a la línea  $a$ . Dado que toda línea es paralela a sí misma, lo único que hace falta para que se satisfaga esta condición es que exista  $a$ . (Aquí no estoy hablando de la *dirección* de  $a$ , sino de la *línea*  $a$ .) El composicionalista dirá, por tanto, que el término singular 'la dirección de  $a$ ' es no vacío bajo, siempre y cuando vivamos en un mundo en el que la línea  $a$  existe.

No solo eso. Si 'la dirección de  $a$ ' es no vacío, podemos 'desencomillar', y así llegar a la conclusión de que existe la dirección de  $a$ . El principio desencomillador que tengo en mente aquí es el siguiente:

'la dirección de  $a$ ' es no vacío si y solo si existe la dirección de  $a$

o, en general,

$t$  es no vacío si y solo si existe  $x$

(donde ' $x$ ' se reemplaza por un término singular en nuestro metalenguaje y ' $t$ ' se reemplaza por un nombre de ese término singular.)

¿Por qué aceptar este principio? Antes de asignarle condiciones de verdad a enunciados contruidos a partir de 'la dirección de  $a$ ', podíamos *mentar* esa expresión en nuestro metalenguaje, pero no podíamos *utilizarlo* en el metalenguaje porque no era todavía significativo. Pero el composicionalista cree que una vez que



hemos completado nuestra asignación de condiciones de verdad, estamos ante un término singular legítimo y, por tanto, un término que puede ser *utilizado*, y no solo mencionado. Eso basta para que estemos en posición de formular, en nuestro metalenguaje, el principio desencomillador, y una vez que el principio ha sido formulado, no debería ser controversial. Debería ser una consecuencia inmediata de nuestro uso del concepto de referencia.

**Pregunta:** ¿Por qué llamar a la propuesta ‘composicionalismo’?

**Respuesta:** Consideremos la pregunta siguiente: ¿Por qué usar un lenguaje con términos singulares? De acuerdo con el metafisicalismo hay una respuesta muy clara a esa pregunta. Necesitamos términos singulares para que la forma lógica de nuestros enunciados atómicos corresponda a la estructura metafísica de los hechos que describen. De acuerdo con el composicionalismo, en cambio, podemos utilizar enunciados con estructuras lógicas muy diferentes para hablar de un mismo hecho. Así que no hay razones metafísicas para utilizar términos singulares. Nos interesan simplemente por sus propiedades composicionales: porque nos permiten generar, de modo composicional, una colección muy rica de enunciados, con condiciones de verdad muy diversas. Si tuviéramos un lenguaje igualmente rico sin términos singulares, no tendríamos ninguna razón para insistir en el uso de términos singulares.<sup>24</sup>

**Pregunta:** ¿Y qué de la rigidez? ¿No son importantes los nombres también por su comportamiento en contextos modales?

**Respuesta:** De acuerdo. Si alguien quisiera usar un lenguaje sin términos singulares habría que ver qué va a hacer con los contextos modales. Aquí no les he dado ninguna razón para pensar que uno pueda emular el comportamiento modal de los términos singulares en un lenguaje sin términos singulares. El composicionalista no tiene que comprometerse con que eso sea posible. Diría solo que en caso de que sea posible, y en caso de que el lenguaje alternativo sea suficientemente rico, no hay ninguna objeción obvia para pensar que podamos abandonar nuestro lenguaje a favor de un lenguaje sin términos singulares.

**Pregunta:** ¿Qué tan fregeana es la propuesta?

**Respuesta:** Bien fregeana. En *Los principios fundamentales de la aritmética*, Frege introduce el Principio Contextual, que dice que no hay que preguntar por el significado de una palabra de manera aislada, sino solamente en el contexto de una proposición.<sup>25</sup> El composicionalismo nos proporciona una manera de interpretar ese principio.

---

<sup>24</sup> Doy una respuesta más amplia a esta pregunta en mi artículo ‘A Guide to the Concept Horse’.

<sup>25</sup> Frege, Gottlob (1884). *Los principios fundamentales de la aritmética*, p. x. Ver también Wright, C. (1983), *Frege's Conception of Numbers as Objects*, Aberdeen University Press, Aberdeen.

Para explicar por qué, volvamos al metafisicalista. Si uno es un metafisicalista, uno cree que el mundo está estructurado metafísicamente. (En este contexto es útil pensar en el mundo como 'la totalidad de hechos, no de cosas'.<sup>26</sup>) Lo que nos dice el metafisicalismo, por tanto, es que hay una manera canónica de dividir en componentes a los *hechos* de los que está constituido el mundo.

El metafisicalista cree, por tanto, que independientemente de cómo es que nosotros representemos al mundo, hay una lista final de los objetos que contiene el mundo, fundamentalmente hablando; son simplemente aquellos componentes -- o, mejor, aquellos componentes objetuales -- que arroje la división canónica de los hechos. El metafisicalista también cree que para que un término refiera tenemos que especificar cuál de esos objetos va a ser el referente de nuestro término singular. En otras palabras, para que un término singular sea no vacío, hay que ir más allá de lo que propone el composicionalista. Tenemos que utilizar un concepto metafísico -- el concepto de estructura metafísica -- para determinar qué objetos contiene el mundo de manera fundamental, y luego decidir cuál de ellos va a ser el referente del nuevo término.

A mí me parece que eso es un error. Tal y como propone Frege, es un error pensar en el significado de un término singular de manera aislada. En particular, es un error pensar que uno puede asignarle referencia a un término como 'la dirección de *a*' simplemente escogiendo un referente -- un referente de entre los objetos que contiene el mundo, fundamentalmente hablando -- sin preocuparse de los enunciados en los que figura el término. ¡Es justamente al revés! Lo que tenemos que hacer es empezar con los enunciados en los que figura el término y decidir qué condiciones de verdad asignarle a esos enunciados. Una vez que tengamos condiciones de verdad, los hechos referenciales son gratis: decir que un término refiere no es más que decir que figura en enunciados verdaderos de la forma correcta.

Consideremos, por ejemplo, el hecho de que la línea *a* y la línea *b* son paralelas. Cuando introducimos los términos 'la dirección de *a*' y 'la dirección de *b*', tenemos una nueva manera de describir este hecho. Ahora lo podemos describir utilizando el enunciado 'la dirección de *a* = la dirección de *b*', y al hacerlo lo dividimos en la dirección de *a*, la dirección de *b* y la relación de identidad, y esto es suficiente para garantizar que 'la dirección de *a*' y 'la dirección de *b*' sean términos no vacíos.

Sería un error pensar que, de acuerdo con esta posición, cuando introducimos una nueva manera de describir al mundo -- los términos 'la dirección de *a*' y 'la dirección de *b*', por ejemplo -- introducimos nuevos objetos al mundo (la dirección de *a* y la dirección de *b*). Los objetos estaban ahí desde el principio.

---

<sup>26</sup> Wittgenstein, Ludwig (1921). *Tractatus Logico-Philosophicus*, 1.1.

Son simplemente parte del hecho de que la línea  $a$  y la línea  $b$  son paralelas. Lo único que ha cambiado es que, gracias a la introducción de nuevos términos, estamos en posición de identificar componentes del hecho que nuestros métodos representacionales anteriores no nos permitían identificar.

He sugerido que Frege sostuvo una posición de este tipo, pero no solo es Frege. A mí me parece que John Burgess, Charles Parsons, Bob Stalnaker y Crispin Wright han sostenido posiciones parecidas.<sup>27</sup> Parte de lo que yo quiero hacer aquí es utilizar herramientas que han desarrollado personas como ellos, y utilizarlas para enfrentar problemas filosóficos.

---

<sup>27</sup> Burgess, John (2005). 'Being explained away', *Harvard Review of Philosophy*, 13, 41–56; Parsons, Charles (1974) 'Sets and classes'. *Noûs*, 8, 1–12; reimpreso en Parsons, Charles (1983). *Mathematics in Philosophy*. Cornell University Press, Ithaca, NY; Stalnaker, Robert C. (1996). 'On what possible worlds could not be', en Morton, Adam and Stich, Stephen (ed.) (1996). *Benacerraf and his Critics* Basil Blackwell, Oxford; reimpreso en Stalnaker, Robert C. (2003). *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*. Clarendon Press, Oxford; Wright, Crispin (1983). *Frege's Conception of Numbers as Objects*. Aberdeen University Press, Aberdeen.

### **Sesión 3: El espacio de posibilidad**

*¿Cómo escoger entre concepciones rivales del espacio de posibilidad?*

---

Les voy a recordar brevemente qué es lo que ha sucedido hasta ahora. La tesis general de estas conferencias es que hay una conexión entre los enunciados de identidad que aceptemos y nuestra concepción del espacio de posibilidad. La idea es simplemente que los enunciados de identidad que aceptamos determinan qué conjunto de distinciones utilizamos para teorizar acerca del mundo, y que nuestra concepción del espacio de posibilidad es simplemente el conjunto de esas distinciones.

La vez pasada consideramos un punto más específico. Consideramos la pregunta de si hay obstáculos lingüísticos para aceptar ciertos enunciados de identidad. Para responder a esta pregunta, contrasté dos posiciones diferentes: el metafisicalismo y el composicionalismo.

De acuerdo con el metafisicalismo, los hechos están metafísicamente estructurados. Hay, por tanto, una manera ‘metafísicamente canónica’ de conceptualizar los hechos en el mundo. Por ejemplo, el hecho de que Ángeles está sentada tiene una división metafísicamente canónica. Tal vez consiste de Ángeles y la propiedad de estar sentada; tal vez consiste de ciertas partículas, de la propiedad (plural) de que las partículas estén organizadas ‘angelísticamente’, y de la propiedad (plural) de que las partículas estén organizadas ‘sentadísticamente’. La tesis central del metafisicalismo es que para que un enunciado atómico sea verdadero su forma lógica tiene que coincidir con la estructura metafísica del hecho que describe. Por ejemplo, si la forma lógica de ‘Ángeles está sentada’ es ‘P(a)’, entonces el enunciado solo puede ser verdadero si la división canónica del hecho de que Ángeles está sentada consiste de un objeto (Ángeles) y una propiedad (la propiedad de estar sentada).

Una consecuencia del metafisicalismo es que hay ciertos enunciados de identidad que no podemos aceptar, simplemente por razones gramaticales. Por ejemplo:

#### **[Muerte]**

Que ocurra la muerte de Sócrates es simplemente que Sócrates muera

Los enunciados ‘Sócrates muere’ y ‘ocurre la muerte de Sócrates’ son atómicos, y es natural pensar que tienen formas lógicas diferentes. Si esto es cierto, el metafisicalista nos diría que ‘Sócrates muere’ y ‘ocurre la muerte de Sócrates’ no pueden describir el mismo hecho, porque ese hecho tiene que tener una cierta forma canónica, y esa forma canónica puede corresponder a la forma lógica de uno u otro de nuestros enunciados, pero no a la de los dos.

Si un metafisicalista quiere aceptar tanto ‘Sócrates muere’ como ‘ocurre la muerte de Sócrates’, tiene que aceptar una posición de acuerdo con la cual hay dos hechos

diferentes. Por un lado está el hecho de que Sócrates muere, y este hecho está dividido canónicamente de una cierta manera; por otro lado está el hecho de que ocurre la muerte de Sócrates, y ese hecho está dividido canónicamente de otra manera. (Y si esto es cierto, [Muerte] no puede ser verdadero.)

El metafisicalismo es la posición que me gustaría rechazar. La posición que me gustaría aceptar es el composicionalismo. Podemos pensar en el composicionalismo como una articulación la tesis fregeana de que uno no debe preguntar por el significado de las palabras de manera aislada, sino solamente en el contexto de un enunciado.<sup>28</sup>

La tesis central del composicionalismo es la siguiente. Supongamos que introducimos un nuevo término singular  $t$  a nuestro lenguaje. ¿Qué hace falta para que  $t$  tenga referente? A diferencia del metafisicalista, el composicionalista no utiliza en concepto de la estructura metafísica de un hecho. El composicionalista cree que lo único que hace falta para que  $t$  refiera es que el mundo satisfaga las condiciones de verdad que le hayan sido asignadas al enunciado ' $\exists x (x = t)$ ', y esta es una tesis no trivial porque el composicionalista también cree que tenemos bastante libertad cuando le asignamos condiciones de verdad a nuestros enunciados. El único requisito sustancial es que nuestra asignación respete inferencias lógicas entre enunciados.

El composicionalismo tiene consecuencias muy sorprendentes. La sesión pasada consideramos una asignación de condiciones de verdad de acuerdo con la cual lo que se requiere del mundo para que el enunciado 'la dirección de  $a$  = la dirección de  $b$ ' sea verdadero es que las líneas  $a$  y  $b$  sean paralelas, y notamos que eso basta para garantizar que 'la dirección de  $a$ ' tenga referencia.

Otra consecuencia del composicionalismo es que nos permite rechazar la idea metafisicalista de que podemos descartar [Muerte] por razones simplemente gramaticales. Como vimos, el metafisicalista diría que 'Sócrates muere' y 'ocurre la muerte de Sócrates' no pueden describir el mismo hecho, porque son enunciados atómicos con formas lógicas diferentes. El composicionalismo, en cambio, elimina esta restricción, y deja abierta la puerta para que [Muerte] sea verdadero.

Les he estado diciendo que debemos rechazar el metafisicalismo, pero no les he dicho por qué. La razón es muy sencilla. Es la misma razón por la que Wittgenstein rechazó la filosofía lingüística que le atribuye a San Agustín. Cuando uno se fija en nuestro *uso* del lenguaje, y no en nuestras preconcepciones acerca de cómo debería funcionar, uno se da cuenta que el uso no está constreñido por ideas metafisicalistas. Cuando tenemos una conversación en el mercado, por ejemplo, no estamos preocupados por si la estructura lógica de lo que decimos corresponda a la estructura metafísica de los hechos. La decisión de si utilizar 'ocurrió la muerte de Sócrates' o 'Sócrates se murió' depende de qué quiera uno enfatizar, o de qué tan importante cree uno que es la muerte de Sócrates, o de qué tener referencias anafóricas queramos tener

---

<sup>28</sup> Ver, Frege, Gottlob (1884). *Los principios fundamentales de la aritmética*, Introducción.

disponibles. Pero no depende de nuestras opiniones metafísicas. Si Ángeles dice, fuera de un contexto filosófico, ‘ocurrió la muerte de Sócrates y todo mundo está triste’, sería ridículo que yo respondiera “Estoy dispuesto a conceder que Sócrates murió, pero eso de que la muerte de Sócrates haya ocurrido me parece controversial.’

\* \* \*

En la sesión de hoy quiero tratar dos temas diferentes, uno epistemológico y uno metafísico. Lo primero que me gustaría hacer es defender una tesis acerca de la epistemología de los enunciados de identidad. Me gustaría responder a la pregunta: ¿bajo qué condiciones hemos de aceptar un enunciado de identidad? (Aquí es importante tener en cuenta que aunque el composicionalismo elimina barreras gramaticales para la aceptación de un enunciado como [Mesa], no nos dice si hemos de aceptar ese enunciado o no.) Lo segundo que quiero hacer es discutir un punto metafísico. Quiero responder a las preguntas ‘¿en qué consiste la verdad de un enunciado de identidad?’ y ‘¿qué quiere decir que una concepción del espacio de posibilidad sea objetivamente correcta?’.

Nuestra pregunta epistemológica -- ¿bajo qué condiciones hemos de aceptar un enunciado de identidad? -- es más complicada de lo que uno podría pensar. El problema es que un enunciado de identidad verdadero tiene siempre condiciones de verdad triviales, y un enunciado de identidad falso tiene siempre condiciones de verdad imposibles.

Consideremos un ejemplo:

**[Calor]**

Que algo esté caliente *es simplemente* que tenga energía cinética alta.

¿Cómo tendría que ser el mundo para que [Calor] sea verdadero? Supongamos que, de hecho, es verdadero. En ese caso, lo que se requiere del mundo es que sea tal que no hay ninguna diferencia entre estar caliente (es decir, tener energía cinética alta), y tener energía cinética alta. En otras palabras, lo que se requiere es que el mundo sea tal que no hay ninguna diferencia entre tener energía cinética alta y tener energía cinética alta. Y eso es trivial. En cualquier situación coherente va a ser el caso que no hay diferencia entre tener energía cinética alta y tener energía cinética alta.

Hemos estado asumiendo que [Calor] es verdadero. Si hubiéramos asumido que es falso un argumento similar nos hubiera llevado a una condición de verdad imposible: la condición de que no haya diferencia entre estar caliente y algo que, de hecho, es diferente a estar caliente (tener energía cinética alta).

La lección de todo esto es que hay una diferencia entre la epistemología de un enunciado común y corriente como ‘la nieve es blanca’ y la epistemología de un enunciado de identidad como [Calor]. Aún cuando conozcamos sus condiciones de verdad, la decisión de si aceptar ‘la nieve es blanca’ no es trivial. Alguien puede saber

que lo que se requiere del mundo para que este enunciado sea verdadero es que la nieve sea blanca sin saber si la nieve es blanca. En cambio, la decisión de si aceptar [Calor] es absolutamente trivial una vez que uno conoce sus condiciones de verdad, porque sabemos de antemano que la condición de verdad trivial está satisfecha y que la imposible no está satisfecha. Lo difícil es saber si [Calor] tiene condiciones de verdad triviales o imposibles.

Para saber cuáles son las condiciones de verdad de un enunciado, necesitamos saber es cuáles son las condiciones de satisfacción de las palabras de las que el enunciado está compuesto. En el caso de [Calor], necesitamos saber qué es lo que hace falta para que algo satisfaga el predicado 'está caliente' y qué es lo que hace falta para que algo satisfaga el predicado 'tiene energía cinética alta'.

**Pregunta:** ¿Conocer las condiciones de satisfacción de una expresión es lo mismo que entender la expresión?

**Respuesta:** No, y esto es muy importante. Para entender una expresión lo único que hace falta es ser un usuario competente del lenguaje con respecto a esa expresión, y uno puede ser competente con el uso de una expresión sin tener información completa acerca de sus condiciones de satisfacción.

Consideremos un ejemplo. Todos somos competentes con el uso de la expresión 'está caliente', pero podríamos no estar de acuerdo en sus condiciones de satisfacción. Como vimos en la primera sesión, un teórico del calórico es alguien que cree que hay una cierta substancia -- el calórico -- tal que estar caliente consiste en tener cantidades suficientes de calórico. (Cuando algo se enfría, es porque está perdiendo calórico.) Si uno cree que existe tal substancia, uno cree que lo que hace falta para que algo satisfaga el predicado 'está caliente' es que tenga cantidades suficientes de calórico. Uno cree, por tanto, que debemos rechazar [Calor]. Hoy día, sin embargo, pensamos que este tipo de teoría no es particularmente fructífera. Aceptamos una teoría molecular del calor, y pensamos que lo que hace falta para que algo satisfaga el predicado 'está caliente' es que tenga energía cinética alta.

La pregunta de si aceptar [Calor] no es puramente lingüística. Es una pregunta que está vinculada a nuestra concepción científica del mundo, y en particular a la pregunta de si creemos que existe una substancia con las características de calórico y la pregunta de si las teorías calóricas del calor son más fructíferas que las teorías moleculares del calor.

En general, lo que hacemos cuando aceptamos un enunciado de identidad es eliminar espacio teórico. Cuando aceptamos [Calor], eliminamos el espacio teórico entre estar caliente y tener energía cinética alta. Dicho de otro modo: nos comprometemos a rechazar como absurda -- y por tanto como teóricamente inútil -- a toda situación en la que hay un objeto que está caliente pero cuyas moléculas no están en movimiento, y

a toda situación en la que un objeto no está caliente pero cuyas moléculas están en alto movimiento.

Según la propuesta que quiero defender aquí, lo que tenemos que hacer para decidir si aceptar un enunciado de identidad es determinar si es más fructífero dejar abierto el espacio teórico relevante, o si es más fructífero cerrarlo. Dejar el espacio abierto tiene ventajas e inconvenientes, y para decidir si aceptar un enunciado de identidad tenemos que sopesarlas. Si dejamos el espacio abierto, tenemos distinciones adicionales con las cuales trabajar -- en el caso de [Calor], la distinción entre estar caliente y tener energía cinética alta -- y esto podría permitirnos formular nuevas teorías. Pero las distinciones adicionales vienen con un costo, porque nos obligan a considerar ciertas preguntas como legítimas, y no siempre es claro que nuestros esfuerzos por responderlas conduzcan a teorías útiles. Si rechazáramos [Calor], por ejemplo, nos veríamos obligados a considerar como legítima la pregunta siguiente: 'Estoy de acuerdo que las partículas de las que está formado el este objeto no están en movimiento, pero ¿por qué no está caliente?'.

El ejemplo de [Calor] podría sugerir que, en general, la decisión de qué enunciados de identidad aceptar depende solo de consideraciones científicas. Pero consideraciones filosóficas también van a tener un papel que jugar. En la primera sesión discutimos

### **[Rojo]**

Que María experimente la sensación de ver rojo es simplemente que María esté en el estado cerebral R.

Cuando uno está tratando de decidir si aceptar un enunciado de identidad como este, no es claro que solo consideraciones científicas sean relevantes. Supongamos que María sabe que mañana estará en el estado cerebral R. Si aceptamos [Rojo], no reconocemos una distinción entre estar en el estado cerebral R y experimentar la sensación de ver rojo, así que no es trivial explicar cómo es que María podría tener un logro cognitivo mañana. Ya hoy sabía, después de todo, que mañana entraría en el estado cerebral R.

Alguien podría pensar que es más fácil dar cuenta del logro conceptual de María si rechazamos [Rojo]. Yo no creo que esto sea cierto,<sup>29</sup> pero ignoremos eso por el momento. Asumamos que rechazar [Rojo] tiene una ventaja clara. El problema es que rechazar [Rojo] también tiene desventajas porque nos veríamos obligados a reconocer como legítimas a preguntas potencialmente incómodas. (Por ejemplo: 'Sé que ahora mismo estoy en el estado cerebral R, pero ¿cómo es que es mi fenomenología es tal y como es, y no a la inversa?') Cuando estamos tratando de decidir si aceptar [Rojo], lo que tenemos que hacer es sopesar consideraciones de estos dos tipos. Por un lado, necesitamos suficientes distinciones para generar buenas teorías; por otro, no queremos tantas distinciones que nos quedemos con preguntas inútiles: preguntas cuyas respuestas no generen teorías que tengan utilidad independiente. Para

---

<sup>29</sup> Véase Lewis, D., 1990, "What Experience Teaches," en *Mind and Cognition: A Reader*, W. Lycan (ed.), Oxford: Blackwell.



encontrar un equilibrio apropiado entre estas presiones opuestas, no queda otra más que arremangarnos la camisa y ponernos a hacer ciencia y filosofía. Cuando se trata de decidir qué enunciados de identidad aceptar, no hay atajos.

La propuesta que estamos desarrollado aquí está fuertemente opuesta a un cierto modo de pensar, de acuerdo con el cual nuestro quehacer científico y filosófico puede dividirse limpiamente en dos partes. La primera parte consiste en decidir *a priori* con qué posibilidades trabajar. La segunda parte consiste en ir al mundo y decidir de entre estas posibilidades cuáles desechar como falsas. De acuerdo con la propuesta que estamos desarrollando aquí, en cambio, tenemos que hacer las dos cosas al mismo tiempo. Entramos al mundo sin saber si queremos contar 'hay agua' y 'hay H<sub>2</sub>O' como una distinción o dos, y por lo tanto sin saber si queremos trabajar con mundos en donde hay agua sin H<sub>2</sub>O, o H<sub>2</sub>O sin agua. Entramos al mundo sin saber si queremos contar 'experimentar la sensación de ver rojo' y 'estar en el estado cerebral R' como una distinción o dos, y por lo tanto sin saber si queremos trabajar con mundos en los que alguien está en el estado cerebral sin tener la fenomenología, o tiene la fenomenología sin estar en el estado cerebral. Es en el proceso de hacer ciencia y hacer filosofía que decidimos, al mismo tiempo, con qué posibilidades trabajar, y cuáles de estas posibilidades desechar como falsas. Es en este sentido que antes decía que uno puede pensar en la propuesta como una reconstrucción post-quineana de Carnap.

**Pregunta:** Antes habías dicho que la manera en que conocemos la verdad de un enunciado como 'la nieve es blanca' es muy distinta a la manera en la que conocemos la verdad de un enunciado de identidad. Pero ahora parece que la diferencia no es tan grande. ¿Podrías decir más?

**Respuesta:** Es cierto. Hay un sentido en el que el problema de si aceptar 'la nieve es blanca' o 'hay agujeros negros' no es tan diferente del problema de si aceptar [Calor] o [Rojo]. En ambos casos las consideraciones relevantes no siempre van a ser *a priori*, y en ambos casos la evidencia podría tener que ver con consideraciones teóricas complejas. La diferencia consiste en que enunciados como 'la nieve es blanca' o 'hay agujeros negros' tienen condiciones de verdad no triviales, así que mientras tengamos claro cuáles son esas condiciones de verdad, tendremos una pregunta clara qué hacerle al mundo. En el caso de los enunciados de identidad la situación es más compleja porque parte de lo que tenemos que decidir es qué condiciones de verdad asignarle al enunciado relevante.

**Pregunta:** No me queda claro que un enunciado de identidad restrinja nuestro espacio de posibilidad. Aún cuando, de hecho, el agua sea H<sub>2</sub>O, no parece absurdo decir 'El vaso tiene agua pero no H<sub>2</sub>O'.

**Respuesta:** Uno podría insistir que si una situación es absurda, entonces debería ser posible determinar *a priori* que se trata de una situación absurda.

Esa no es, sin embargo, la concepción del absurdo que yo tengo en mente aquí. Para ver qué es lo que tengo en mente, consideremos el siguiente enunciado de identidad:

**[Agua]**

Que algo esté compuesto de agua es simplemente esté compuesto de  $H_2O$

Supongamos que [Agua] es verdadero, y consideremos una situación en la que algo está compuesto de agua (es decir, de  $H_2O$ ) pero no de  $H_2O$ . Tenemos, en otras palabras, una situación en la que algo está compuesto de  $H_2O$  y no está compuesto de  $H_2O$ , y eso es absurdo. Es este el sentido de 'absurdo' con el que estamos trabajando aquí. Cuando pensamos en el absurdo de esta manera, el hecho de que tengamos una concepción del espacio lógico que cuente ciertas situaciones como absurdas no significa que podamos decidir *a priori* qué cuenta como absurdo y qué no.

Tampoco significa que no podamos cambiar de opinión, y cambiar nuestra concepción del absurdo. Hay por lo tanto un sentido en el que siempre permanece 'abierta' la pregunta de si sería absurdo suponer que algo que está hecho de agua está compuesto de  $H_2O$ . Una nueva teoría química, por ejemplo, podría convencernos de que conviene distinguir entre estar compuesto de agua y estar compuesto de  $H_2O$ .

Los límites que el espacio de posibilidad no son los límites de lo que no podemos excluir *a priori*. Los límites del espacio de posibilidad son los límites de lo absurdo, en el sentido de 'absurdo' con el que estamos trabajando aquí.

**Pregunta:** El conjunto de posibilidades con el que estás trabajando aquí, ¿es el mismo que el de un filósofo como David Lewis?

**Respuesta:** Lo que he estado haciendo hasta ahora es defender una manera muy general de pensar en posibilidad. Estoy defendiendo una concepción de la posibilidad de acuerdo con la cual los límites de la posibilidad dependen de qué enunciados de identidad aceptemos. Pero algo que no he hecho todavía es defender enunciados de identidad particulares, y el espacio de posibilidad con el que terminemos va a depender de qué enunciados de identidad aceptemos.

Es en principio posible trabajar con la concepción de posibilidad que estoy proponiendo aquí, y aceptar enunciados de identidad que arrojen un conjunto de posibilidades parecido al que acepta Lewis. Pero mi posición no arroja este resultado inmediatamente, y creo que esa es una gran ventaja. De acuerdo con mi propuesta, nuestra concepción de la posibilidad está determinada por los

enunciados de identidad que aceptemos y, por tanto, por las teorías que utilizamos para navegar el mundo.

Desde el punto de vista de Lewis, en cambio, las posibilidades están caracterizadas utilizando consideraciones metafísicas (la recombinación de propiedades perfectamente naturales), consideraciones que están bien separadas nuestro quehacer científico (y, en general, de las teorías que utilizamos para navegar el mundo). Lewis mismo admite esta separación, cuando señala que no estamos en posición de identificar a las propiedades puramente naturales porque su naturaleza no está vinculada a las disposiciones que somos capaces de estudiar cuando hacemos ciencia.<sup>30</sup>

De acuerdo con la concepción de posibilidad que estamos considerando aquí, en cambio, construimos las posibilidades a partir del aparato teórico que utilizamos para navegar el mundo.<sup>31</sup> Esto es importante porque, como veremos en la última sesión, no da una manera de explicar cómo es que podemos conocer los límites de la posibilidad.

\* \* \*

He argumentado que la pregunta de si aceptar un enunciado de identidad como [Calor] o como [Rojo] no es una pregunta meramente terminológica. Pero creo que hay algunos casos en los que sí nos enfrentamos a decisiones puramente terminológicas, y creo que hay casos los que la decisión depende de una mezcla consideraciones terminológicas y consideraciones no terminológicas. Voy a considerar tres ejemplos, empezando con una caso en el que la decisión es claramente terminológica y terminando con un caso en el que la decisión es claramente no terminológica.

El primer ejemplo es una discusión imaginaria entre dos zoólogos. Nuestros zoólogos están totalmente de acuerdo acerca de qué animales comparten su linaje, y acerca de quiénes son capaces de reproducirse y generar descendencia fértil. Están en desacuerdo, sin embargo, acerca de cuál es la manera más útil de agrupar animales. El zoólogo 1 cree que las teorías más útiles son las que enfatizan consideraciones que tienen que ver con el *linaje* de los diferentes animales. Este zoólogo podría pensar, en particular, que debemos aceptar:

### **[Especie 1]**

Que dos individuos sean de la misma especie es simplemente que compartan un cierto tipo de linaje

---

<sup>30</sup> Véase Lewis, David. 'Ramseyan Humility', en Braddon-Mitchel, David y Robert Nola (eds.) (2009). *Conceptual Analysis and Philosophical Naturalism*. MIT Press, Cambridge, MA.

<sup>31</sup> Ver Stalnaker, Robert C. (2008) *Our Knowledge of the Internal World*, Oxford University Press: Oxford.

El zoólogo 2, en cambio, cree es más útil clasificar animales con respecto a sus habilidades reproductivas. Tal vez diría que debemos aceptar:

**[Especie 2]**

Que dos individuos sean de la misma especie es simplemente que sean parte de un grupo capaz de reproducirse y producir descendencia fértil.

Tal y como hemos descrito el caso de los zoólogos, la decisión de si aceptar [Especie 1] o [Especie 2] es meramente terminológica. Esto es porque cualquier distinción que pueda hacer uno de los zoólogos la puede hacer también el otro. Cuando el zoólogo 1 dice '*a* y *b* son de la misma especie', el zoólogo 2 puede simplemente decir '*a* y *b* comparten un cierto tipo de linaje'; y cuando el zoólogo 2 dice '*a* y *b* son de la misma especie', el zoólogo 1 puede decir '*a* y *b* son miembros de un grupo capaz de reproducirse y generar descendencia fértil'.

No digo que la discusión entre nuestros zoólogos sea totalmente vacía. Es posible que verdaderamente resulte más útil utilizar la palabra 'especie' de una manera que de otra, y que valga la pena que los zoólogos piensen con cuidado cómo la quieren usar. Creo, sin embargo, que el interés de la discusión no va más allá de eso.

Nuestro segundo ejemplo es una discusión imaginaria entre un químico y un loco. Supongamos que el loco considera que la composición química de ciertas sustancias depende de su temperatura. Tal vez cree que aunque el alcohol etílico está normalmente compuesto de hidrógeno y oxígeno, cuando la temperatura alcanza exactamente  $\sqrt{2}^{\circ}\text{C}$ , pasa a estar compuesto de platino y cesio; y tal vez piensa que el agua normalmente está compuesta de  $\text{H}_2\text{O}$ , pero que, cuando la temperatura alcanza exactamente  $\pi^{\circ}\text{C}$ , pasa a estar compuesta de oro y mercurio. El loco debe, por tanto, rechazar:

**[Agua]**

Que algo esté compuesto de agua es simplemente esté compuesto de  $\text{H}_2\text{O}$

De lo contrario, no tendría acceso a las posibilidades que necesita para teorizar. Necesita, en particular, una posibilidad en la que hay una porción de agua cuya composición química no es  $\text{H}_2\text{O}$ , porque la temperatura es exactamente  $\pi^{\circ}\text{C}$ .

Ahora bien, si el loco rechaza [Agua], ¿qué enunciado de identidad debe aceptar? Tal vez es algo así como lo siguiente:

**[Loco]**

Que algo esté compuesto de agua es simplemente que sea un líquido inodoro, incoloro con tales-y-tales propiedades adicionales.

El químico, a diferencia del loco, está dispuesto a aceptar [Agua]. Consideremos la pregunta siguiente. ¿Es posible expresar la teoría del loco utilizando las categorías del químico?

Creo que la respuesta a esta pregunta depende de si el loco esté en posición de hacer explícito lo que quiere decir con 'tales-y-tales propiedades adicionales'. Si no está en posición de hacerlo, el químico no sabría cómo usar sus propias categorías para articular lo que el loco quiere decir cuando utiliza la palabra 'agua'. Supongamos, por ejemplo, que el loco dice 'el vaso está lleno de agua'. ¿Cómo podría el químico expresar, utilizando sus propias categorías, lo que quiso decir el loco? Simplemente repetir 'el vaso está lleno de agua' no es suficiente, porque desde la perspectiva del químico 'el vaso está lleno de agua' tiene las mismas condiciones de verdad que 'el vaso está lleno de  $H_2O$ '. El químico podría decir, 'el vaso está lleno de un líquido inodoro e incoloro', pero eso no coincide con lo que quería decir el loco, porque el loco no cree que ser inodoro e incoloro baste para ser agua. Faltan las 'tales-y-tales propiedades adicionales', y no ayudaría que el loco señalara hacia una muestra de agua y dijera 'las propiedades faltantes son precisamente las propiedades esenciales de *esto de aquí*', porque desde el punto del químico las propiedades esenciales de la muestra consisten en estar compuesto de  $H_2O$ .

Mientras que los enunciados de identidad que aceptan los zoólogos del ejemplo anterior generan concepciones isomorfas del espacio de posibilidad, no es claro que los enunciados de identidad que aceptan el químico y el loco generen concepciones isomorfas del espacio de posibilidad. Esto es importante porque, como hemos visto, la pregunta de cuál de dos concepciones isomorfas uno ha de aceptar no parece particularmente urgente. En cambio, una decisión entre concepciones potencialmente no isomorfas puede ser crucial, porque puede afectar el conjunto de distinciones que tenemos disponibles para hacer ciencia.

Nuestro último ejemplo involucra a un lógico clásico y un lógico intuicionista. El lógico clásico cree que la inferencia de  $[\sim\sim p]$  a  $p$  es lógicamente válida, y el lógico intuicionista cree que no lo es. El lógico clásico aceptaría, por tanto:

### **[Doble Negación]**

Que no sea el caso que no  $p$  es simplemente que sea el caso que  $p$ .

Una consecuencia de este enunciado de identidad es que no hay espacio teórico entre que no sea el caso que Ángeles no esté sentada y que Ángeles esté sentada. Supongamos que yo digo 'no es el caso que Ángeles no esté sentada', y alguien me responde 'De acuerdo, pero ¿está Ángeles sentada?'. Quien acepte [Doble Negación] dirá que esa pregunta ya ha quedado respondida, porque ya se aclaró que no es el caso que Ángeles no esté sentada.

El intuicionista, en cambio, está comprometido a rechazar [Doble Negación] porque quiere mantener abierto un espacio teórico entre  $p$  y que no sea el caso que no  $p$ . ¿Por qué mantener abierto este espacio teórico? Hay quien piensa que podría ayudarnos a resolver ciertas paradojas. Es importante tener en mente, sin embargo, que aún si el rechazo de [Doble Negación] resultara tener ventajas, también tendría costos, pues no

es inmediatamente obvio cómo entender el espacio teórico entre  $p$  y que no sea el caso que no  $p$ .

Yo soy un lógico clásico, y genuina y auténticamente no entiendo cómo tendría que ser el mundo para que esté claro que no es el caso que Ángeles no está sentada sin que esté claro que Ángeles está sentada. No digo que no se me pueda hacer cambiar de opinión. Se me podría convencer si se me convenciera de la utilidad de abrir espacio teórico entre  $p$  y  $[\sim\sim p]$ . Tal vez se me podría convencer de que el espacio adicional nos ayudaría a resolver paradojas, o tal vez Dummett podría convencerme de que se requiere para hacer metafísica.<sup>32</sup> En ese caso adoptaría un nuevo sistema de distinciones, algunas de las cuales me parecen absurdas desde mi perspectiva actual.

Estamos, por tanto, en un caso bien distinto al de los dos zoólogos. Estamos en un caso en el que el lógico clásico, utilizando sus propias categorías, ve las distinciones del intuicionista como absurdas. (El intuicionista tendría una queja análoga. Así como lógico clásico cree que el intuicionista está haciendo distinciones sin diferencia, el intuicionista cree que el lógico clásico está confundiendo dos distinciones que son diferentes.)

**Pregunta:** ¿Qué relación hay entre la lógica que aceptamos y nuestra concepción del espacio de posibilidad?

**Respuesta:** Creo que nuestra concepción de la consecuencia lógica es parte de lo que define nuestra concepción del espacio de posibilidad. De hecho hay una manera fácil de convertir nuestras opiniones acerca de qué lógica es la correcta en opiniones acerca de qué enunciados de identidad aceptar. Supongamos que queremos contar a  $\phi$  como verdad lógica. Entonces deberíamos aceptar el siguiente enunciado de identidad:

que sea el caso que  $\phi$  es simplemente que sea el caso que  $T$

en donde  $T$  es una tautología. Este método tiene la ventaja de ser muy general, pero dependiendo de la lógica que adoptemos podría haber maneras más interesantes de generar enunciados de identidad. Acabamos de ver, por ejemplo, que el lógico clásico aceptaría y el lógico intuicionista negaría [Doble Negación].

Una razón por la que es interesante el caso de la lógica es que ilustra los límites de la racionalidad. Algunos de nuestros movimientos cognitivos pueden calificarse como racionales o no racionales. Un bayesiano, por ejemplo, diría que un movimiento cognitivo es racional en tanto que resulta de

---

<sup>32</sup> Dummett, M. *The Logical Basis of Metaphysics* London: Duckworth, and Cambridge MA: Harvard University Press, 1991.

condicionalizar sobre nuestra evidencia.<sup>33</sup> No es claro, sin embargo, que podamos hablar de racionalidad o irracionalidad cuando se trata de cambiar de lógica.

No digo que no sea posible cambiar de lógica. Si a mí se me muestra que la lógica intuicionista, o la lógica paraconsistente, puede hacer trabajo teórico suficientemente interesante, estaría dispuesto a cambiar de lógica. Pero uno nunca podría usar un argumento *lógico* para justificar este tipo de movimiento cognitivo. Desde la perspectiva de una lógica dada uno nunca puede llegar a la conclusión de que una lógica distinta es la correcta. Utilizando una herramientas clásicas, por ejemplo, siempre llegaremos a la conclusión de que la lógica intuicionista está haciendo distinciones sin diferencia.<sup>34</sup>

Cambiar de lógica y, de manera más general, cambiar de concepción del espacio lógico, es un proceso a-racional. Es un poco como cambiar de caja de herramientas. Uno puede experimentar con un nuevo conjunto de herramientas, y decidir si le gusta lo suficiente como para trabajar con él. La decisión de si quedarse con las nuevas herramientas es, en parte, un asunto *práctico*: es un asunto que depende, en parte, de qué manera de trabajar nos parezca más cómoda. De acuerdo con el tipo de posición que he estado defendiendo aquí, lo mismo sucede con la decisión de si aceptar una concepción del espacio de posibilidad: la decisión es, en parte, un asunto *práctico*. La racionalidad regula el movimiento cognitivo dentro de una concepción del espacio de posibilidad, pero no puede regular los cambios entre concepciones diferentes del espacio de posibilidad.

Desde nuestro punto de vista, el único espacio de posibilidad que no rebasa los límites del absurdo es el nuestro. Todos los espacios rivales utilizan distinciones sin diferencia, o confunden distinciones diferentes. Al mismo tiempo, somos capaces de imaginarnos cambiando de concepción del espacio posibilidad: somos capaces de imaginarnos cambiando de concepción del absurdo. Eso ocurriría, en particular, si cambiamos de lógica.

Desde mi punto de vista actual -- el punto de vista de un lógico clásico -- el lógico intuicionista está haciendo distinciones sin diferencia, y está, por tanto, comprometido con una posición absurda. Al mismo tiempo, podría haber consideraciones teóricas que me llevaran a aceptar una lógica intuicionista, y por lo tanto a cambiar mi concepción del absurdo.

Tal vez uno podría pensar que existe tal cosa como 'la perspectiva de Dios': un espacio de súper-posibilidad, en donde cada punto es una concepción diferente

---

<sup>33</sup> Hay una buena introducción al bayesianismo en Talbott, William, "Bayesian Epistemology", *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2011/entries/epistemology-bayesian/>.

<sup>34</sup> Hartry Field desarrolla una propuesta de este tipo en Field, Hartry (2009) 'Epistemology without Metaphysics' *Philosophical Studies* 143.

del espacio de posibilidad. Tal vez Dios sería capaz de ver la 'coherencia' -- en algún sentido de 'coherencia' -- de cada una de esas concepciones del espacio lógico, y tal vez Dios podría emitir un juicio acerca de cuál es la 'objetivamente correcta.' A mí esto me parece un error. Yo lo que creo es que no posible -- ni para Dios ni para nosotros -- describir al mundo sin tomar la perspectiva de alguna concepción particular del espacio de posibilidad. Estamos en el Barco de Neurath.<sup>35</sup>

**Pregunta:** El hecho de que me parezca que alguien más está comprometido con un absurdo no significa que no nos podamos comunicar. ¿Es así?

**Respuesta:** Correcto. Lo que hace falta para comunicarnos es mucho menos que lo que hace falta para estar de acuerdo en una concepción del espacio de posibilidad. Tanto el químico y como el loco utilizan la palabra 'agua', y esto les permite cierto nivel de comunicación. Si, por ejemplo, uno le dice al otro '¡Pásame aquella botella de agua!', el hecho de que le asignen diferentes condiciones de satisfacción a la palabra 'agua' no va a ser un obstáculo para la comunicación. Hay ciertos proyectos teóricos, sin embargo, para los cuales las diferencias entre el químico y el loco sí importan. Por ejemplo, solo el loco vería valor en el proyecto de determinar cómo cambia la composición química del agua conforme varía la temperatura.

\* \* \*

Les he ofrecido una propuesta acerca de la *epistemología* de los enunciados de identidad. He argumentado que la decisión de qué enunciados de identidad aceptar debe estar basada en qué tan útil sería una teoría construida sobre el sistema de distinciones que resultaría de aceptar los enunciados de identidad relevantes. Más específicamente, he argumentado que aceptar un enunciado de identidad tiene ventajas y desventajas, y que lo que tenemos que hacer es sopesar consideraciones de esos dos tipos. Hay algunos casos en los que la discusión acerca de qué enunciados de identidad aceptar es meramente verbal, pero hemos visto que hay otros casos en donde la discusión es sustancial porque las concepciones del espacio de posibilidad que resultarían de aceptar diferentes enunciados de identidad no son isomorfas, y por lo tanto no arrojan los mismos recursos teóricos.

Ahora quisiera dejar de lado la discusión epistemológica, y considerar la pregunta de si tiene sentido decir que un enunciado de identidad es 'objetivamente verdadero'. Voy a comenzar con algunas consideraciones preliminares. Lo primero que me gustaría hacer es recalcar algo que ya habíamos discutido: la conexión entre los enunciados de identidad, el absurdo y la posibilidad metafísica. Todos los ingredientes ya están sobre la mesa, lo único que tenemos que hacer es juntarlos.

---

<sup>35</sup> La analogía aparece en Quine, W.V.O *Word and Object*, MIT Press, Cambridge MA (1960), p. 3.



Hablemos primero de la noción de posibilidad. Diferentes filósofos han dicho tantas cosas muy diferentes acerca de esta noción. En particular, hay diferentes opiniones acerca de qué está haciendo el término ‘metafísica’ cuando alguien dice “posibilidad metafísica”.<sup>36</sup>

Una manera de entender ‘metafísica’ es como determinando un cierto nivel de exigencia.<sup>37</sup> De acuerdo con esta propuesta, tenemos un espacio de posibilidad máxima, algo así como un espacio de posibilidades lógicas. Dentro de este gran espacio maximal, podemos distinguir entre nociones más o menos estrechas de posibilidad, representados como círculos concéntricos, centrados en la posibilidad ‘actualizada’. El círculo más pequeño consiste de las posibilidad físicas, es decir, posibilidades que comparten con la posibilidad actualizada todas las leyes físicas. El siguiente círculo consiste de las posibilidades metafísicas, es decir, posibilidades que no necesariamente comparten con nosotros las leyes físicas, pero sí comparten las ‘leyes metafísicas’. El siguiente círculo consiste de posibilidades conceptuales, es decir, posibilidades que no necesariamente comparten con nosotros las leyes físicas o metafísicas, pero que sí comparten con nosotros la ausencia de contradicciones conceptuales. El círculo más amplio es el espacio completo.

De acuerdo con esta manera de pensar, el papel que juega la palabra ‘metafísica’ en ‘posibilidad metafísica’ juega un papel análogo ‘física’ en ‘posibilidad física’. Lo que hace es especificar uno de los círculos concéntricos en el espacio maximal; especifica, en otras palabras, un cierto nivel de exigencia.

Esa es una manera de pensar en posibilidad metafísica. Pero a mí no me parece particularmente útil, y aquí me gustaría trabajar con una propuesta una alternativa. De acuerdo con la propuesta alternativa, el papel de la palabra ‘metafísica’ en ‘posibilidad metafísica’ es distinguir entre dos tipos de posibilidad: posibilidad *de repraesentatione* y posibilidad *de mundo*. Para que dos enunciados expresen la misma posibilidad *de mundo* lo único que hace falta es los enunciados tengan la misma condición de verdad; lo único que hace falta, en otras palabras, es que el requerimiento que su verdad le impone al mundo sea el mismo. Noten, sin embargo, que puede haber representaciones bien diferentes que compartan condiciones de verdad. (Por ejemplo, ‘Héspero es Fósforo’ y ‘Héspero es Héspero’.)

El que dos enunciados expresen la misma posibilidad *de repraesentatione*, en cambio, depende no solo de las condiciones de verdad de los enunciados sino también de las particularidades de los enunciados mismos. La consistencia lógica es un ejemplo de posibilidad *de repraesentatione*, pues el que un enunciado cuente como lógicamente consistente depende no solo de las condiciones de verdad del enunciado sino también

---

<sup>36</sup> Hay una buena discusión de estos temas en Rosen, Gideon (2006). ‘The limits of contingency’, en MacBride, F. (ed.) (2006). *Identity and Modality*. Oxford University Press, Oxford.

<sup>37</sup> Véase, por ejemplo, Kment, Boris (2006) ‘Counterfactuals and the analysis of necessity’ *Philosophical Perspectives*, **20**, 237–302; Cameron, Ross (2010) ‘The grounds of necessity’ *Philosophical Compass*, **4**, 348–358; y Sider, Theodore (2012). *Writing the Book of the World*. Oxford University Press, Oxford and New York.

de las particularidades del enunciado. 'Héspero no es Fósforo', por ejemplo, es lógicamente consistente, pero 'Héspero no es Héspero' no lo es, a pesar de que los dos enunciados tienen la misma condición de verdad. Lo que se requiere del mundo para que uno sea verdadero es precisamente lo que hace falta para que el otro sea verdadero.

De acuerdo con la noción de posibilidad metafísica con la que me gustaría trabajar aquí, el papel de 'metafísica' en 'posibilidad metafísica' es aclarar que estamos hablando de posibilidad *de mundo*, y no de posibilidad *de repraesentatione*. La propuesta es, por tanto, que la posibilidad metafísica es la posibilidad *de mundo* más incluyente que hay.<sup>38</sup>

Desde este punto de vista, hay una relación bien estrecha entre qué enunciados de identidad aceptamos, qué espacio de posibilidad aceptamos y qué contamos como absurdo. En particular:

- Un enunciado de primer orden describe una posibilidad metafísica si y sólo si es lógicamente consistente con el conjunto de enunciados de identidad verdaderos.
- Un enunciado de primer orden describe un absurdo (en el sentido de 'absurdo' que antes habíamos discutido) si y sólo si es lógicamente inconsistente con el conjunto de enunciados de identidad verdaderos.

Por tanto, decir que un enunciado expresa una necesidad metafísica es decir que su negación es inconsistente con los enunciados de identidad verdaderos; decir que un enunciado es inconsistente con los enunciados de identidad verdaderos es decir que expresa un absurdo; y decir que expresa un absurdo es decir que no expresa una posibilidad metafísica.

El resultado es un conjunto de nociones -- posibilidad, identidad y absurdo -- que parecían diferentes pero que terminan estando estrechamente relacionadas. Más específicamente, las nociones de posibilidad, identidad y absurdo se complementan unas a otras. No es que comencemos entendiendo una de las nociones claramente, y la utilicemos para aclarar las demás. Lo que sucede es que cada una de las tres nociones viene con ciertos constreñimientos teóricos. Cuando establecemos conexiones entre ellas, el resultado es que combinamos los constreñimientos, y esto arroja luz sobre las tres nociones al mismo tiempo.

Otra razón por la que es útil establecer conexiones entre las tres nociones es nos da una manera muy cómoda de hablar. Nos permite, en particular, pensar en el espacio de posibilidad como el conjunto de posibilidades metafísicas.

---

<sup>38</sup> Hasta donde puedo ver, esta es la propuesta que tenía en mente Kripke en *Naming and Necessity*, Harvard University Press, Cambridge, MA (1980).

\* \* \*

Estamos listos para considerar el problema de qué podría querer decir que un enunciado de identidad sea objetivamente verdadero.

Aquí hay una manera muy natural de pensar en la noción de verdad. Comenzamos con una concepción del espacio de posibilidad. Cuando alguien asevera un enunciado, la aseveración tiene dos efectos. Primero, divide al espacio de posibilidad en mundos posibles que verificarían la aseveración, y mundos posibles que no la verificarían; segundo, la aseveración excluye a los mundos posibles en la segunda categoría. Podemos decir entonces que el enunciado es verdadero si y sólo si el mundo actual no ha quedado excluido.

Cuando pensamos en la noción de verdad de esta manera, resulta que un enunciado de identidad va a contar como verdadero con respecto a una concepción del espacio de posibilidad si y solo si la concepción fue construida a partir de ese enunciado de identidad. Por ejemplo, si yo usé [Agua] para construir mi concepción del espacio de posibilidad, todo mundo posible en mi espacio va a ser uno en el que no hay diferencia entre estar compuesto de agua y estar compuesto de  $H_2O$ . Por tanto, todo mundo en mi espacio va a ser uno en el que [Agua] es verdadero. Y si no hubiera usado [Agua] para construir mi concepción del espacio de posibilidad, todo mundo posible en mi espacio sería tal que estar compuesto de agua no es lo mismo que estar compuesto de  $H_2O$ . Por tanto, todo mundo en mi espacio sería uno en el que [Agua] es falso.

Supongamos ahora que queremos saber si un cierto enunciado de identidad es objetivamente verdadero. Sabemos que nuestro enunciado es verdadero con respecto a algunas concepciones del espacio de posibilidad pero no otras. ¿Con respecto a cuál de ellas debemos evaluar verdad objetiva? La respuesta natural es: con respecto a la concepción *objetivamente correcta* del espacio de posibilidad.

¿En qué consiste, pues, la corrección objetiva de una concepción del espacio de posibilidad? La posición que he estado defendiendo aquí no me compromete a ninguna respuesta particular a esa pregunta, pero les confieso que solamente hay una respuesta que soy capaz de entender. Voy a tratar de articular esa respuesta, y los voy a invitar a que la acepten. Lo que no voy a hacer, sin embargo, es presentar un argumento de que esa es la única propuesta posible. De hecho, yo preferiría una alternativa. Mi sospecha, sin embargo, es que no hay una alternativa viable, y que es difícil escapar del tipo propuesta que estoy por articular.

De acuerdo con la propuesta que quiero defender, adoptar una concepción del espacio de posibilidad es siempre una decisión práctica, pero es una decisión práctica que depende en parte de cómo es el mundo. Si el mundo fuera diferente, podría ser que lo que ahora es una buena decisión práctica dejaría de ser una buena decisión práctica. Aquí nos puede ayudar una analogía con los juegos.

Cuáles son 'las reglas del juego' depende solo de qué juego estemos jugando: una vez que se fija el juego, quedan fijas las reglas. Cuál es 'el marcador del juego', en cambio, depende de dos cosas; depende de qué juego estemos jugando y también depende de ciertos hechos empíricos. (Depende, por ejemplo, de si la reina de corazones está sobre la mesa, o de si el balón ha entrado en la portería.) La analogía que les propongo es que aseverar un enunciado de identidad es como hacer una aseveración acerca de las reglas del juego, y aseverar un enunciado contingente ('la nieve es blanca', por ejemplo) es como hacer una aseveración acerca del marcador del juego. La idea es que, así como no tiene sentido hablar de 'el juego con reglas objetivamente correctas', tampoco tiene sentido hablar de 'la concepción del espacio de posibilidad basada en enunciados de identidad objetivamente verdaderos'.

La decisión de qué reglas adoptar para jugar un juego es una decisión práctica, pero es una decisión práctica que depende en parte de cómo es el mundo. Cuando decidimos qué juego jugar, nuestra decisión depende de si tenemos ganas de ejercicio, de si todo el mundo se acuerda de las reglas del *Bridge*, de qué tan borracho está el tío Heriberto, y así sucesivamente. De la misma manera, la decisión de qué concepción del espacio de posibilidad aceptar es una decisión práctica, pero es una decisión práctica que depende en parte de cómo es el mundo. Supongamos, por ejemplo, que estamos tratando de decidir si aceptar [Rojo]; es decir, estamos tratando de decidir si aceptar la idea de que experimentar la sensación de ver rojo es simplemente estar en el estado cerebral R. Esta es una decisión práctica, pero es una decisión práctica que depende en parte de cómo es el mundo; depende en parte de cómo funcionan nuestros cerebros, y del tipo de experiencias que tenemos.

Es importante notar que esta manera de pensar no nos compromete con la conclusión de que no hay tal cosa como la verdad. A lo que estamos comprometidos es a la tesis de que la verdad es siempre relativa a una concepción del espacio lógico, y a la tesis de que la decisión de qué concepción aceptar es irremediabilmente práctica. Hay algo que depende de decisiones prácticas y algo que no depende de decisiones prácticas. Lo que depende de decisiones prácticas es qué distinciones utilizar para teorizar acerca del mundo; en otras palabras, qué concepción del espacio de posibilidad adoptar. Lo que no depende de decisiones prácticas es cuáles de las posibilidades en el espacio de posibilidad que hemos adoptado están actualizadas y cuáles no. Eso lo determina el mundo.

Aquí hay otra manera de pensar en el asunto. De acuerdo con la posición que estoy defendiendo, hay un hecho objetivo acerca de cómo es el mundo; pero cuando tratamos de *describir* al mundo, no nos queda más que introducir distinciones, y no hay un hecho objetivo acerca de qué distinciones hemos de utilizar. Esa es una decisión que depende, en parte, de consideraciones prácticas. Es importante, sin embargo, no confundir la idea de que no hay un hecho objetivo acerca de qué distinciones utilizar para describir al mundo con la idea de que no haya un hecho objetivo acerca de cómo es el mundo. La falta de objetividad viene del proyecto de describir el mundo, no de la manera en la que es el mundo.

**Pregunta:** ¿Se sigue de esas posición que una pregunta como ‘¿Cuál es la temperatura actual en la Ciudad de México?’ no tiene respuesta objetivamente correcta?

**Respuesta:** No. De acuerdo con mi posición hay un hecho objetivo acerca de cómo es el mundo y, en particular, acerca de cuál es la temperatura actual en la Ciudad de México. Las cosas se complican, sin embargo, cuando tratamos de *describir* al mundo. El problema surge porque que la única manera que tenemos de describir al mundo es introduciendo distinciones, y para mí no es claro que tenga sentido hablar de un sistema de distinciones objetivamente correcto. Al mismo tiempo, creo que una vez que adoptamos un sistema de distinciones, hay una respuesta objetiva a la pregunta ‘¿de qué lado de esas distinciones está el mundo?’.

Una vez que hemos adoptado la distinción entre maneras de ser del mundo de acuerdo con las cuales la temperatura actual en la Ciudad de México es 20°C, y maneras de ser del mundo de acuerdo con las cuales no es el caso que la temperatura sea 20°C, hay una respuesta objetiva a la pregunta ‘¿de qué lado de esta distinción está el mundo’. Pero es solo desde la perspectiva de nuestro sistema de distinciones que tiene sentido hacer esa pregunta.

**Pregunta:** Se me ocurre que tal vez podríamos utilizar diferentes concepciones del espacio de posibilidad para modelar las maneras de pensar de diferentes comunidades, y los cambios de metodología científica a través del tiempo. ¿Sería esa una ventaja de la posición que estás defendiendo?

**Respuesta:** Estoy de acuerdo con todo lo que dices, excepto el optimismo. Es cierto que uno podría pensar que diferentes comunidades adoptan concepciones diferentes del espacio de posibilidad.<sup>39</sup> Es importante notar, sin embargo, que -- desde nuestra perspectiva y con nuestros recursos teóricos -- no es fácil hacerle justicia a las concepciones rivales del espacio de posibilidad. Hemos visto, por ejemplo, que no es fácil que un lógico clásico le haga justicia a la perspectiva teórica de un lógico intuicionista, y que no es fácil que el químico le haga justicia a la perspectiva teórica del loco. Esto significa que en tanto que creamos que necesitamos una concepción rival del espacio de posibilidad para modelar las maneras de pensar de una comunidad, no va a ser fácil hacerle justicia.

**Pregunta:** ¿No podríamos aceptar un enunciado de identidad en algunos contextos pero rechazarlo en otros?

---

<sup>39</sup> Una discusión clásica de estos temas es Kuhn, Thomas S. (1996). *The Structure of Scientific Revolutions* (tercera edición) University of Chicago Press.

**Respuesta:** Absolutamente. Es una idealización pensar que una persona, o una comunidad, trabaja siempre con un único espacio de posibilidad. Lo que sucede en la práctica es que con respecto a proyectos diferentes aceptamos enunciados de identidad diferentes, según nos sea útil. No es que mi teoría me comprometa con esto, pero es mi mejor hipótesis acerca de cómo es que de hecho operamos los humanos.

Una ventaja de permitir variaciones en los enunciados de identidad que una acepta es que nos permite apropiarnos de un truco de Lewis que a mí me parece muy útil.<sup>40</sup> Supongamos que construyo una estatua de bronce de una persona con un brazo levantado. ¿Son objetos diferentes la estatua y la porción de bronce del que la estatua está compuesta? ¿O se trata de un solo objeto? De acuerdo con Lewis hay un solo objeto, pero según utilicemos la expresión ‘estatua’ o ‘porción de bronce’ hacemos relevantes diferentes relaciones de similitud entre mundos posibles. Eso le permite decir que ‘la estatua hubiera sobrevivido la destrucción del brazo levantado’ es verdadero, pero que ‘la porción de bronce hubiera sobrevivido la destrucción del brazo levantado’ no es verdadero.

Nosotros podemos utilizar un truco similar. Supongamos que la estatua es el mismo objeto que la porción de bronce que la constituye, y llamemos a ese objeto ‘X’. Supongamos también que el brazo levantado es el mismo objeto que la porción de bronce que lo constituye, y llamemos a ese objeto B. Con respecto a proyectos diferentes aceptamos distintos enunciados de identidad. A veces aceptamos el siguiente enunciado de identidad y a veces lo rechazamos:

**[Bronce]**

Parte de lo que es ser el objeto X es tener a B como parte.

Cuando decimos ‘la porción de bronce no hubiera sobrevivido la destrucción del brazo levantado’, estamos aceptando [Bronce], y cuando decimos ‘la estatua hubiera sobrevivido la destrucción del brazo levantado’ estamos rechazando [Bronce].

---

<sup>40</sup> Ver Lewis, David (1986a). *On the Plurality of Worlds*. Blackwell, Oxford y Nueva York, sección 4.5.

## Sesión 4: Las matemáticas y el espacio de posibilidad

*¿Es correcto pensar en las proposiciones matemáticas como descripciones del mundo?*

---

En la sesión de hoy me gustaría argumentar que podemos utilizar el conjunto de herramientas que hemos estado desarrollando en sesiones anteriores para desarrollar una buena filosofía de las matemáticas.

De acuerdo con la propuesta que quiero defender el siguiente enunciado de identidad es verdadero:

### **[Dinosaurios]**

Que el número de los dinosaurios sea cero es simplemente que no haya dinosaurios.

La mejor manera de entender en qué consiste esta tesis es utilizando la metáfora de Dios. Alguien que rechaza [Dinosaurios] podría pensar que cuando Dios hizo al mundo y se aseguró de que no hubiera dinosaurios, tenía que hacer algo *adicional* para que el número de los dinosaurios fuera cero. Faltaba que Dios abriera el cajón de los números, seleccionara al cero, lo agregara al mundo, y se asegurara de que este nuevo objeto numerara a los dinosaurios. Solo entonces se habría asegurado de que el número de los dinosaurios sea cero.

Alguien que acepta [Dinosaurios], en cambio, cree que cuando Dios hizo al mundo y se aseguró de que no hubiera dinosaurios, ya había terminado. Ya había hecho todo lo que tenía que hacer para que el número de los dinosaurios fuera cero. La razón es que el hecho de que no haya dinosaurios es el mismo que el hecho de que número de los dinosaurios sea cero. En otras palabras, 'el número de los dinosaurios es cero' y 'no hay dinosaurios' son representaciones diferentes d el mismo hecho.

Uno podría generalizar [Dinosaurios] como sigue:

### **[Números]**

Que el número de las Fs sea  $n$  es simplemente que haya  $n$  Fs.<sup>41</sup>

Lo que voy a hacer en la sesión de hoy es defender una filosofía de las matemáticas basada en [Números]. El primer paso de mi argumento es uno que ya discutimos en sesiones anteriores. Es la tesis de que no hay razones *lingüísticas* para resistir [Números]. Defendí esa idea en la sesiones 2 y 3, cuando discutimos la diferencia entre el metafisicalismo y el composicionalismo. El metafisicalismo tiene la consecuencia de que la estructura semántica de [Números] es suficiente para

---

<sup>41</sup> Al igual que en la primera sesión, estoy pensando en [hay  $n$  Fs] como una abreviación de  $[\exists!_n x (Fx)]$ .

garantizar su falsedad. Argumenté, sin embargo, que debemos rechazar el metafisicalismo a favor del composicionalismo. El composicionalismo no implica que [Números] sea verdad, pero sí implica que la estructura semántica de [Números] no basta para garantizar su falsedad.

El segundo paso de mi argumento es la tesis de que debemos aceptar [Números]. De acuerdo con la epistemología de los enunciados de identidad que discutimos en la sesión anterior, cuando uno acepta un enunciado de identidad, el resultado es que uno elimina cierto espacio teórico. Los casos en los que hemos de aceptar el enunciado de identidad son casos en donde las ventajas de eliminar ese espacio teórico son suficientes para compensar las desventajas. Voy a argumentar, por tanto, que las ventajas de aceptar [Números] son suficientes para compensar las desventajas.

El paso final de mi argumento va a consistir en mostrar que uno puede desarrollar una filosofía de las matemáticas seria basada en [Números]. [Números] es una tesis sugerente, pero no constituye, por sí sola, una filosofía de las matemáticas. Necesitamos, por ejemplo, una *semántica* para los enunciados de la matemática, una semántica que le haga justicia a [Números]. También necesitamos mostrar que [Números] es consistente con una metafísica razonable, y necesitamos desarrollar una epistemología para las matemáticas basada en [Números].

El proyecto de dar una epistemología de las matemáticas lo vamos a dejar para la próxima sesión, pero el resto del argumento lo vamos a discutir en la sesión de hoy.

\* \* \*

Consideremos el segundo paso del argumento. ¿Cuáles serían las ventajas y desventajas de aceptar [Números]? Una ventaja fundamental de aceptar [Números] es que quedaríamos exentos de responder a la siguiente pregunta incómoda:<sup>42</sup>

Sé que no hay dinosaurios. Lo que me gustaría entender ahora es si *también* es cierto que el número de los dinosaurios es cero. Me gustaría entender, además, cómo es que uno podría estar justificado en tener una opinión al respecto, dado que no gozamos de ningún tipo de acceso causal al mundo de los objetos abstractos.

Alguien que acepta [Números] diría que cuando llegamos a la conclusión de que no hay números, ya habíamos llegado a la conclusión de que el número de los dinosaurios es cero, porque *no hay diferencia* entre lo uno y lo otro: que el número de los dinosaurios sea cero es simplemente que no haya dinosaurios. Quien hace la pregunta incómoda está, por tanto, introduciendo una presuposición falsa a la discusión. Está presuponiendo que hay espacio teórico entre que no haya dinosaurios y que el

---

<sup>42</sup> Paul Benacerraf famosamente planteó una pregunta de este tipo en 'Mathematical truth' *Journal of Philosophy*, 70, 661–79, 1973.



número de los dinosaurios sea cero, y es un error pensar que exista tal espacio teórico. Esto significa que la pregunta incómoda no es una realmente pregunta legítima.

Creo que un defecto crucial en la literatura sobre filosofía de las matemáticas es que se suele presuponer -- y presuponer sin argumento -- que por el simple hecho de que dos enunciados tengan formas gramaticales diferentes están describiendo hechos diferentes.<sup>43</sup> Se presupone, en particular, que 'no hay dinosaurios' y 'el número de los dinosaurios es cero' describen hechos distintos. Esto ha generado un sinfín de teorías dedicadas llenar el espacio teórico entre los dos hechos, y, cada vez más, uno tiene la impresión de que esas teorías no van a ningún lado. La gran ventaja de atacar el problema utilizando enunciados de identidad, como estamos haciendo aquí, es que podemos tomar la decisión de cerrar ese espacio teórico, y así enfrentar los retos de la filosofía de la matemática utilizando otro conjunto de herramientas.

Hemos considerado una ventaja importante de aceptar [Números]. ¿Qué desventajas tendría? Si elimináramos el espacio teórico entre que no haya Fs y el que el número de las Fs sea cero, tendríamos que contar como absurda una situación en la que no hay dinosaurios pero es falso que el número de los dinosaurios sea cero. No podríamos, en otras palabras, incluir esa situación entre las posibilidades que tenemos disponibles en nuestro trabajo teórico. ¿Pero qué tanta utilidad teórica tendría una situación así? Hasta donde puedo ver, sería totalmente inútil. La única consecuencia de incluir situaciones de este tipo en nuestro espacio de posibilidad es que tendríamos que preguntarnos por qué no están actualizadas.

Me parece claro, por tanto, que las desventajas de aceptar [Números] son muchísimo menos importantes que las ventajas de aceptarlo. Si esto es así, tenemos buenas razones para aceptar [Números].

(Vale la pena notar, por cierto, que esta es la primera vez que realmente consideramos un argumento para aceptar un enunciado de identidad, y no solo un argumento para pensar que no hay obstáculos gramaticales para aceptar un enunciado de identidad.)

\* \* \*

Me gustaría distinguir entre dos tesis diferentes. La primera es el *platonismo matemático*, es decir, la tesis de que existen objetos matemáticos. (Uno podría agregar, si uno quisiera, la tesis de que los objetos matemáticos son abstractos, pero yo aquí me voy a enfocar en la tesis básica de que hay objetos matemáticos.)

La segunda tesis es el *trivialismo matemático*, es decir, la tesis de que las verdades de la matemática pura tienen condiciones de verdad triviales, y las falsedades de la matemática pura tienen condiciones de verdad imposibles; en otras palabras: no se

---

<sup>43</sup> Una excepción importante es el neo-fregeanismo. Ver, por ejemplo, Hale, Bob and Wright, Crispin (2001). *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*. Clarendon Press, Oxford.

requiere nada del mundo para que sea verdadera una verdad de la matemática pura, y se requeriría algo absurdo del mundo para que fuera verdadera una falsedad de la matemática pura.

La posición que quiero defender aquí es la conjunción del platonismo matemático y el trivialismo matemático. Lo que quiero defender, en otras palabras, es el *trivialismo platónico*.

Alguien que, como yo, acepta [Números], tiene buenas razones para ser platonista y buenas razones para ser trivialista. ¿Por qué ser trivialista? Discutimos el argumento en la primera sesión. Es un argumento que usa [Números] como premisa, y muestra que sería absurdo suponer que no hay números. Aquí lo tienen:

Supongamos, para reducción al absurdo, que no hay números. Que el número de los números sea cero es simplemente que no haya números. Por tanto, el número de los números es cero. Por tanto, el cero existe. Por tanto, hay números. ¡Contradicción!

Alguien que acepta [Números] cree, por tanto, que la exigencia de que el mundo contenga números es una exigencia trivial. No se requiere nada del mundo para que quede satisfecha.

¿Por qué aceptar el platonismo matemático? Supongamos que uno cree que no hay dinosaurios. Si uno acepta [Números], uno debe creer que el número de los dinosaurios es cero, y de esto se sigue que el cero tiene que existir. Dado que el cero es un número, se sigue que hay números y, por tanto, que el platonismo matemático es verdadero.

Vale la pena distinguir entre la posición que estoy defendiendo aquí y una posición de acuerdo con la cual 'el número de los dinosaurios es cero' no es literalmente verdadero. ¿En qué consistiría esta segunda posición? Tal vez la idea es que cuando aseveramos 'el número de los dinosaurios es cero', estamos hablando de números de manera metafórica, como cuando decimos que alguien tiene mariposas en el estómago.<sup>44</sup> O tal vez la idea es que 'el número de los dinosaurios es cero' es algo así como una abreviación sintáctica de 'no hay dinosaurios'. Alguien que acepte este tipo de posición diría que, estrictamente hablando, no estamos comprometidos con la existencia de los números, pero que podemos usar vocabulario aritmético para comunicar hechos que no tienen nada que ver con los números.<sup>45</sup>

Es importante tener claro que alguien así no aceptaría [Números]. Lo que es característico de mi posición no es que no haya números. Yo creo que los números

---

<sup>44</sup> Hay un buen desarrollo de este tipo de posición en Yablo, Stephen (2001) 'Go Figure: A Path Through Fictionalism' *Midwest Studies in Philosophy*, XXV, 72–102.

<sup>45</sup> Para una versión más sofisticada de este tipo de posición, ver Hofweber, T. (2005a). 'Number Determiners, Numbers and Arithmetic'. *The Philosophical Review*, 114:179–225.

existen. Lo que es característico de mi posición es que creo que la existencia de números es trivial: no requiere nada sustancial del mundo.

\* \* \*

Lo que me gustaría hacer ahora es decir algo sobre el tipo de metafísica a la que está comprometido el trivialismo platónico. Uno podría pensar, por ejemplo, que alguien que acepta [Números] está obligado a abandonar la idea de que hay un hecho objetivo acerca de cómo es el mundo, de que el mundo es independiente de nuestra manera de describirlo. Tal vez cuando describimos al mundo utilizando 'el número de los dinosaurios es cero' el mundo contiene números; pero cuando lo describimos diciendo 'no hay dinosaurios', el mundo no contiene números.

Eso sería un error. De acuerdo con el trivialismo platónico, el mundo es como es independientemente de cómo lo describimos. Esto significa, en particular, que hay números independientemente de si alguien asevera enunciados que contienen vocabulario aritmético. Si no hubiera vocabulario matemático, no hablaríamos explícitamente de objetos matemáticos, pero los objetos matemáticos seguirían existiendo.

Pero ¿no terminaríamos con una posición de acuerdo con la cual la existencia de los objetos está, en algún sentido, constituida por nuestro uso del lenguaje? ¿No estamos comprometidos con la idea de que el hecho de que el cero existe es, en algún sentido, dependiente de que describamos al mundo utilizando enunciados como 'el número de los dinosaurios es cero'? No. Lo que está constituido por nuestro uso del lenguaje es la existencia de *términos singulares*. Aunque es cierto que necesitamos términos singulares para hacer explícita la existencia de los números, los números mismos existen independientemente de qué enunciados utilicemos para describir el hecho de que hay números (o, equivalentemente, el hecho de que el número de los dinosaurios es cero).

¿No está comprometido el trivialismo platónico con que el mundo sea una masa desestructurada? No. La razón es muy sencilla. De acuerdo con el trivialismo platónico, es estricta y literalmente verdadero que el número de los dinosaurios es cero; y, hasta donde yo soy capaz de ver, eso es lo único que hace falta para que el mundo esté estructurado. Vale la pena notar, en particular, que es suficiente para garantizar que el mundo contenga objetos (como el número cero), y propiedades (como la propiedad de numerar a los dinosaurios).

Tal vez alguien podría creer que para que el mundo esté *realmente* estructurado hace falta más que eso. Tal vez lo que hace falta que el mundo esté provisto de lo que en sesiones anteriores habíamos llamado 'estructura metafísica'. Tal vez es gracias a esta 'estructura metafísica' que el mundo es más que una masa amorfa. Si es así como entendemos el asunto, el trivialismo platónico es *neutral* con respecto a la pregunta de si el mundo es una masa amorfa, porque el trivialismo platónico es neutral con respecto a la pregunta de si hay tal cosa como la 'estructura metafísica'.

La moraleja de todo esto es que aunque es cierto que el trivialismo platónico está comprometido con la idea de que un hecho puede conceptualizarse de maneras diferentes, no hay ninguna razón para pensar que esta idea nos comprometa con una metafísica problemática. El trivialismo platónico es compatible con una metafísica bastante común y corriente.

**Pregunta:** De acuerdo con el trivialismo platónico, ¿son o no abstractos los objetos matemáticos?

**Respuesta:** Entre los enunciados que podemos formular utilizando vocabulario aritmético, hay algunos como ‘hay infinitos números primos’ que son bien útiles, pero hay otros enunciados que solo a un filósofo se le ocurriría aseverar; por ejemplo, el ‘ $2 = \text{Julio César}$ ’ o ‘ $2$  está guardado en mi clóset’. Utilizando una frase de Wittgenstein, yo lo que diría es que en estos casos el lenguaje se ha ido de vacaciones.<sup>46</sup> Estamos ante enunciados que no juegan ningún papel en nuestro quehacer matemático, y mi opinión es que no tenemos ninguna razón para pensar que tienen condiciones de verdad bien definidas.

Creo que lo mismo sucede con enunciados como ‘ $2$  es un objeto abstracto’. Es un caso en el que el lenguaje se ha ido de vacaciones.

Un metafisicalista está comprometido con la idea de que incluso enunciados como ‘ $2 = \text{Julio César}$ ’ o ‘ $2$  es un objeto abstracto’ tienen que tener condiciones de verdad. Pero una de las dichas del composicionalismo es que solo tenemos que asignarle condiciones de verdad a los enunciados que queramos tener en ‘activo’, es decir, a los enunciados que quiera tener disponibles para la comunicación. Podemos dejar fuera a los enunciados que no prometan ser parte de una práctica teórica interesante.

**Pregunta:** ¿Cómo se distingue tu posición de un convencionalismo de acuerdo con el cual las verdades matemáticas son verdaderas por convención?

**Respuesta:** Creo que el convencionalismo tiene un elemento de verdad.

El metafisicalista diría que introducir un nuevo sistema axiomático es una empresa arriesgada, porque la verdad de nuestros nuevos axiomas va a depender de qué objetos arroje la estructura metafísica de mundo. Si tenemos mala suerte y la estructura metafísica del mundo no arroja números (o, por lo menos, una secuencia infinita de objetos), los axiomas de la aritmética estarán condenados a ser falsos.

---

<sup>46</sup> *Investigaciones Filosóficas*, párrafo 48.

De acuerdo con la posición que he estado defendiendo aquí, en cambio, es bastante fácil introducir axiomas verdaderos. Todo lo que uno tiene que hacer es estipular que uno va a usar cierta terminología de manera que los axiomas relevantes resulten verdaderos. Es fácil verificar que siempre y cuando los axiomas sean consistentes (o conservativos, en el caso de la matemática aplicada) será posible generar una asignación de condiciones de verdad que satisfaga las requerimientos composicionalistas de la segunda sesión. El composicionalista diría, por tanto, que al introducir un nuevo sistema axiomático no tenemos que realmente preocuparnos de cómo es el mundo (al menos en el caso de las matemáticas puras), y en esto estaría de acuerdo con el convencionalista.

\* \* \*

De acuerdo con el trivialismo platónico, todas las verdades de la matemática pura tienen condiciones de verdad triviales. Es tentador pensar que eso no puede estar bien. ¿No es obvio que cuando afirmo ' $2 + 2 = 4$ ' comunico algo bien diferente que cuando afirmo 'hay infinitos números primos'?

No voy a negar que esos dos enunciados juegan papeles bien diferentes en la comunicación, y en nuestras vidas teóricas. Tampoco voy a negar que uno puede creer que  $2 + 2 = 4$  sin creer que hay infinitos números primos. Creo que la conclusión que debemos derivar de todo esto, sin embargo, es que la noción de condiciones de verdad con la que estamos trabajando aquí no basta por sí sola para dar una teoría de la comunicación en matemáticas, o para dar una teoría de la creencia en matemáticas. Necesitamos herramientas filosóficas adicionales, y voy a hablar de algunas de ellas en la próxima sesión.

Cuando los filósofos teorizan acerca del significado, suelen comenzar con una noción informal de significado y presuponer que esa noción es la herramienta correcta para atacar varios proyectos diferentes; en particular: (1) el proyecto de entender cómo funciona la comunicación lingüística, (2) el proyecto de entender cómo funcionan nuestras creencias, y (3) el proyecto de entender qué se requiere del mundo para que un enunciado sea verdadero. Aquí estamos procediendo de modo muy diferente. Hemos puesto sobre la mesa una cierta herramienta -- la noción de condiciones de verdad -- y hemos trabajado duro para entender cómo funciona. Ahora que la tenemos un poco más clara, estamos empezando a utilizarla para atacar ciertos proyectos filosóficos, pero no estamos presuponiendo que la noción de condiciones de verdad es apropiada para atacar todos nuestros problemas filosóficos. Vamos a tratar de determinar cuándo es apropiada y cuando no, y utilizarla solo cuando resulte apropiada.

\* \* \*

Me gustaría hablar un poco sobre el papel que juegan las *paráfrasis* en el trivialismo platónico. Cuando uno se expone a un principio como [Números] es tentador pensar que la propuesta es algo así como lo siguiente:

Siempre que se me dé un enunciado en el lenguaje de la aritmética, yo voy a ser capaz de identificar un enunciado que no tenga un vocabulario aritmético y que describa el mismo hecho.

Esa no es la propuesta. Hay filósofos que han tratado de desarrollar propuestas de este tipo,<sup>47</sup> pero es resulta ser que hay dificultades de principio para llevar a cabo este proyecto. Hay un teorema que muestra que es *imposible* especificar un método general de paráfrasis para el lenguaje de la aritmética, dados supuestos razonables.<sup>48</sup> Si esto no es obvio, es porque algunos de los filósofos que discuten estos temas enfocan su discusión en enunciados de juguete como ‘el número de los niños es igual al número de las niñas’. Los ejemplos de juguete son bien fáciles de parafrasear en un lenguaje de orden superior. Pero eso no es lo que estamos buscando. Lo que estamos buscando es un método de paráfrasis general, y la horrible verdad es que, dadas restricciones razonables, no hay tal método.

Vale la pena señalar, sin embargo, que si el lenguaje en el que damos las paráfrasis es suficientemente exótico -- y una de nuestras ‘restricciones razonables’ es que no lo sea -- entonces sí que es posible especificar un método general de paráfrasis. Es posible, por ejemplo, definir un método general para parafrasear enunciados aritméticos en un lenguaje de orden  $\omega+3$  (es decir, un lenguaje que contiene cuantificadores de todo orden finito y además de cuatro órdenes infinitos).<sup>49</sup> Aquí no voy a tratar de defender la idea de que lenguajes de ese tipo son legítimos. El punto que quiero hacer es que la existencia de paráfrasis depende de qué recursos lingüísticos tengamos a nuestra disposición. Eso significa que la existencia de paráfrasis no es una buena manera de evaluar una filosofía de las matemáticas como la que estamos desarrollando aquí. Lo que estamos haciendo aquí es defender una tesis acerca de la naturaleza de los hechos matemáticos, y esta no es una tesis que no debería depender de los recursos lingüísticos que tengamos a nuestra disposición.

Es útil contrastar mi posición con la de un nominalista, es decir, con la de alguien que cree que no hay números. Supongamos que nuestro nominalista cree que los enunciados matemáticos presuponen la existencia de los números (y son, por tanto, falsos), pero también cree que podemos utilizar un enunciado matemático para comunicarnos, siempre y cuando pensemos en el enunciado como abreviación sintáctica de algún enunciado no matemático. Un nominalista como este sí que

---

<sup>47</sup> Ver, por ejemplo, Field, Hartry (1980). *Science Without Numbers*. Basil Blackwell y Princeton University Press, Oxford y Princeton; Hodes, Harold T. (1984). ‘Logicism and the ontological commitments of arithmetic’ *Journal of Philosophy*, 81, 123–149; y Yablo, Stephen (2002) ‘Abstract objects: A case study’ *Noûs*, 36, supp. 1, 255–286. Ofrezco una discusión detallada del tema en el capítulo 7 de *The Construction of Logical Space*.

<sup>48</sup> Esto lo muestro en ‘On specifying truth-conditions’. *The Philosophical Review*, 117, 385–443, 2008. Los supuestos razonables son: (1) que el lenguaje en el que se dan las paráfrasis sea un lenguaje de orden  $n$ , para  $n$  finito, (2) que la función de paráfrasis sea recursiva, (3) que respete valores de verdad, y (4) que esto último no dependa del supuesto de que el dominio del lenguaje en el que producimos nuestras paráfrasis sea infinito.

<sup>49</sup> Esto lo muestro en la sección 7.4 de *The Construction of Logical Space*.

necesita encontrar paráfrasis no matemática para cada enunciado matemático que quiera tener disponible para la comunicación.

Nosotros, en cambio, estamos en una situación bien diferente. De acuerdo con el trivialismo platónico, un enunciado como ' $2 + 2 = 4$ ' o 'hay infinitos números primos' es literal y estrictamente verdadero. Si lo queremos utilizar para decir cosas verdaderas, no necesitamos, por tanto, encontrar una paráfrasis.

Esto no significa, sin embargo, que estemos exentos de dar una *semántica* para los enunciados matemáticos. Aún cuando no sea parte del proyecto identificar paráfrasis no matemáticas para los enunciados matemáticos, es importante poder explicar cómo funciona el lenguaje matemático. En particular, es importante que estemos en posición de hacer lo siguiente:

Especificar una teoría semántica que le asigne a cada símbolo en el lenguaje de la aritmética un valor semántico y que nos permita derivar, a través de reglas composicionales, condiciones de verdad para cada enunciado del lenguaje (o por lo menos todo enunciado del lenguaje que queramos tener en activo).

No solo eso. Dado que somos trivialistas, queremos que todo enunciado verdadero de la matemática pura termine con condiciones de verdad triviales, y que todo enunciado falso de la aritmética pura termine con condiciones de verdad imposibles.

Si tuviéramos un método para parafrasear enunciados aritméticos utilizando vocabulario puramente lógico (y si se tratara de un método composicional), podríamos utilizarlo para construir nuestra teoría semántica. Pero, como hemos visto, ese proyecto es imposible, dados supuestos razonables.

Afortunadamente, existe una manera diferente de construir una teoría semántica como la que queremos. La observación crucial es que aunque nuestro proyecto es el de producir una teoría semántica que asigne condiciones de verdad trivialistas para el lenguaje objeto, esto no significa que nuestra metateoría -- la teoría que utilizamos para generar la asignación de condiciones de verdad -- no puede utilizar vocabulario matemático.

Aquí hay una manera de pensar en el asunto. Estamos distinguiendo entre los recursos que se necesitan para especificar condiciones de verdad en el metalenguaje y los compromisos de las condiciones de verdad especificadas para enunciados del lenguaje objeto. Como parte del ejercicio de especificar condiciones de verdad vamos a utilizar una teoría que está comprometida a objetos matemáticos, pero las condiciones de verdad especificadas van a ser condiciones de verdad trivialistas. Estamos, por tanto, desarrollando nuestra teoría semántica desde la perspectiva de alguien que entiende el lenguaje matemático, y que va a utilizar recursos matemáticos para especificar condiciones de verdad trivialistas.

**Pregunta:** En la semántica que vas a proponer, ¿le asignamos condiciones de verdad a enunciados como '2 = Julio César'?

**Respuesta:** No. La semántica le asigna condiciones de verdad a todo enunciado que podría interesarle a un matemático, pero no para enunciados que solo podrían interesarle a un filósofo.

\* \* \*

Los detalles de la semántica que les quiero proponer están el apéndice, pero la idea básica es muy sencilla. Consideremos la notación '[...]<sub>w</sub>', que se lee 'es verdad en el mundo posible *w* que ...'. Esto nos permite decir, por ejemplo, lo siguiente:

$\exists x (x \text{ es el moderador de mi sesión} \ \& \ [x \text{ tiene bigote}]_w)$

(Leer: hay algo tal que esa cosa es el moderador de mi sesión, y en el mundo *w* esa cosa tiene bigote)

Imaginemos que esta fórmula está precedida por la frase 'sea *w* tal que satisface la condición siguiente'. ¿Qué requisito se le está imponiendo a *w*? La única condición que tiene que satisfacer *w* es que Moisés Vaca -- la persona que, de hecho, está moderando mi sesión -- tenga bigote. No hace falta, sin embargo, que en *w* haya moderadores. Lo que estamos haciendo es que, nosotros -- los teóricos -- estamos identificando un objeto desde nuestra propia perspectiva, a través de la descripción, 'ser el moderador de esta sesión'. Una vez que hemos utilizado esa descripción para identificar un objeto en el mundo actual, decimos, de ese objeto, que existe en *w* y tiene bigote.

Contrastemos la fórmula anterior con la siguiente:

$[\exists x (x \text{ es el moderador de mi sesión} \ \& \ x \text{ tiene bigote})]_w$

(Leer: en el mundo *w* hay algo tal que esa cosa es el moderador de mi sesión y tiene bigote.)

Esta fórmula le impone una condición bien diferente a *w*. Lo que se requiere de *w* es que contenga un objeto que modere mi sesión y que use bigote. Para que *w* satisfaga esta condición no hace falta que *w* contenga a Moisés, pero sí hace falta que contenga sesiones y moderadores, y hace falta que el moderador de esa sesión -- quien podría ser, por ejemplo, Raymundo Morado -- tenga bigote.



Volvamos ahora a nuestra semántica trivialista, y enfoquémonos en el enunciado ‘el número de los dinosaurios es cero’, o, en su versión formalizada:

$$\#_x (\text{Dinosaurio}(x)) = 0$$

Este es un enunciado de nuestro lenguaje objeto, y queremos asignarle condiciones de verdad. Consideremos dos asignaciones de condiciones de verdad diferentes, ambas expresadas en nuestro metalenguaje:

### **Asignación Externa**

el número de las  $z$ s tales que  $[z \text{ es un dinosaurio}]_w = 0$

### **Asignación Interna**

$[\text{el número de las } z\text{s tales que } z \text{ es un dinosaurio} = 0]_w$

Estos dos enunciados metalingüísticos son bastante parecidos. Como en el caso de los ejemplos anteriores, difieren solo en la ubicación de los corchetes.

¿Cómo tendría que ser  $w$  para satisfacer la condición especificada en la Asignación Externa? Aquí hay una manera de pensar en el asunto. Primero nos fijamos qué cosas satisfacen, en  $w$ , la condición de ser dinosaurio. Luego contamos estos objetos y les asignamos un número, pero esto último lo hacemos en nuestro metalenguaje. En particular, no suponemos que  $w$  contiene números. Lo único que suponemos es que nosotros, como teóricos, somos capaces de hablar de números, y de asociar un número con los objetos que existen en  $w$  y satisfacen una cierta condición. Finalmente, nos preguntamos, desde nuestra propia perspectiva, si el número que le hemos asignado a los dinosaurios de  $w$  es igual a cero. Si sí, entonces concluimos que la condición especificada por la Asignación Externa ha quedado satisfecha por  $w$ ; en caso contrario, concluimos que la condición no ha quedado satisfecha. El resultado de todo esto es que lo que se requiere de  $w$  para que quede satisfecha la condición especificada por la Asignación Externa es que no contenga planetas.

La Asignación Interna es bien diferente. Para que se satisfaga esta condición lo que se requiere de  $w$  es que contenga un número, que ese número numere a los objetos que existen en  $w$  y sean dinosaurios, y que ese número sea igual a cero.

La semántica trivialista que les estoy proponiendo (y que hago explícita en el apéndice) especifica condiciones de verdad para ‘ $\#_x (\text{Dinosaurio}(x)) = 0$ ’ utilizando la Asignación Externa, y especifica condiciones de verdad para otros enunciados del metalenguaje utilizando análogos de esta misma idea. Esto permite que -- aún cuando nosotros, los teóricos, tengamos que hablar de objetos matemáticos para especificar condiciones de verdad -- las condiciones de verdad que se le asignan a enunciados del lenguaje objeto nunca hablan de números. A cada verdad de la aritmética pura se le asignan condiciones de verdad triviales, y a cada falsedad de la aritmética pura se le asignan condiciones de verdad imposibles.

**Pregunta:** Un trivialista piensa que la existencia de los números es trivial. Si esto es así, ¿por qué se necesita una semántica tan complicada? ¿No podríamos, por ejemplo, asignarle a ' $2 + 2 = 4$ ' la condición de que el mundo sea tal que  $2 + 2 = 4$ , y argumentar que esta es una condición trivial?

**Respuesta:** Tienes toda la razón. Para el trivialista no hay ninguna diferencia entre la condición de que el mundo sea tal que  $2 + 2 = 4$  y la condición trivial; y no hay ninguna diferencia entre la condición de que no haya dinosaurios y la condición de que el número de los dinosaurios sea cero. Desde el punto de vista del trivialista, por tanto, una asignación homofónica de condiciones de verdad -- una que le asigne a ' $2 + 2 = 4$ ' la condición de que el mundo sea tal que  $2 + 2 = 4$  -- arroja los resultados correctos. Pero ese no es un proyecto interesante. El proyecto interesante -- y el proyecto que nos interesa aquí -- es el de ver si es posible especificar condiciones de verdad trivialistas *sin presuponer que el trivialismo es correcto*. Una razón por la que este segundo proyecto es interesante es que la semántica resultante se puede utilizar para explicarle a un no trivialista -- o por lo menos a un no trivialista que esté dispuesto a utilizar vocabulario matemático -- cómo asignarle condiciones de verdad trivialistas a los enunciados aritméticos.

Lo último que les quiero decir antes de terminar es que es posible dar una semántica de este tipo no sólo para la aritmética, sino también para la teoría de conjuntos.<sup>50</sup> Esto es significativo porque casi toda la matemática contemporánea puede desarrollarse dentro de la teoría de conjuntos. No les voy a dar los detalles aquí, pero la idea básica es muy sencilla. Decimos por ejemplo: que exista un conjunto que tiene a Susana como elemento es simplemente decir que existe Susana. (Una complicación es que tenemos que proceder de manera que se eviten las paradojas conjuntistas. Afortunadamente, hay manera de hacerlo.)

## Apéndice

### *Una semántica trivialista para el lenguaje de la aritmética*

Denotación de términos aritméticos:

1.  $\delta_{\sigma W}(n_i) = \sigma(n_i)$
2.  $\delta_{\sigma W}(0) = \text{el número cero}$
3.  $\delta_{\sigma W}(S(t)) = \delta_{\sigma W}(t) + 1$
4.  $\delta_{\sigma W}(t_1 + t_2) = \delta_{\sigma W}(t_1) + \delta_{\sigma W}(t_2)$
5.  $\delta_{\sigma W}(t_1 \times t_2) = \delta_{\sigma W}(t_1) \times \delta_{\sigma W}(t_2)$
6.  $\delta_{\sigma W}(\#x_i(\phi(x_i))) = \text{el número de zs tales que } \text{Sat}(\phi(x_i), \sigma Z/x_i, w)$

---

<sup>50</sup> Esto lo muestro en el capítulo 3 de *The Construction of Logical Space*.

7.  $\delta_{\sigma w} (\#n_i (\phi (n_i)))$  = el número de  $ms$  tales que  $Sat(\phi(n_i), \sigma m/x_i, w)$

Denotación de términos no aritméticos:

1.  $\delta_{\sigma w} (x_i) = \delta (x_i)$
2.  $\delta_{\sigma w} (\text{César}) = \text{Cayo Julio César}$
3.  $\delta_{\sigma w} (\text{Tierra}) = \text{el planeta Tierra}$

Satisfacción:

Donde  $[...]_w$  se lee 'es verdad en  $w$  que  $\phi$ ',

1.  $Sat(\exists n_i \phi, \sigma, w) \leftrightarrow$  hay un número  $m$  tal que  $Sat(\phi, \sigma m/n_i, w)$
2.  $Sat(\exists x_i \phi, \sigma, w) \leftrightarrow$  hay una  $z$  tal que  $([\exists y(y = z)]_w \ \& \ Sat(\phi, \sigma z/x_i, w))$
3.  $Sat(t_1 = t_2, \sigma, w) \leftrightarrow \delta_{\sigma w} (t_1) = \delta_{\sigma w} (t_2)$
4.  $Sat(\text{Planeta}(t), \sigma, w) \leftrightarrow [\delta_{\sigma w} (t) \text{ es un planeta}]_w$  (donde  $t$  es un término no aritmético)
5.  $Sat(\phi \ \& \ \psi, \sigma, w) \leftrightarrow Sat(\phi, \sigma, w) \ \& \ Sat(\psi, \sigma, w)$
6.  $Sat(\neg \phi, \sigma, w) \leftrightarrow \neg Sat(\phi, \sigma, w)$

Aquí estoy simplificando de diversas maneras. Una de ellas es que estoy presuponiendo que, en el metalenguaje, tenemos la capacidad de cuantificar sobre objetos meramente posibles. Es posible, sin embargo, evitar esta presuposición. Hay más detalles en mi artículo 'On Specifying Truth-Conditions', *Philosophical Review*, 117, 385–443, 2008.

## Sesión 5: Cómo creer en verdades necesarias

*¿Cómo hemos de modelar logros cognitivos que no dividan al espacio de posibilidad*

---

En la sesión anterior defendí el siguiente enunciado de identidad:

### [Números]

Que el número de las Fs sea  $n$  es simplemente que haya  $n$  Fs.

y lo utilicé para defender el trivialismo platónico, es decir, la conjunción del platonismo matemático (la tesis de que existen los objetos matemáticos), y el trivialismo matemático (la tesis de que las verdades de la matemática pura tienen condiciones de verdad triviales, y las falsedades de la matemática pura tienen condiciones de verdad imposibles).

Lo que me gustaría hacer en la sesión de hoy es hablar de algo que se nos había quedado pendiente: el proyecto de desarrollar una *epistemología* para las matemáticas que resulte satisfactoria desde la perspectiva de un trivialista platónico.

Si el trivialismo platónico es verdadero, los enunciados verdaderos de la matemática pura describen hechos triviales. ¿Cómo dar cuenta, entonces, del hecho innegable de que el quehacer matemático es altamente no trivial? Probar un teorema es un logro cognitivo. ¿En qué consiste este logro cognitivo? Cuando hacemos un descubrimiento empírico (cuando descubrimos, por ejemplo, que el Instituto es azul) parte de lo que sucede es que descartamos regiones de nuestro espacio de posibilidad (por ejemplo, la región en la que el Instituto no es azul). Pero si las verdades de la matemática pura tienen condiciones de verdad triviales, no es claro que podamos modelar los logros cognitivos en matemáticas de esta manera, porque no hay ninguna región en nuestro espacio de posibilidad en la que no estén verificadas. El objetivo de la sesión de hoy es enfrentar este problema.

**Pregunta:** Estoy dispuesto a conceder que un enunciado como ' $p$  si y solo si  $p$ ' tiene condiciones de verdad triviales, pero las verdades de la matemática pura requieren que el mundo contenga números, y eso no parece nada trivial.

**Respuesta:** Estoy defendiendo una propuesta bien controversial. El trivialismo platónico exige que rechacemos el paradigma quineano: la tesis de que los compromisos ontológicos de un enunciado son una buena manera de medir qué tan demandantes son las condiciones de verdad del enunciado, o para medir qué tanto se exige del mundo para que el enunciado sea verdad.

El paradigma quineano es razonable desde el punto de vista de un metafisicista. En cambio, desde el punto de vista del composicionalista, (y, en particular, desde el punto de vista del platónico trivialista), es importante separar la pregunta de qué compromisos ontológicos tiene un enunciado de la pregunta qué tan demandantes son sus condiciones de verdad.

El puede ilustrarse con un enunciado como 'hay números'. De acuerdo con el criterio de compromisos ontológicos de Quine,<sup>51</sup> este enunciado está comprometido con la existencia de números. Si aceptáramos el paradigma quineano, por tanto, tendríamos que concluir que el mundo debe que satisfacer un requisito altamente no trivial para que el enunciado sea verdadero. Hemos visto en sesiones anteriores, sin embargo, que quien acepta [Números] tiene un argumento contundente para pensar que un mundo sin números sería inconsistente, y, por tanto, un argumento para pensar que el requisito de que el mundo contenga números es un requisito trivial. Tenemos, en otras palabras, un enunciado que está comprometido con la existencia de números pero tiene condiciones de verdad triviales.

Vale la pena tener en mente que el problema que vamos a discutir en la sesión de hoy es un problema que surge aunque uno no acepte el trivialismo platónico. La razón es muy sencilla. Incluso si rechazáramos el trivialismo platónico sería razonable pensar que las verdades *lógicas* tienen condiciones de verdad triviales, y esto nos deja con el problema de explicar en qué consisten los triunfos cognitivos del quehacer lógico.

El enunciado ' $p$  si y solo si  $p$ ' que mencionabas antes es una tautología muy sencilla. Consideremos un ejemplo un poco más complicado, como ' $(p \rightarrow q) \vee (q \rightarrow p)$ '. Este enunciado también es una tautología, pero no es inmediatamente obvio que lo sea. Tenemos, por tanto, un enunciado que tiene condiciones de verdad triviales a pesar de que el proyecto de determinar si el enunciado es verdadero es claramente no trivial. Surge por lo tanto el problema de explicar en qué consiste el logro cognitivo de quien lleva a cabo ese proyecto.

Si alguno de ustedes no está de acuerdo con el platonismo trivial, puede simplemente pensar en la sesión de hoy como una discusión sobre epistemología de la lógica. El problema de fondo es el mismo.

**Pregunta:** Creo que hay que hacer una distinción entre no requerir nada del mundo y requerir solo lo que trae el mundo por ser mundo. El requisito de que el mundo contenga números es un requisito del segundo tipo pero no del primero. ¿No es, en ese sentido, un requisito más fuerte que el requisito trivial que impone una tautología?

---

<sup>51</sup> Quine, Willard Van Orman (1948). 'On What There Is', *Review of Metaphysics*, 2. Reprinted in Quine, Willard Van Orman (1953) *From a Logical Point of View*. Harvard University Press, Cambridge, MA.

**Respuesta:** Creo que todo depende de cómo midamos fuerza. Una manera de medir fuerza la fuerza de una aseveración es diciendo que entre más pequeña es la región del espacio de posibilidad que verificaría a la aseveración más fuerte es. De acuerdo con esta manera de pensar en fuerza, el requisito de que el mundo contenga números es exactamente tan fuerte como el requisito que impone una tautología, por lo menos desde el punto de vista del trivialismo platónico.

Tal vez uno podría articular una noción diferente de fuerza. Uno podría decir, por ejemplo, que entre más enunciados de identidad aceptemos, más fuerte es nuestra concepción del mundo. Alguien que acepte [Números] estaría, por tanto, comprometido con una concepción del mundo más fuerte que alguien que no lo acepte. De acuerdo con esta manera de pensar en fuerza, sin embargo, una concepción del mundo más fuerte que la nuestra debería parecernos no solo falsa sino inaceptable, porque la veríamos como asimilado distinciones que de hecho son diferentes.

**Pregunta:** Parece que hay una tensión en el trivialismo platónico. La tesis platonista sugiere que los objetos matemáticos están en el mundo, como las sillas y las mesas. Pero de acuerdo con el trivialismo, el mundo podría ser de cualquier manera, y las verdades matemáticas seguirían siendo verdaderas. Eso sugiere que el trivialista platónico es solo platonista de dientes para fuera. ¿Qué caso tiene postular objetos matemáticos como parte del mundo, si eso no le impone constreñimientos a la manera de ser del mundo?

**Respuesta:** Creo que el metafisicalista estaría de acuerdo contigo. El metafisicalista cree que nuestro discurso objetual (es decir, nuestro uso de términos singulares y variables que ocupan las posiciones de términos singulares) está metafísicamente cargado. En otras palabras, el hecho mismo de utilizar términos singulares significa que nuestras aseveraciones le imponen constreñimientos sustanciales al mundo. De acuerdo con el composicionalista, en cambio, esto no es así. El uso de términos singulares no es más que un mecanismo para generar, de manera composicional, una gran variedad de enunciados. No podemos presuponer, por tanto, que por el hecho mismo de utilizar términos singulares sea suficiente para que nuestras aseveraciones le impongan constreñimientos sustanciales al mundo.

Lo que está en el fondo de todo esto es que el metafisicalista y el composicionalista tienen concepciones diferentes de la noción de objeto. De acuerdo con el metafisicalista la noción de objeto es, al final del día, una noción metafísica. Quiere decir algo así como 'elemento de la conceptualización metafísicamente canónica de los hechos que constituyen al mundo'. En cambio, si el composicionalista tratar de explicar qué es un objeto, lo más que podría hacer es aseverar un enunciado como ' $\forall x$  (Objeto(x))'. De acuerdo con

el composicionalista, la noción de objeto es una noción esencialmente lingüística; en términos fregeanos, ser un objeto es simplemente ser la referencia de un término singular posible.

El trivialista platónico acepta [Números]. De esto se sigue que tiene que ser un composicionalista, y que su noción de objeto tiene que ser la del composicionalista. Estoy de acuerdo en que se trata de una noción de objeto menos fuerte que la del metafisicalista, pero a mí esto me parece una virtud de la propuesta. Creo que los filósofos se han metido en problemas por asumir que la noción de objeto está metafísicamente cargada.

\* \* \*

El reto epistémico tradicional en filosofía de las matemáticas es el dilema de Benacerraf.<sup>52</sup> Lo que dice el dilema es que solo podemos explicar cómo es posible el conocimiento matemático si hacemos una de dos cosas: o bien postulamos una semántica no estándar para los enunciados de la aritmética (una de acuerdo con la cual no estén realmente comprometidos con objetos matemáticos), o bien explicamos cómo es posible tener conocimiento del mundo abstracto y causalmente inerte en el que habitan los objetos matemáticos.

De acuerdo con el trivialismo platónico, este no es un dilema genuino, porque padece de una presuposición falsa. Presupone que hay espacio teórico entre el hecho de que no haya dinosaurios y el hecho de que el número de los dinosaurios sea cero. Presupone, por tanto, un espacio teórico que tendríamos que llenar de alguna manera (tal vez postulando que tenemos una facultad cuasi-perceptual que nos permite 'observar' el mundo abstracto de los objetos matemáticos).<sup>53</sup>

Alguien que acepta [Números] no cree que haya tal espacio teórico. Cuando uno sabe que el mundo es tal que no hay dinosaurios, uno ya conoce el hecho que tenía que conocer para saber que el número de los dinosaurios es cero. Es posible que uno no se de cuenta de que ya conoce el hecho relevante, pero para librar ese obstáculo lo único que hace falta es convencerse de que [Números] es correcto. No hace falta investigar 'un mundo abstracto y causalmente inerte'.

El trivialista platónico no se enfrenta, por tanto, al reto epistémico tradicional de la filosofía de las matemáticas. Eso no significa, sin embargo, que no se enfrente a un reto epistémico. Lo que sucede es que el trivialismo platónico cambia el reto tradicional por uno diferente. El nuevo reto es el que mencionábamos antes: si las verdades la matemática pura tienen condiciones de verdad triviales, ¿cómo podemos modelar los logros cognitivos del quehacer matemático?

---

<sup>52</sup> Benacerraf, Paul 'Mathematical truth' *Journal of Philosophy*, 70, 661–79, 1973.

<sup>53</sup> Aquí una lectura relevante es Gödel, K., 1944. 'Russell's Mathematical Logic', en Benacerraf, P. y Putnam, H. (eds.), 1983. *Philosophy of Mathematics: Selected Readings*, Cambridge: Cambridge University Press, segunda edición.

El objetivo de la sesión de hoy es enfrentarnos a este nuevo reto. (Como vimos antes, es un reto al que no solo se enfrenta el trivialista platónico; es un reto al que nos enfrentamos todos cuando tratamos de modelar los logros cognitivos del quehacer lógico.)

\* \* \*

Hay por lo menos dos maneras diferentes de pensar en la noción de proposición. De acuerdo con la primera, usamos la noción de proposición para rastrear *trozos de información*, es decir, para rastrear regiones del espacio de posibilidad. Si pensamos en el espacio de posibilidad como el conjunto de mundos posibles, podemos, por tanto, modelar una proposición como un conjunto de mundos posibles. Una consecuencia de esta manera de pensar es que 'Héspero es Héspero' y 'Héspero es Fósforo' expresan la misma proposición, porque están verificados por la misma región del espacio de posibilidad.<sup>54</sup>

De acuerdo con la segunda manera de pensar en la noción de proposición, las proposiciones se utilizan para modelar *logros cognitivos*, no trozos de información. Esto significa que 'Héspero es Héspero' y 'Héspero es Fósforo' no expresan la misma proposición porque corresponden a logros cognitivos diferentes: alguien puede creer que Héspero es Héspero sin creer que Héspero es Fósforo. Quien piensa en la noción de proposición de esta manera podría pensar que las proposiciones son algo así como 'pensamientos fregeanos'.<sup>55</sup>

Para nuestros propósitos no es importante que escojamos entre estas dos nociones de proposición. Lo que importa es que tengamos claro que hay dos tareas que son importantes. Una tarea es la de rastrear información acerca del mundo; otra es la de rastrear logros cognitivos. Al final del día, necesitamos que nuestro herramental teórico nos permita hacer las dos cosas.

En sesiones anteriores hemos estado trabajando duro para caracterizar el espacio de posibilidad. El resultado de nuestros esfuerzos es que tenemos a nuestra disposición las herramientas necesarias para modelar de trozos de información. Podemos pensar en ellos como regiones del espacio de posibilidad, o como conjuntos de mundos posibles. Lo que vamos a hacer aquí es comenzar con estas herramientas, y utilizarlas para construir la herramientas necesarias para modelar diferencias entre distintos logros cognitivos.

Consideremos un ejemplo. Godínez sabe que tiene un terreno cuadrado. Está, sin embargo, en una situación curiosa. Si le preguntáramos '¿Cuánto mide tu terreno de largo?', diría 'nueve metros'; pero si le preguntáramos '¿Cuánto tiene de área?', diría 'sesenta y cuatro metros cuadrados'. Sobra decir que no hay tal cosa como un terreno cuadrado que tenga nueve metros de largo y sesenta y cuatro metros cuadrados de

---

<sup>54</sup> Famosamente, Bob Stalnaker piensa en proposiciones de esta manera. Ver, por ejemplo, Stalnaker, Robert C. (1984). *Inquiry*. MIT Press, Cambridge, MA.

<sup>55</sup> Ver, por ejemplo, Frege, G. 'The Thought: A Logical Inquiry', *Mind*, 65 (1956), pp. 289-311.



área (a menos que estemos en un espacio no Euclidiano, pero eso es algo que voy a ignorar aquí).

¿Cómo modelar el sistema cognitivo de Godínez? Una opción es utilizar proposiciones en el segundo sentido: proposiciones como rastreadores de logros cognitivos. Podríamos decir, por ejemplo, que hay dos proposiciones distintas que Godínez cree: por un lado, la proposición de que el terreno mide nueve metros de largo; por otro, la proposición de que el terreno tiene sesenta y cuatro metros cuadrados de área. Hay algo, sin embargo, que hemos dejado fuera. Nada de lo que hemos dicho hasta ahora nos permite contestar a la pregunta de cómo es el mundo de acuerdo con Godínez. ¿Dónde se posicionaría Godínez en el espacio de posibilidad? Las proposiciones que le hemos atribuido a Godínez no bastan para contestar a esta pregunta, porque no hay ningún punto en el espacio de posibilidad en el que el terreno de Godínez tiene nueve metros de largo y sesenta y cuatro metros cuadrados de área.

Aquí hay otra manera de pensar en el asunto. Imaginemos que Dios quiere hacer que el mundo sea exactamente como Godínez cree que es. ¿Cómo habría de proceder? Ni siquiera Dios podría construir un mundo en el que Godínez tiene un terreno de nueve metros de largo y sesenta y cuatro metros cuadrados de área, porque ese sería un mundo inconsistente. Peor aún: un mundo inconsistente no le haría justicia a las creencias de Godínez, porque Godínez no cree que vive en un mundo inconsistente.

Para hacerle justicia a las creencias de Godínez, Dios necesita más información. Tal vez Godínez midió cuidadosamente el largo de su terreno, pero cometió un error matemático tratando de calcular el área. En ese caso es natural pensar que de acuerdo con Godínez el mundo es tal que su terreno tiene nueve metros de largo (y por lo tanto ochenta y un metros cuadrados de área). O tal vez Godínez midió cuidadosamente el área de su terreno, y cometió un error matemático tratando de calcular el área. En ese caso es natural pensar que de acuerdo con Godínez el mundo es tal que su terreno tiene ocho metros de largo (y por lo tanto sesenta y cuatro metros cuadrados de área).

La moraleja es que cuando nos limitamos a utilizar proposiciones como rastreadores de logros cognitivos hay algo que dejamos fuera.

Si nos limitáramos a pensar en proposiciones como rastreando trozos de información, sin embargo, nos enfrentaríamos a un problema dual. Consideremos, por ejemplo, una situación en la que Godínez midió el largo de su terreno cuidadosamente, pero se equivocó calculando el área. Tal vez en ese caso podríamos atribuirle a Godínez la información de que su terreno tiene nueve metros de largo (y, por tanto, ochenta y un metros cuadrados de área). El problema es que eso, por sí solo, no nos permite saber qué es lo que Godínez está en posición de hacer con el trozo relevante de información. Imaginemos que va a la tienda a comprar azulejos para cubrir su terreno. ¿Cuántos azulejos de metro cuadrado va a comprar? El hecho de que Godínez posee la información de que su terreno tiene ochenta y un metros cuadrados de área no es suficiente para concluir que va a comprar ochenta y un azulejos. De hecho, Godínez está en posición de utilizar la información relevante para decidir cuántos metros de

barda comprar si quiere bardar un lado de su terreno, pero no está en posición de usarla para decidir cuántos azulejos comprar. Saber qué información posee un sistema cognitivo -- saber cómo es el mundo de acuerdo con ese sistema cognitivo -- no es suficiente para saber qué tipo de uso está en posición de hacer de esa información.

La conclusión de todo esto que para modelar la situación cognitiva de una persona no basta utilizar proposiciones como rastreadores de logros cognitivos y no basta con utilizar proposiciones como rastreadores de pedazos de información. Necesitamos recursos de los dos tipos.

**Pregunta:** ¿No podríamos decir que las matemáticas juegan un papel en la organización de nuestro marco conceptual?

**Respuesta:** Creo que sí. De acuerdo con la posición que voy a defender aquí, los logros cognitivos en matemáticos consisten, en parte, en la introducción de 'puentes' entre los distintos propósitos con respecto a los cuales podría estar disponible un cierto pedazo de información.

Cuando, por ejemplo, Godínez descubre que  $8 \times 8 = 64$ , esto le da un puente para lograr que información que tenía disponible para el propósito de responder a la pregunta '¿Tiene tu terreno ocho metros de largo?' esté también disponible para el propósito de responder a la pregunta '¿Tiene tu terreno sesenta y cuatro metros cuadrados de área?'.

\* \* \*

Para modelar el sistema cognitivo de Godínez, tenemos que reconocer dos tipos diferentes de logros cognitivos. El primero consiste en la adquisición de información acerca de cómo es el mundo. El segundo consiste en la adquisición de una capacidad para utilizar esa información para proyectos particulares.

Godínez tiene información acerca del tamaño de su terreno, y para ciertos propósitos tiene acceso a esa información. (Tal vez tiene acceso a la información para propósito de responder a la pregunta "¿qué tan largo es tu terreno?", o para el propósito de decidir cuántos metros de barda comprar para bardar su terreno.) Hay, sin embargo, propósitos diferentes con respecto a los cuales Godínez no tiene acceso a esa misma información. (Tal vez el propósito de contestar a la pregunta "¿cuál es el área de tu terreno?", o el propósito de decidir cuántos azulejos comprar.)

¿Cómo podríamos utilizar la maquinaria que les estoy proponiendo para modelar el logro cognitivo de Godínez cuando logra calcular (correctamente) el área de su terreno a partir del largo de su terreno? Según la propuesta que quiero defender aquí, el logro cognitivo de Godínez consiste no en la adquisición de información adicional acerca de cómo es el mundo, sino en la adquisición de una habilidad para utilizar

información que ya estaba en su sistema cognitivo para nuevos propósitos; por ejemplo, el propósito de contestar a la pregunta '¿cuál es el área de tu terreno?' o el propósito de determinar cuántos azulejos comprar.

Aquí hay una manera más general de poner el punto. La pregunta a la que nos enfrentamos es: ¿cómo modelar los logros cognitivos en lógica y matemáticas? Tenemos sobre la mesa herramientas de dos tipos: por un lado, la noción de un trozo de información (es decir, una región del espacio de posibilidad, o un conjunto de mundos posibles); por otro, una noción de accesibilidad---la accesibilidad de un trozo de información para los propósitos de un cierto tipo de tarea. De acuerdo con mi propuesta, necesitamos la segunda herramienta para modelar los logros cognitivos en lógica y matemáticas. Estos logros deben ser modelados, en parte, como la adquisición de habilidades de *transferencia* de información, es decir, como habilidades para conseguir que información que ya estaba disponible para los propósitos de un cierto tipo de tarea esté también disponible para tareas de otros tipos.

Me gustaría enfatizar que no estoy tratando de convencerlos de que las herramientas que estoy utilizando aquí corresponden a nuestra práctica común y corriente de atribución de creencias. Estamos simplemente utilizando los conceptos teóricos que desarrollamos en sesiones anteriores para modelar ciertos tipos de logros cognitivos. Reconozco que no hay una conexión obvia entre nuestra práctica de atribución de creencias y el modelo que les estoy proponiendo aquí, pero más adelante voy a tratar de decirles cómo es que creo que están relacionados.

También me gustaría enfatizar que muchas de las ideas que estoy poniendo sobre la mesa tienen su origen en el trabajo de Bob Stalnaker, y en un proyecto en el que llevo varios años trabajando Adam Elga. Ninguno de ellos tiene la culpa de los errores que seguramente estoy cometiendo aquí, pero mucho de lo que estoy diciendo aquí se lo debo a ellos.

**Pregunta:** ¿Podrías decir más acerca de qué es un 'propósito'?

**Respuesta:** Pensemos en el proyecto de modelar un sistema cognitivo como un proyecto de interpretación radical. Hacemos dos cosas al mismo tiempo. Por un lado, postulamos un conjunto de 'propósitos'; por otro postulamos una función que le asigna a cada propósito la información que el sujeto tiene disponible para ese propósito. ¿Y qué hace falta para que estos postulados sean correctos? Decimos que una asignación de propósitos e información-relativa-a-propósitos es 'correcta' en tanto que resulte útil para predecir y describir el comportamiento nuestro sujeto.

Estoy proponiendo, por tanto, un cierto tipo de instrumentalismo. Vale la pena señalar, sin embargo, que mi instrumentalismo no es esencial a la posición general que he estado defendiendo aquí. A final del día, nuestra respuesta a la pregunta de cómo establecer criterios de identidad para 'propósitos' dependerá de nuestra respuesta a lo que Brentano llamó el problema de la

intencionalidad.<sup>56</sup> La propuesta que estoy defendiendo aquí es compatible con diferentes respuestas al problema de la intencionalidad, incluyendo respuestas no instrumentalistas.

\* \* \*

He estado desarrollando un modelo para dar cuenta de nuestros logros cognitivos en lógica y matemáticas. Es interesante notar que podríamos redescibir el modelo hablando de modos de presentación. El término 'modo de presentación' no es muy útil por sí solo; es simplemente una etiqueta para un concepto teórico que queda por desarrollar, pero la maquinaria que hemos estado utilizando aquí nos da una buena manera de dar contenido a la etiqueta.

Aquí está la propuesta. Decir que un trozo de información se cree bajo un cierto modo de presentación es simplemente decir que el sujeto tiene acceso a esa información para los propósitos de una cierta tarea. En otras palabras, los diferentes modos de presentación de un trozo de información corresponden a diferentes tipos de propósitos con respecto a los cuales un sujeto dado podría tener acceso a esa información.

Supongamos, por ejemplo, que Godínez tiene acceso a la información de que su terreno tiene nueve metros de largo (es decir, ochenta y un metros cuadrados de área) para los propósitos de decidir cuánta barda comprar. Esto lo podríamos modelar diciendo que Godínez conoce la proposición-como-pedazo-de-información de que el terreno tiene nueve metros de largo bajo el modo de presentación correspondiente a la tarea de decidir cuánta barda comprar. Y cuando Godínez hace un cálculo que le permite utilizar la información relevante para decidir cuántos azulejos comprar, podemos decir que conoce esa misma proposición bajo un nuevo modo de presentación.

Hay modos de presentación que corresponden a tareas lingüísticas. Godínez puede, por ejemplo, conocer una cierta proposición bajo el modo de presentación que corresponde a la tarea de responder a la pregunta '¿Cuál es el área de tu terreno, en metros cuadrados?'. Los modos de presentación lingüísticos son un caso particular, pero un caso particular que va a resultar bien importante cuando discutamos la matemática pura.

Hay otros modos de presentación que no tienen nada que ver con el lenguaje. Podríamos imaginar, por ejemplo, que Godínez es un sujeto no-lingüístico, y que a pesar de ello es capaz de decidir cuántos metros de barda comprar para bardar su terreno. Aquí hay otro ejemplo. Yo tengo cierta información en mi aparato cognitivo; a saber, la información de que un cierto tipo de procedimiento es un método eficaz para

---

<sup>56</sup> Ver Brentano, Franz, *Psychology from an Empirical Standpoint*, London: Routledge and Kegan Paul. Ver también Stalnaker, Robert C. (1984). *Inquiry*. MIT Press, Cambridge, MA.

atar agujetas. Conozco la proposición relevante -- el conjunto de mundos posibles en los que ese procedimiento es eficaz -- bajo un modo de presentación que corresponde a la tarea de atar mis propias agujetas. Pero supongamos que no soy capaz de atar las agujetas de Raymundo Morado, cuyo zapato está posicionado a la inversa que el mío, dado que Raymundo está sentado delante de mí. En ese caso diríamos que la proposición relevante no está disponible bajo el modo de presentación que corresponde a la tarea de atar las agujetas de Raymundo. Tenemos, pues, un ejemplo en el que consideraciones lingüísticas son claramente irrelevantes.

También hay casos híbridos. Cuando alguien descubre que Héspero es Fósforo, adquiere información empírica que no tiene nada que ver con el lenguaje; en particular, la información de que hay un planeta y no dos. Pero, en ciertos casos, uno también adquiere información lingüística. Uno adquiere la información de que dos términos singulares de nuestro lenguaje son co-referenciales.

\* \* \*

Ha llegado el momento de discutir el caso de las matemáticas puras. Mi propuesta es bastante sencilla, pero las herramientas que voy a utilizar son bastante diferentes a los que utilizamos cuando describimos logros matemáticos en el lenguaje ordinario. Es importante tener en cuenta, sin embargo, que nuestro proyecto no consiste en regimentar la manera ordinaria de hablar. No estamos partiendo de atribuciones comunes y corrientes de creencias matemáticas, y utilizándolas como las piezas básicas de una teoría más cuidadosa. Vamos a construir nuestro modelo utilizando herramientas teóricas que hemos estado desarrollando en sesiones anteriores.

La propuesta tiene dos partes. Consideraremos, primero, el logro cognitivo de reconocer a un axioma como verdadero y, después, el logro cognitivo de llevar a cabo una deducción.

Cuando alguien reconoce un axioma como verdadero, suceden dos cosas. Primero, el sujeto adquiere cierta información. ¿Qué información? Información *lingüística* -- específicamente, la información de que el vocabulario matemático se utiliza de modo tal que el axioma relevante resulta verdadero. Lo segundo que sucede es que esta información pasa a estar disponible para ciertos propósitos. ¿Qué propósitos? Una familia de tareas lingüísticas de las que un ejemplo paradigmático es la tarea de contestar a la pregunta '¿Cuáles son los axiomas?' o '¿Es verdadero tal-y-tal enunciado?', donde 'tal-y-tal enunciado' es el axioma relevante.

Consideremos ahora el caso de la deducción. ¿Cómo vamos a modelar el logro cognitivo de haber determinado que un cierto enunciado es consecuencia lógica de otros? Lo modelamos como una habilidad para *transmitir* información. Supongamos, por ejemplo, que uno ha deducido el enunciado  $\pi$  a partir de los axiomas. Esto lo vamos a modelar como la adquisición de una habilidad para usar la información lingüística que antes habíamos asociado con los axiomas -- información acerca de

cómo se utiliza el vocabulario matemático -- para nuevos propósitos. ¿Qué propósitos? Una familia de tareas lingüísticas de las que un ejemplo paradigmático es la tarea de responder a la pregunta '¿Es el caso que  $\pi$ ?', o la tarea de aseverar el enunciado  $\pi$  en el contexto adecuado.

\* \* \*

Una objeción inmediata a esta propuesta es que parece demasiado lingüística. Este es un punto que vale la pena discutir.

Supongamos que le pregunto a Pedro si hay algún tapir en el aula. ¿Qué información necesitaría poseer Pedro para estar en posición de contestar a mi pregunta? Necesita dos tipos de información. Primero, necesita información lingüística, porque para poderme dar una respuesta, tiene que empezar por *entender* mi pregunta. En particular, tiene que entender la palabra 'tapir'. (Escogí 'tapir' porque tal vez algunos de ustedes no sepan qué es un tapir, y eso haría especialmente obvio que Pedro necesita información lingüística para estar en posición de responder a mi pregunta.) El segundo tipo de información que Pedro necesita es información no-lingüística acerca de cómo es el mundo; necesita saber si hay o no algún tapir en esta aula.

Una característica de nuestro ejemplo es que para alguien como Pedro, que entiende la palabra 'tapir', el procesamiento lingüístico que se requiere para responder a mi pregunta es absolutamente transparente. Cuando Pedro contesta a mi pregunta, su atención consciente estará enteramente enfocada en la pregunta no-lingüística de si hay algún tapir en el aula.

Consideremos ahora un caso en el que el procesamiento lingüístico que se requiere para entender una pregunta es un poco más complejo. Imaginen que le pregunto a Pedro si hay tapires que no sean tales que no sea el caso que no sean no-mamíferos. Como en el caso anterior, Pedro está en posición de contestar a mi pregunta. En este caso, sin embargo, Pedro le va a tener que dedicar bastante más tiempo al proyecto de desenredar mi oración para entender qué es lo que estoy preguntando. De hecho, mi pregunta consiste de una oración con dos dobles negaciones, así que lo que estoy preguntando es simplemente si hay tapires que sean mamíferos, y una vez que Pedro vea eso, no tendrá dificultad alguna en contestar a mi pregunta. Pedro tenía desde el principio la información de que los tapires son mamíferos, pero es solo una vez que ha llevado a cabo el procesamiento lingüístico relevante que está en posición de utilizar esa información para el propósito de contestar a mi pregunta.

Lo que quisiera proponer aquí es que el tipo de logro cognitivo que llevamos a cabo cuando probamos un teorema matemático involucra procesamiento lingüístico de este tipo. Así como Pedro tuvo que desenredar mi pregunta para entender qué es lo que le estaba preguntando, parte de lo que hacemos cuando nos convencemos de una verdad lógica o matemática es que desenredamos la oración correspondiente.

Supongamos, por ejemplo, que le hago a Pedro una pregunta de la forma ' $\zeta(p > q) \vee (q > p)?$ ', donde ' $>$ ' es el condicional material. Con suficiente esfuerzo, Pedro podría desenredar mi pregunta. Lo que descubriría es que ' $(p > q) \vee (q > p)$ ' es una tautología, y, por tanto, que tiene condiciones de verdad triviales. Este es un caso en el que prácticamente todo el esfuerzo de Pedro consistiría en procesamiento lingüístico, porque una vez que ha identificado qué es lo que estoy preguntando, y se da cuenta de que sus condiciones de verdad son triviales, no tendrá que hacer ninguna investigación adicional acerca de cómo es el mundo para estar en posición de responder a mi pregunta.

Cuando le pregunté a Pedro si hay algún tapir en el aula, el procesamiento lingüístico correspondió a una parte relativamente pequeña del esfuerzo que tuvo que llevar a cabo Pedro para responder a mi pregunta. Cuando le pregunté si hay tapires que no sean tales que no sea el caso que no sean no-mamíferos, Pedro tuvo que dedicarle una mayor proporción de sus esfuerzos al procesamiento lingüístico. De acuerdo con la propuesta que estoy defendiendo aquí, las preguntas lógicas y matemáticas son ejemplos extremos de este tipo de fenómeno. Son ejemplos en los que prácticamente todo el esfuerzo se va en llevar a cabo el procesamiento lingüístico necesario para determinar cuáles son las condiciones de verdad del enunciado relevante.

Es importante notar, sin embargo, que no es parte de la propuesta que cuando hacemos matemáticas estamos hablando *sobre* lenguaje. No es parte de la propuesta que el lenguaje es el *asunto* sobre el que tratan las matemáticas.

Creo que los logros cognitivos en matemáticos son, en parte, logros lingüísticos, pero eso no significa que la matemática sea sobre el lenguaje. Según mi posición, las verdades de la matemática pura tienen contenidos triviales. De eso se sigue que la matemática no es sobre ningún tema particular (y, por tanto, que no es sobre los aspectos lingüísticos del mundo). La razón por la que no es trivial determinar si un enunciado matemático es verdadero es que no es trivial determinar si su contenido es trivial. Hacerlo requiere procesamiento lingüístico de exactamente el mismo tipo que Pedro tendría que llevar a cabo para desenredar mi pregunta con dos dobles negaciones, pero así como la pregunta que le hice a Pedro no es sobre aspectos lingüísticos del mundo, tampoco son sobre aspectos lingüísticos del mundo los enunciados de la matemática pura.

Uno podría objetar que el modelo que les estoy proponiendo no hace justicia a la fenomenología de las matemáticas. Cuando trabajamos en un problema matemático, la sensación que tenemos no es la de estarnos preguntando cómo funciona el lenguaje relevante. La sensación que tenemos -- o por lo menos la que tengo yo -- es la de investigar el mundo de los objetos matemáticos.

Es cierto que de acuerdo con mi propuesta los logros lógicos y matemáticos son en buena medida logros lingüísticos, pero creo que cuando de procesamiento lingüístico se trata, nuestra fenomenología no es un mecanismo muy fiable para determinar qué es lo que realmente está ocurriendo. En el caso de la matemática pura creo que a lo

que responde nuestra fenomenología es a los trucos heurísticos que utilizamos para facilitar el procesamiento lingüístico. Cuando uno hace matemáticas es natural imaginarse una colección de objetos, y apoyarse en ellos para desenredar las oraciones con las que uno está trabajando. Nuestra fenomenología responde a elementos heurísticos de este tipo.

\* \* \*

En la vida cotidiana decimos cosas como 'Pedro cree que hay infinitos números primos', pero no hablamos ni de regiones del espacio de posibilidad, ni de modos de presentación. Lo que me gustaría hacer ahora es vincular el tipo de modelo que hemos estado discutiendo aquí con las atribuciones comunes y corrientes de creencias matemáticas.

Consideremos una atribución de creencias de la forma  $[S \text{ cree que } p]$ . De acuerdo con una posición muy básica, lo que hacemos cuando aseveramos un enunciado como este es que nos comprometemos con que la proposición que expresa el enunciado  $p$  esté entre las creencias de  $S$ .

De acuerdo con la propuesta que me gustaría defender aquí, las atribuciones de creencia funcionan de manera un poco más compleja, porque que  $p$  juega dos papeles diferentes. El primer papel consiste en identificar una región del espacio de posibilidad (es decir, un trozo de información). Una manera de usar  $p$  para identificar un trozo de información es fijándonos en el conjunto de mundos posibles que verificarían a  $p$ . Vale la pena notar, sin embargo, que esa no es la única manera; también podríamos usar  $p$  para identificar lo que Stalnaker llama 'la proposición diagonal'.<sup>57</sup>

La segunda tarea de  $p$  es darle a la audiencia una idea del tipo de propósitos con respecto a los cuales el sujeto tiene acceso al trozo de información que antes habíamos usado a  $p$  para identificar. En otras palabras, parte de la razón por la que puede haber diferencia entre aseverar  $[Pedro \text{ cree que } p]$  y aseverar  $[Pedro \text{ cree que } q]$  aún cuando  $p$  y  $q$  sean lógicamente equivalentes es que el uso de  $p$  en lugar de  $q$  le dice algo a la audiencia acerca del tipo de propósitos con respecto a los cuales Pedro es capaz de utilizar la información relevante. Paradigmáticamente, el uso de  $p$  sugiere que Pedro es capaz de utilizar la información para responder a la pregunta  $[\text{¿Es cierto que } p?]$ .

En el caso general, las atribuciones de creencias requieren que  $p$  juegue estos dos papeles. Hay, sin embargo, casos particulares en los que uno de los dos papeles es más importante que el otro. Cuando alguien dice, por ejemplo, 'Pedro cree que la nieve es blanca', el trabajo principal de 'la nieve es blanca' consiste en identificar un trozo de

---

<sup>57</sup> La proposición diagonal que le corresponde a  $p$  es el conjunto de mundos  $w$  tal que  $w$  es compatible con las presuposiciones de la conversación y tal que, en  $w$ , el vocabulario de  $p$  se usa de tal modo que  $p$  resulta verdadero). Ver Stalnaker, Robert C. (1979) 'Assertion' *Syntax and Semantics* 9, 315–322.



información; en particular, la información de que la nieve es blanca (o, en una versión más elaborada de la propuesta, la información de que 'la nieve es blanca' es un enunciado verdadero). En este caso, el segundo papel no es tan importante como el primero, porque el hecho de que Pedro y yo seamos miembros de la misma comunidad hace que me sea bastante fácil imaginar el tipo de propósitos con respecto a los cuales Pedro tiene acceso a la información relevante.

En el otro extremo están las atribuciones de creencias matemáticas. Aquí el primero de los dos papeles de  $p$  no es tan importante como el segundo, porque la región del espacio lógico que identificamos a partir de  $p$  no es particularmente interesante (generalmente será información acerca de cómo se usa nuestro vocabulario matemático). En cambio, el segundo papel de  $p$  es crucial. Necesitamos saber para qué propósitos tiene el sujeto acceso a información acerca de cómo se usa el vocabulario matemático.

\* \* \*

Antes de terminar, me gustaría que pensáramos en la propuesta desde un punto de vista un poco más abstracto.

De acuerdo con el modelo que hemos estado discutiendo, el sistema cognitivo de una persona contiene cierta información, pero el sistema no tiene acceso irrestricto a esa información. Tiene acceso a la información para algunos propósitos pero no para otros. Podemos pensar en un sistema cognitivo de este tipo como un sistema *fragmentado*. Hablando metafóricamente, podemos pensar en sistema como una *colección* de individuos, cada uno correspondiendo a un cierto tipo de propósito para el que podría requerirse un cierto trozo de información.

En lugar de decir que el sujeto tiene un trozo de información disponible para ciertos propósitos, podríamos decir, por tanto, que el 'fragmento' del sujeto que le corresponde a esos propósitos posee esa información (y que la posee sin más). Podríamos, pues, redescubrir la propuesta de la sesión de hoy como la idea de que los logros cognitivos en lógica y matemáticas consisten, en parte, de la adquisición de una habilidad para transmitir información entre diferentes fragmentos de nuestro sistema cognitivo.

La idea de un sistema cognitivo fragmentado nos permite modelar fenómenos que van más allá de los logros cognitivos en lógica y matemáticas. Consideremos un ejemplo. Pedro me contó que alguna vez una historia sobre un desayuno muy abundante: chilaquiles de pollo, huevos rancheros, una canasta de pan dulce, y un montón de cosas más. Lo curioso del asunto es que mientras desayunaba todo eso, Pedro le explicaba a un amigo que él siempre desayuna muy ligero. Podemos modelar la situación de Pedro de manera muy natural utilizando las herramientas que hemos desarrollado aquí. Podemos decir, en particular, que el sistema cognitivo de Pedro está fragmentado. El fragmento de Pedro que corresponde a la tarea de describir sus hábitos alimenticios posee la información (o, en este caso desinformación) de que

Pedro siempre desayuna muy ligero. Al mismo tiempo, el fragmento de Pedro que corresponde a la tarea de responder a una pregunta como '¿Qué estás desayunando?' posee la información de que en esta ocasión está disfrutando de un desayuno muy abundante.

Ejemplos como estos, en los que un sujeto tiene creencias inconsistentes, nos dan razones independientes para pensar en nuestros sistemas cognitivos como fragmentados. Un ejemplo diferente en el que es útil hablar de fragmentación es el caso de María y el tomate. Como vimos en sesiones anteriores, María ha estado encerrada en una habitación blanco y negro toda su vida, pero sabe todo lo que la ciencia del color tiene que enseñar. ¿Qué es lo que aprende María cuando ve un tomate rojo por primera vez? La fragmentación nos permite dar un cierto tipo de respuesta a esta pregunta, una respuesta que es valiosa porque es compatible con el fisicalismo.

Supongamos que el fisicalista acepta el siguiente enunciado de identidad:

**[Rojo]**

Que María experimente la sensación de ver rojo es simplemente que María esté en el estado cerebral R.

Supongamos además que a María le habían avisado de antemano que vería un tomate rojo al mediodía. Dado que María sabe todo lo que la ciencia del color tiene que enseñar, esto le permite inferir que al mediodía su cerebro estará en el estado cerebral R. Pero si [Rojo] es verdad, la región del espacio de posibilidad en la que María está en el estado cerebral R es exactamente la región en la que experimenta la sensación de ver rojo.

¿Cómo podríamos modelar lo que sucede cuando María ve el tomate? Lo que les quiero proponer aquí es que el sistema cognitivo de María está fragmentado. María posee toda la información relevante -- la información de que estará en el estado cerebral R al mediodía -- antes de ver el tomate, pero antes de ver el tomate la información estaba solo disponible para ciertos propósitos. Estaba disponible para el propósito de responder a preguntas como '¿en qué estado cerebral estarás al mediodía?', pero no estaba disponible para otros propósitos. Supongamos, por ejemplo, que antes de la entrega del tomate, le enseño a María una pelota roja y le pregunto '¿Así va a ser la sensación que vas a experimentar al medio día, cuando se te enseñe el tomate?'. María no estaría en posición de responder a mi pregunta, a pesar de que posee toda la información relevante. Sabemos que posee la información relevante porque sabemos que posee la información de que está experimentando ahora un cierto tipo de sensación (equivalentemente: la información de que está en el estado cerebral R), y sabemos que posee la información de que al mediodía estará en el estado cerebral R (equivalentemente: la información de que a las doce experimentará esa sensación particular). Juntos, estos dos pedazos de información implican que la sensación que experimenta María ahora es la misma que experimentará al mediodía.

El logro cognitivo que experimenta María cuando finalmente mira el tomate puede, por tanto, modelarse diciendo que María expande el rango de propósitos -- y por lo tanto el rango de fragmentos -- con respecto a los cuales tiene disponible información pertinente al estado cerebral R. Es solo una vez que María ha visto el tomate que es capaz de utilizar la información de que al mediodía estuvo en el estado cerebral R para responder a la pregunta '¿Es así la sensación que experimentaste al mediodía?' mientras le muestro una pelota roja.

## Sesión 6: Mundos posibles

*¿Cómo podríamos recuperar un espacio de mundos posibles a partir de nuestra concepción del espacio de posibilidad?*

---

Lo que me gustaría hacer en la sesión de hoy es articular una concepción de la posibilidad metafísica basada en las herramientas que hemos estado desarrollando en sesiones anteriores. Después me gustaría contrastar mi propuesta con la de David Lewis.<sup>58</sup>

Antes habíamos visto que hay dos maneras distintas de entender la palabra 'metafísica' cuando ocurre en la expresión 'posibilidad metafísica'. Una opción es entenderla como jugando un papel análogo al que juega 'física' en 'posibilidad física'; es decir, como especificando un cierto nivel de exigencia. La idea es que empezamos con un espacio de posibilidad más amplio que el de la posibilidad metafísica (el espacio de posibilidad *lógica*, por ejemplo). Luego estipulamos que, así como las posibilidades físicas son aquellos puntos en el espacio más amplio que no difieren del mundo actual en sus leyes físicas, las posibilidades metafísicas son aquellos puntos en el espacio más amplio que no difieren del mundo actual en sus 'leyes metafísicas'.

Aquí vamos a entender 'metafísica' en 'posibilidad metafísica' de modo bien distinto. Vamos a usar la palabra no para indicar un nivel de exigencia, sino para enfatizar el hecho de que estamos hablando de posibilidad *de mundo* en lugar de posibilidad *de representación*. ¿Cuál es la diferencia entre estos dos tipos de posibilidad? Las posibilidades *de mundo* son maneras en las que podría ser el mundo; las posibilidades *de representación*, en cambio, son descripciones del mundo que contamos como 'coherentes', dada una noción más o menos exigente de coherencia. Es natural pensar, por ejemplo, que no es *de mundo* posible que Héspero sea distinto de Fósforo. La razón es muy sencilla: un mundo en el que Héspero (es decir, Venus) es distinto de Fósforo (es decir, Venus), es un mundo en el que Venus es distinto de Venus, lo cual es absurdo. Uno podría decir, en cambio, que Héspero es Fósforo es posible *de representación*, argumentando que el enunciado 'Héspero es Fósforo' es lógicamente consistente.

Tal y como la vamos a entender aquí, la noción de posibilidad metafísica es simplemente la noción más amplia de posibilidad *de mundo* que hay. No es que el precio de ir más allá de la posibilidad metafísica sea violar una 'ley metafísica'. El precio de ir más allá de la posibilidad metafísica es terminar en el absurdo: uno deja de considerar maneras consistentes en las que el mundo podría ser.

---

<sup>58</sup> Lewis, David (1986a). *On the Plurality of Worlds*. Blackwell, Oxford and New York. Eduardo García Ramírez tiene una magnífica traducción al castellano de este texto que está publicada por el Instituto de Investigaciones Filosóficas.

\* \* \*

Imaginemos que  $L$  es un lenguaje de primer orden, y que  $L\Diamond$  es el resultado de agregarle a  $L$  el operador modal ' $\Diamond$ ', que se lee 'es metafísicamente posible que sea el caso que' (o, más amigablemente, 'podría ser el caso que'). Si en  $L$  podemos decir, por ejemplo, ' $\exists x$  (Elefante ( $x$ ))' ('existe algo tal que esa cosa es un elefante'), en  $L\Diamond$  podemos decir, por ejemplo, ' $\exists x \Diamond$  (Elefante ( $x$ ))' (es decir, 'existe algo tal que esa cosa podría ser un elefante'). El operador de necesidad metafísica, ' $\Box$ ', se puede definir en  $L\Diamond$  como ' $\sim\Diamond\sim$ ', y se lee 'es metafísicamente necesario que sea el caso que'.

El objetivo principal de la sesión de hoy es ofrecer un método general para responder a la pregunta siguiente: ¿Qué enunciados de  $L\Diamond$  hemos de aceptar?

Dicho de manera muy informal, la propuesta que me gustaría defender es la siguiente:

El operador ' $\Diamond$ ' ha de ser caracterizada tan liberalmente como sea posible, pero no tan liberalmente que terminemos en el absurdo.

Esto no es más que una reformulación de lo que ya había dicho antes: dado que la noción de posibilidad metafísica es la noción de posibilidad *de mundo* más amplia que hay, los límites de la posibilidad metafísica son simplemente los límites del absurdo.

El objetivo de hoy desarrollar esta idea con un poco más de detalle.

\* \* \*

Algunas asignaciones de propiedades a objetos constituirían un absurdo. (Por ejemplo: la asignación, a este vaso, de la propiedad de contener agua pero no  $H_2O$ .) Otras no. (Por ejemplo: la asignación, a este vaso, de la propiedad de contener agua.)

Una primera aproximación a la propuesta que me gustaría defender es que vamos a contar como posible a cualquier situación que pueda ser expresada en  $L\Diamond$ , siempre y cuando no involucre asignaciones de propiedades a objetos que constituyan un absurdo.

Necesitamos, por tanto, un método general para distinguir entre asignaciones de propiedades a objetos que constituyen un absurdo y asignaciones que no. Mi propuesta, a grandes rasgos, es que la *única* manera de que una asignación de propiedades a objetos constituya un absurdo es que viole alguna de las condiciones siguientes:

1. La asignación debe respetar *identidades* entre propiedades;

2. la asignación debe respetar las *esencias* de los objetos.<sup>59</sup>

Veamos qué quiere decir cada una de estas condiciones.

Primero: ¿qué quiere decir que se respete una identidad entre propiedades? Consideremos un ejemplo. Asumamos que estar compuesto de agua es simplemente estar compuesto de H<sub>2</sub>O, y, por tanto, que la propiedad de estar compuesto de agua es la misma que la propiedad de estar compuesto de H<sub>2</sub>O. Lo que nos dice nuestra primera condición es que una asignación de propiedades a objetos constituiría un absurdo si dijera que un objeto tiene la propiedad de contener agua pero no la propiedad de contener H<sub>2</sub>O.

Segundo: ¿qué quiere decir que no se respete la esencia de algún objeto? Yo no tengo opiniones muy firmes acerca de cuáles son las propiedades esenciales de los objetos, pero asumamos por el momento que ser elefante es esencial a los elefantes. En otras palabras: si algo es un elefante, parte de lo que es ser esa cosa es ser elefante. Supongamos ahora que estamos considerando admitir dos asignaciones de propiedades a Heriberto. De acuerdo con una de las asignaciones, Heriberto tiene la propiedad de ser elefante; de acuerdo con la otra, no la tiene. Lo que nos dice nuestra segunda condición es que no podemos aceptar ambas asignaciones sin caer en el absurdo. (Para ver por qué, comencemos con la asignación de acuerdo con la cual Heriberto tiene la propiedad de ser un elefante. Si esta no es una asignación absurda, sería metafísicamente posible que Heriberto fuera un elefante. Dado que ser elefante es esencial a los elefantes, esto nos lleva a la conclusión de que parte de lo que es ser Heriberto es ser elefante, lo cual es inconsistente con que haya asignaciones de acuerdo con las cuales Heriberto no sea un elefante. Así que la segunda asignación tiene que ser absurda.)

La propuesta que estoy defendiendo aquí tiene una parte menos controversial y una parte más controversial. La parte menos controversial es que violar alguna de las condiciones anteriores es *suficiente* para que una asignación de propiedades a objetos caiga en el absurdo. La parte más controversial -- y el corazón de la propuesta -- es que la *única* manera de que una asignación de propiedades a objetos caiga en el absurdo es que viole alguna de las condiciones anteriores.

**Pregunta:** Hay algo que no entiendo. Para ser abuelo es esencial tener hijos, ¿se sigue de eso que es metafísicamente imposible que mi abuelo hubiera existido sin hijos?

**Respuesta:** No. Hay que distinguir entre los componentes esenciales de una propiedad y las propiedades esenciales de un objeto. La propiedad de tener hijos es un componente esencial de la *propiedad* de ser abuelo, pero esto no

---

<sup>59</sup> Kit Fine desarrolla una versión de esta idea en Fine, Kit (1994). *Essence and modality. Philosophical Perspectives*, 8, 1–16.

implica que los objetos que tienen la propiedad de ser abuelo, o la propiedad de tener hijos, tengan estas propiedades de manera esencial.

Aquí es útil introducir un poco de terminología. Digamos que una propiedad *constitutiva* es una propiedad tal que si es posible que algo la tenga entonces es imposible que esa cosa exista sin tenerla. Lo que revela tu pregunta es que aunque es razonable pensar que ni la propiedad de ser abuelo ni la propiedad de tener hijos es una propiedad constitutiva, todo esto es compatible con que la propiedad de tener hijos sea una parte esencial de la propiedad de ser abuelo.

**Pregunta:** ¿Crees que hay propiedades constitutivas?

**Respuesta:** No es una tesis con la que estoy comprometido, pero es importante para mí que la existencia de propiedades constitutivas compatible con el programa que les estoy proponiendo.

En la sesión de hoy no voy a tratar de defender enunciados modales particulares. Lo que voy a defender es un sistema para decidir qué enunciados modales aceptar.

\* \* \*

Tenemos sobre la mesa una versión preliminar -- y muy informal -- de la propuesta. La idea es que vamos a contar como posible a cualquier situación que pueda ser expresada en  $L\Diamond$ , siempre y cuando no involucre una asignación de propiedades que viole las condiciones 1 o 2.

Una característica interesante de las condiciones 1 y 2 es que pueden reformularse utilizando enunciados de identidad. Específicamente, uno podría decir que una situación que puede ser expresada en  $L\Diamond$  solo es absurda si viola alguna de las condiciones siguientes:

1'. Es consistente con todos los enunciados de identidad verdaderos

(Aquí estoy pensando en enunciados de identidad simples, en los que no hay iteraciones del operador de identidad)

2'. Es consistente con todos los enunciados de identidad *condicionales* verdaderos

(Un enunciado de identidad condicional es un enunciado de la forma 'Supongamos que algo es P, entonces parte de lo que es ser ese individuo es ser P', en donde P es un predicado de  $L$ .)

Esto nos pone en posición de articular una versión ligeramente menos informal de la propuesta. Vamos a contar como posible a cualquier situación que pueda expresarse en  $L_{\Diamond}$ , siempre y cuando sea compatible con los enunciados de identidad 'básicos', donde los enunciados básicos son los que figuran en las condiciones 1' y 2' (es decir, son enunciados de identidad condicionales y no condicionales que hayan sido contruidos a partir de vocabulario en  $L$  y en los que no hay iteraciones del operador de identidad).

Si pudiéramos desarrollar una versión más formal de esta propuesta, tendríamos una conclusión muy emocionante. Habríamos llegado a la conclusión de que fijar las condiciones de verdad de enunciados de identidad básicos es suficiente para fijar los valores de verdad de todos los enunciado en  $L_{\Diamond}$ .

\* \* \*

Este tipo de propuesta requiere una hipótesis altamente no trivial. Podemos verlo notando que cada enunciado de identidad tiene una contraparte modal. En particular, la contraparte del enunciado de identidad

$$P(x) \equiv_x Q(x)$$

es el enunciado modal

$$\Box \forall x (P(x) \leftrightarrow Q(x))$$

y la contraparte del enunciado de identidad condicional

$$\frac{P(z)}{\text{-----}} \\ z = x \equiv_x P(x)$$

(leer: supongamos que algo es  $P$ ; entonces parte de lo que es ser ese individuo es ser  $P$ )

es el enunciado modal

$$\Box (\forall z (P(z) \rightarrow \Box \forall x (z = x \leftrightarrow P(x)))$$

Tal y como yo los entiendo aquí, un enunciado de identidad y su contraparte modal tienen las mismas condiciones de verdad. La propuesta presupone, por tanto, que uno puede usar un subconjunto bien pequeño de los enunciados de  $L_{\Diamond}$  -- aquellos que son contrapartes de algún enunciados de identidad básico -- para fijar los valores de verdad de *todos* los enunciados en  $L_{\Diamond}$ .



Consideremos, por ejemplo, el enunciado ' $\Diamond(\text{Bigote}(\text{Pedro}))$ ' ('es posible que Pedro tenga bigote'). Ni este enunciado ni su negación es consecuencia lógica de enunciados de  $L\Diamond$  que son contrapartes de los enunciados de identidad básicos. Una versión exitosa de nuestra propuesta, sin embargo, arrojaría la conclusión de que, a pesar de ello, fijar los valores de verdad de los enunciados de identidad básicos es suficiente para fijar el valor de verdad de ' $\Diamond(\text{Bigote}(\text{Pedro}))$ '.

**Pregunta:** ¿No podríamos decir simplemente que un enunciado de  $L\Diamond$  cuenta como verdadero siempre y cuando sea consistente con todos los enunciados básicos verdaderos?

**Respuesta:** No es tan sencillo, desafortunadamente. El problema es que puede suceder que tanto un enunciado como su negación sean consistentes con todos los enunciados básicos verdaderos, y no queremos concluir de ello que hay enunciados verdaderos con negaciones verdaderas.

Aquí hay una propuesta ligeramente diferente que captura parte de lo que tienes en mente. Uno podría fijarse en un conjunto inicial de 'mundos' y eliminar aquellos que no satisfagan la contraparte modal de algún enunciado básico verdadero. Esto nos deja con conjunto de mundos 'buenos', y uno podría decir un enunciado de  $L\Diamond$  es verdadero a secas si y solo si es verdadero en todos los modelos 'buenos'.

Desafortunadamente, no es claro qué mundos hemos de contar como 'buenos'. Para ver por qué, consideremos una propiedad constitutiva: la propiedad de ser elefante, por ejemplo. Asumamos, pues, que el enunciado de identidad condicional

$$\begin{array}{l} \text{Elefante}(z) \\ \text{-----} \\ z = x \equiv_x \text{Elefante}(x) \end{array}$$

es verdadero, y, por tanto, que también es verdadera su contraparte modal:

$$\Box(\forall z (\text{Elefante}(z) \rightarrow \Box \forall x (z = x \leftrightarrow \text{Elefante}(x)))$$

Ahora imaginemos que entre nuestros 'mundos' iniciales hay uno de acuerdo con el cual Heriberto es un elefante y uno de acuerdo con el cual Heriberto no es un elefante. Para respetar la verdad del enunciado modal que captura la constitutividad de ser elefante, tenemos que eliminar uno de estos mundos. ¿Pero cuál? Las contrapartes modales de los enunciados básicos verdaderos no nos dan una respuesta inmediata a esta pregunta.

La lección de todo esto es que la hipótesis de que es posible usar una asignación de valores de verdad de los enunciados básicos para fijar el valor de verdad de todo enunciado de  $L_{\Diamond}$  es altamente no trivial.

**Pregunta:** Cuando hablas de 'mundos' en este contexto, ¿qué quieres decir?

**Respuesta:** Estoy usando 'mundo' en el sentido de la teoría de modelos para lenguajes modales que desarrolló Kripke. Un mundo es simplemente un objeto matemático. Es una pareja que consiste de un dominio, y de una asignación de referencias al vocabulario del lenguaje.

Hay, sin embargo, un detalle que vale la pena mencionar. La manera usual de hacer semántica de mundos posibles para lenguajes modales presupone que podemos hablar, en el metalenguaje, de objetos meramente posibles. Esto es potencialmente problemático porque no es claro que existan objetos meramente posibles. En el último número de *Crítica* tengo un artículo en el que propongo un truco para evitar este problema.<sup>60</sup> Desarrollo una noción de mundo que nos permite hacer semántica kripkeana sin apelar a objetos meramente posibles. Estrictamente hablando, es ese el sentido de 'mundo' que tengo en mente aquí.

\* \* \*

Hemos visto que nuestra propuesta requiere una hipótesis no trivial. Felizmente, hay un teorema que muestra que la hipótesis es verdadera.<sup>61</sup> Lo llamo el Teorema de Extensión. Informalmente, el teorema dice lo siguiente:

### **El Teorema de Extensión**

Cualquier asignación de valores de verdad a los enunciados de identidad básicos puede extenderse a una asignación de valores de verdad para todos los enunciados  $L_{\Diamond}$ .

Más específicamente, lo que muestra el teorema es que los enunciados de identidad básicos pueden utilizarse para asignarles valores de verdad a todos los enunciados de  $L_{\Diamond}$  de modo tal que se cumplen dos condiciones: (1) todas las contrapartes modales de enunciados básicos verdaderos cuentan como verdaderos, y (2) los enunciados de  $L_{\Diamond}$  que expresan 'posibilidades' (en un sentido técnico de 'posibilidad') cuentan como verdaderos siempre que sean compatibles con los enunciados básicos verdaderos.

---

<sup>60</sup> 'An Actualist's Guide to Quantifying-In', *Crítica: Revista Hispanoamericana de Filosofía* 44, 2012, pp. 3-34.

<sup>61</sup> Discuto el teorema en el Capítulo 5 de *The Construction of Logical Space*, y lo pruebo en el Apéndice C.

\* \* \*

¿Qué conclusiones epistemológicas podemos extraer de la propuesta?

En sesiones anteriores sugerí que la manera de decidir qué enunciados de identidad aceptar es haciendo ciencia (y en algunos casos haciendo filosofía, pero no necesariamente metafísica modal). En la sesión de hoy hemos visto que los valores de verdad de los enunciados de identidad básicos pueden utilizarse para fijar los valores de verdad de todos los enunciados de  $L\Diamond$ . Esto sugiere que la manera de decidir qué enunciados de  $L\Diamond$  aceptar es también haciendo ciencia (y en algunos casos filosofía, pero no necesariamente metafísica modal).

Esta es una conclusión muy emocionante, porque sugiere que hay una ruta no misteriosa para dar cuenta de nuestro conocimiento modal.

**Pregunta:** En alguna versión anterior de tu trabajo sugerías que uno no debe pensar en el operador de identidad como ‘más fundamental’ que el operador de posibilidad metafísica. ¿Ha cambiado tu posición?

**Respuesta:** Mi posición sigue siendo la misma. No me parece que ninguna de estas dos nociones sea ‘más fundamental’ que la otra. Lo que sí me parece primario, desde un punto de vista epistemológico, es el quehacer científico. Me parece esencial que seamos capaces de vincular nuestro quehacer científico con el operador de identidad y con el operador de posibilidad (aunque no sea importante cuál de esos vínculos ocurra primero). La razón por la que el vínculo es tan importante es por que es eso lo que hace que las nociones no sean misteriosas, y que no estén totalmente fuera de nuestro alcance epistémico.

\* \* \*

Lo que me gustaría hacer ahora es contrastar mi propuesta con la de David Lewis. La concepción de mundos posibles de Lewis puede dividirse en dos componentes centrales. Muchos de ustedes están familiarizados con el primero de ellos:

L1

Que sea posible que  $P$  es simplemente que exista un mundo en el que es el caso que  $P$

Una manera de pensar en este principio es como proponiendo una *reificación* de la posibilidad. Dado que es posible que Sebas tenga bigote, el principio implica que existe un mundo Lewisiano en el que es el caso que Sebas tiene bigote.

La propuesta de Lewis tiene, sin embargo, un segundo componente, y es posible que a algunos de ustedes les resulte sorprendente:

L2

Que sea el caso que *P* en un mundo es simplemente que el mundo *represente* que *P* (con respecto a una cierta relación de contraparte).

Aquí está lo que dice Lewis al respecto en *Sobre la Pluralidad de los Mundos*:

¿Cómo es que un [un mundo Lewisiano] representa de Humphrey que existe. Un [mundo Lewisiano] puede hacerlo teniendo a Humphrey mismo como parte. Es así como nuestro mundo representa, de Humphrey, que existe. Pero para que otros mundos representaran de la misma manera que Humphrey existe, Humphrey tendría que ser una parte común de muchos mundos superpuestos... rechazo tal superposición... Hay una mejor manera de que un [mundo Lewisiano] represente, de Humphrey, que existe... puede tener un Humphrey propio, una contraparte de carne y hueso de nuestro Humphrey, un hombre muy parecido a Humphrey en sus orígenes, sus características intrínsecas, o su papel histórico. (p. 194)

La idea es la siguiente. Hay mundos Lewisianos en los que Humphrey mismo no existe pero en los que hay otros individuos que tienen cierto parecido al verdadero Humphrey, al Humphrey de nuestro mundo. Si el parecido es del tipo correcto, diremos que el individuo relevante es una *contraparte* de Humphrey. De acuerdo con la propuesta de Lewis, lo que tiene que ocurrir para que un mundo represente la proposición de que Humphrey existe es que el mundo contenga una contraparte de Humphrey.

Lewis también cree que la relación de contraparte puede variar con el contexto. En un contexto en el que el parecido físico es importante, podría suceder que solo individuos con gran parecido físico a Humphrey cuenten como sus contrapartes; en un contexto en el que son más importantes los orígenes, podría suceder que solo individuos con orígenes similares a los de Humphrey cuenten como sus contrapartes.

Cuando uno discute la posición de Lewis, es importante distinguir entre mundos y posibilidades. Aunque los mundos Lewisianos *representan* posibilidades, los mundos Lewisianos no son posibilidades. Una manera de convencerse de este punto es notando que un solo mundo Lewisiano puede representar dos posibilidades diferentes. Supongan, por ejemplo, que Helena y Julia son gemelas idénticas. Hubiera podido ocurrir que naciera una sin la otra. Si Helena y Julia son suficientemente parecidas, el mismo mundo Lewisiano representa la posibilidad de que Helena existe sin hermanas (con respecto a la relación de contraparte que es relevante en un contexto) y la posibilidad de que Julia existe sin hermanas (con respecto a la relación de contraparte que es relevante en otro contexto).

**Pregunta:** Utilizas los principios L1 y L2 para caracterizar la posición de Lewis, pero ¿no habría otros filósofos que también aceptarían L1 y L2?

**Respuesta:** Absolutamente. Lewis no es el único que cree que hay mundos, o que la relación entre mundos y posibilidades es una relación de representación. Una posición diferente a la de Lewis que también acepta L1 y L2 es el representacionismo lingüístico, de acuerdo con el cual los mundos son conjuntos de enunciados y la posibilidad representada por un mundo es la proposición expresada por los enunciados que el mundo tiene como elementos.

Lo que es específico a la posición de Lewis son los detalles de la relación de representación. Lewis, a diferencia de otros representacionistas, cree que la manera en la que un mundo representa la posibilidad de que yo tengo bigote es conteniendo a un señor de carne y hueso que se parece a mí de ciertas maneras, y que tiene bigote.

**Pregunta:** ¿Es Lewis un *haecceitista* acerca de posibilidades? ¿Cree que hay posibilidades que conciernen individuos particulares?

**Respuesta:** Tal y como defines 'haecceitismo', Lewis sí que es un haecceitista. Cree, por ejemplo, que hay posibilidades que me conciernen a mí: aquellas que son representadas por mundos Lewisianos que contienen contrapartes mías.

Hay, sin embargo, un sentido diferente en el que Lewis no es un haecceitista, y ese es el sentido de 'haecceitismo' que él enfatiza en sus textos. Aunque Lewis es un haecceitista con respecto a *posibilidades*, no es un haecceitista con respecto a los *mundos Lewisianos* que representan a estas posibilidades. Lewis cree que hay diferentes posibilidades que me atañen a mí, pero no cree que hay diferentes mundos Lewisianos que me *contienen* a mí.

El haecceitismo posibilístico -- la idea de que hay posibilidades diferentes que atañen al mismo individuo -- no debería ser una tesis controversial. Uno podría, por ejemplo, escribir una obra de teatro sobre Sebas, de acuerdo con la cual Sebas tiene bigote. Esa obra de teatro tiene un cierto contenido y el contenido atañe a Sebas. Mientras la obra no caiga en el absurdo, ese contenido corresponde a una posibilidad que atañe a Sebas, una posibilidad de acuerdo con la cual *él mismo* tiene bigote. Al mismo tiempo, podría haber una representación de ese contenido -- por ejemplo, una serie de actores diciendo ciertas cosas en un cierto contexto -- en la que Sebas no esté involucrado. (Sebas no tiene por qué ser actor de teatro.) Lo que tendríamos entonces es una representación en la que no figura Sebas, pero en la que se representa una posibilidad en la que Sebas sí figura. La posición de Lewis es parecida.

\* \* \*

Muchas de las tesis Lewisianas que hemos discutido hasta este momento son compatibles con mi propuesta. Esto es cierto, en particular, de las tesis L1 y L2.

L1 es un enunciado de identidad: que algo sea posible es simplemente que haya un mundo en el que eso sea verdadero. Aunque L1 es compatible con mi propuesta, me gustaría enfatizar es que no estoy *comprometido* con esa idea. (Es cierto que la prueba los resultados técnicos de los que hablé antes habla de 'mundos', pero ahí los mundos son simplemente parte de la maquinaria técnica que usamos para asignarles valores de verdad a los enunciados de  $L\Diamond$ . No juegan un papel esencial en la concepción de posibilidad metafísica que estoy desarrollando.)

¿Hemos de aceptar L1? Como en el caso de todo enunciado de identidad, creo que aceptar L1 tiene ventajas e inconvenientes. La ventaja es que hay ciertas preguntas incómodas que no tendríamos que contestar; por ejemplo: 'Estoy de acuerdo en que es posible que Sebas tiene bigote, pero ¿tenemos razones para pensar que hay un mundo posible en el que tiene bigote?'. La desventaja es que uno pierde distinciones con las cuales trabajar; en particular, uno pierde la distinción entre que algo sea posible y que exista un mundo de acuerdo con el cual eso es verdadero. En este caso particular yo diría que las ventajas son más que suficientes para compensar las desventajas, y, por tanto, que tenemos buenas razones para aceptar L1. Pero, como dije antes, esto no es algo a lo que estoy comprometido aquí.

L2 es una tesis representacionalista. Nos dice es que no hay diferencia entre que sea cierto que P en un mundo y que ese mundo represente que P. Como en el caso de L1, no me parece que L2 sea esencial a mi propuesta. Si uno quisiera, sin embargo, uno podría defender una tesis representacionalista con respecto al conjunto de mundos posibles que arroja la prueba del Teorema de Extensión. Uno podría decir que un mundo posible (en el sentido de la teoría de modelos) representa que P si y solo si P es verdadero en ese mundo posible de acuerdo con la semántica de Kripke.

Ese punto no me parece particularmente interesante. El punto que sí me parece interesante es que resulta ser que uno puede probar un teorema muy sorprendente.<sup>62</sup> Uno puede demostrar que hay un sentido preciso en el que -- cuando restringimos nuestra atención al proyecto de hacer semántica -- todo lo que puede hacerse utilizando mundos Lewisianos también puede hacerse utilizando los mundos posibles con los que estamos trabajando aquí.

Me gustaría hacer un pequeño paréntesis para explicar este punto con un poco más de detalle. Una de las razones por las que es útil la metafísica de Lewis es que nos permite hacer semántica intencional, es decir, semántica de mundos posibles. (Más específicamente, Lewis muestra que es posible asignarle a cada una de las unidades

---

<sup>62</sup> Ver 'An Actualist's Guide to Quantifying-In', *Crítica: Revista Hispanoamericana de Filosofía* 44, 2012, pp. 3-34. También discuto el teorema en el Capítulo 6 de *The Construction of Logical Space*, y lo pruebo en el Apéndice B.

léxicas del lenguaje un objeto conjuntista construido a partir de habitantes de mundos Lewisianos, y que es posible hacerlo de tal modo que exista un método composicional para generar condiciones de verdad para cada enunciado del lenguaje a partir de esas asignaciones.<sup>63)</sup>

Una de las cosas que se sigue del teorema que tengo en mente es que el proyecto semántico de Lewis se puede llevar a cabo utilizando solo 'mundos' en el sentido de la teoría de modelos. No solo eso. Como antes había mencionado, hay una manera de hablar de 'mundos' (en el sentido de la teoría de modelos) que no presupone que tengamos la capacidad de cuantificar sobre objetos meramente posibles. Se sigue del teorema, por tanto, que el proyecto semántico de Lewis puede llevarse a cabo utilizando solo objetos conjuntistas contruidos a partir de individuos que existen en el mundo actual.

Es importante tener en cuenta, sin embargo, que el teorema está restringido al proyecto semántico. Cuando vamos más allá del proyecto semántico, hay cosas que se pueden hacer usar utilizando la maquinaria de Lewis pero no se pueden hacer con la mía. Aquí hay un ejemplo metafísico. Lewis propone caracterizar a las propiedades como conjuntos de habitantes de mundos Lewisianos. La propiedad de ser filósofo, por ejemplo, tiene como miembros a todos los filósofos del mundo Lewisiano actual, y a todos los filósofos en mundos Lewisianos no actuales. Esa noción de propiedad tiene la característica de transcender el poder expresivo de nuestro lenguaje. Permite la existencia de propiedades que no corresponden a ningún predicado que seamos capaces de expresar. Eso es algo que mi propuesta no nos permite hacer. Las construcciones con las que yo estoy trabajando están limitadas por el poder expresivo de nuestro lenguaje. (La razón por la que el teorema funciona cuando restringimos nuestra atención al proyecto semántico es que ahí las únicas propiedades que necesitamos son propiedades que pueden ser expresadas en el lenguaje con el que estamos trabajando.)

\* \* \*

Lo que es realmente característico de la posición de Lewis, y lo que a mí me parece problemático, no es L1. Tampoco es L2, considerado a secas. Lo que a mí me parece problemático es la manera particular en la que Lewis desarrolla L2, la manera particular en la que cree que los mundos representan posibilidades.

El problema emerge porque Lewis cree que los mundos son distribuciones espacio-temporales de propiedades 'perfectamente naturales' a través del espacio y el tiempo. (La noción de una propiedad perfectamente natural es un concepto teórico de Lewis; no voy a entrar aquí en los detalles, pero la idea es que las propiedades perfectamente naturales son siempre propiedades intrínsecas.<sup>64)</sup>

---

<sup>63</sup> Lewis, David (1970). 'General semantics'. *Synthese*, **22**, 18–67.

<sup>64</sup> Lewis, David (1983a). New work for a theory of universals. *Australasian Journal of Philosophy*, **61**, 343–377.

Una consecuencia de este modo de pensar es que hay un cierto sentido en el que Lewis está comprometido con una reducción de lo modal a lo no modal. Consideremos, por ejemplo, el hecho modal de que Humphrey hubiera podido ganar la elección presidencial estadounidense de 1968. Lewis cree que este hecho puede descomponerse en condiciones que afectan a diferentes mundos. Tal vez cree que se puede descomponer en (1) la condición de que nuestro mundo contenga a Humphrey, y (2) la condición de que haya algún mundo en el que una contraparte de nuestro Humphrey gana una cierta elección. Pero Lewis también cree que todo lo que sucede en un mundo superviene en su distribución espacio-temporal de propiedades perfectamente naturales, así que cree que, al final del día, el hecho de que Humphrey hubiera podido ganar la elección se reduce a la presencia de ciertas distribuciones de propiedades perfectamente naturales en el espacio de los mundos Lewisianos. Si suponemos que las distribuciones espacio-temporales de propiedades perfectamente naturales son hechos no modales, esto implica que el hecho modal de que Humphrey hubiera podido ganar la elección puede descomponerse en hechos no modales.

Mi propuesta, en cambio, no está comprometida con una reducción de lo modal a lo no modal, por lo menos no de manera inmediata. Todo depende de los enunciados de identidad que uno acepte. Para que la propuesta resultara en una reducción como la de Lewis, uno tendría que aceptar enunciados de identidad como el siguiente:

Que Humphrey gane la elección presidencial de 1968 es simplemente que tal-y-tal sea el caso.

en donde el 'tal-y-tal' se reemplaza por un enunciado enteramente no modal.

No es obvio, sin embargo, que sea posible encontrar un 'tal-y-tal' apropiado, porque es difícil encontrar palabras en nuestro lenguaje que no tengan algún componente modal. Es natural pensar, por ejemplo, que parte de lo que es ser Humphrey es ser de la especie *Homo sapiens*, y que parte de ser una elección es ser el producto de una sociedad. Es natural pensar, incluso, que los conceptos de la física fundamental tienen componentes modales: parte de lo que es tener carga negativa es tener la disposición de repeler a otros objetos con carga negativa; parte de lo que es tener masa es tener la disposición de resistir la aceleración.

Tal vez Lewis tiene razón. Tal vez hay una reducción de lo modal a lo no modal. Lo que me parece mal es presuponer esta idea desde el principio. Un aspecto fundamental de mi proyecto es que la decisión de qué enunciados de identidad aceptar está estrechamente vinculada con nuestro quehacer científico. La pregunta de si aceptar enunciados de identidad que nos comprometan a una reducción de lo modal a lo no modal es, por tanto, una pregunta que no puede desvincularse de nuestro quehacer científico.

No es claro, en cambio, que las propiedades perfectamente naturales de Lewis estén vinculadas a nuestro quehacer científico. Lewis algunas veces dice cosas que sugieren que las propiedades fundamentales de la física serían ejemplos de propiedades



perfectamente naturales. Tal vez querría insistir que la propiedad de tener carga negativa es una propiedad perfectamente natural. Pero, ¿qué es lo que de hecho nos dice la ciencia acerca de la carga negativa? Todo lo que nos dice la ciencia tiene que ver con las características disposicionales de objetos con carga negativas. La física nos dice que las partículas con cargas similares tienen la disposición de repelerse, y que las partículas con cargas diferentes tienen la disposición de atraerse. Lewis cree, en cambio, que las propiedades perfectamente naturales son totalmente intrínsecas: el que uno tenga esa propiedad no depende de nada de lo que uno tiene a su alrededor. Eso significa que hay mundos Lewisianos en los que objetos con carga negativa se atraen y mundos Lewisianos en los que se repelen. ¿En cuál de esos mundos vivimos? No podemos usar las herramientas científicas usuales para contestar esa pregunta porque el único método que nos da la ciencia para determinar si dos partículas tienen la misma carga consiste en observar sus propiedades disposicionales. Lewis está, por tanto, comprometido con la idea de que la identidad de las propiedades perfectamente naturales trascienden a nuestras mejores teorías científicas.<sup>65</sup>

Eso no muestra que está mal la teoría de Lewis, pero sí muestra que Lewis está comprometido con un cierto tipo de anti-naturalismo. Hay una falta de continuidad entre nuestro teorizar científico y lo que está a la base del edificio modal de Lewis. La posición resultante tiene ventajas e inconvenientes. La ventaja es que uno puede aceptar una posición fundacionista, de acuerdo con la cual todo se reduce, al final del día, a distribuciones espacio-temporales de propiedades perfectamente naturales. La desventaja es que uno está obligado a aceptar una concepción del espacio de posibilidad basada en consideraciones metafísicas que no están muy claramente vinculadas a nuestras teorías científicas, y a mí ese vínculo me parece muy importante. Creo que para que la posibilidad metafísica no resulte un concepto misterioso, es indispensable que nuestra concepción del espacio de posibilidad esté determinada, entre otras cosas, por nuestras mejores teorías científicas.

De acuerdo con mi posición -- y aquí estoy influenciado por Stalnaker<sup>66</sup> -- construimos el espacio de posibilidad a partir del mundo con el que nos encontramos. Es en el proceso de teorizar acerca del mundo en el que vivimos que construimos nuestra concepción del espacio de posibilidad.

**Pregunta:** En vista de lo que dices acerca de las nociones de la física fundamental, ¿no estaríamos comprometido con la idea de que las leyes físicas son metafísicamente necesarias?

**Repuesta:** Efectivamente. Me parece razonable pensar que parte de lo que es tener carga negativa es tener la disposición de rechazar otras partículas con carga negativa. Si esto es cierto, mi propuesta implica que es metafísicamente

---

<sup>65</sup> Este es un punto que Lewis mismo admite. Véase su artículo 'Ramseyan Humility', en Braddon-Mitchell, David and Nola, Robert (eds.) (2009). *Conceptual Analysis and Philosophical Naturalism*. MIT Press, Cambridge, MA.

<sup>66</sup> Ver Ver Stalnaker, Robert C. (2008) *Our Knowledge of the Internal World*, Oxford University Press: Oxford. También: Stalnaker, Robert C. (2012) *Mere Possibilities*, Princeton University Press: Princeton, N.J.

imposible que dos partículas con carga negativa no tengan la disposición de rechazarse.

Vale la pena enfatizar, sin embargo, que esta conclusión depende de que aceptemos enunciados de identidad particulares. No es algo con lo que uno estaría comprometido simplemente por aceptar la concepción de posibilidad metafísica que he estado defendiendo aquí.

También vale la pena enfatizar que aún cuando uno acepte enunciados de identidad del tipo que llevaría a la conclusión de que las leyes físicas son metafísicamente necesarias, uno no habrá excluido posibilidades de acuerdo con las cuales hay objetos que se parecen mucho a los electrones y se atraen en lugar de repelerse. Es solo que estos objetos no son electrones, y no tienen carga negativa.

\* \* \*

Hemos visto que la concepción de posibilidad metafísica que he estado defendiendo aquí no arroja una reducción de lo modal a lo no modal, por lo menos no de manera inmediata. Hay, sin embargo, un resultado que la propuesta sí arroja, y que a mí me resulta muy sorprendente.<sup>67</sup> Lo llamo el 'Teorema de Reducción'. Es un corolario del Teorema de Extensión, y dice lo siguiente.

### **Teorema de Reducción**

Dado cualquier enunciado  $\phi$  de  $L\Diamond$  existe un enunciado  $\phi'$  construido exclusivamente a partir de vocabulario de  $L$  (y, por tanto, sin operadores modales) tal que  $\phi$  y  $\phi'$  tienen las mismas condiciones de verdad. (Es importante tener en cuenta, sin embargo, que  $\phi'$  podría tener longitud infinita.)

Consideremos el siguiente enunciado:

$\exists x (\text{Mamífero}(x) \ \& \ \Diamond \text{Humano}(x))$   
(Hay algo que es un mamífero y podría ser humano)

¿Cómo podríamos especificar las condiciones de verdad de este enunciado sin utilizar operadores modales? Una opción sería utilizar el razonamiento siguiente:

¿Cuáles son las cosas que hubieran podido ser humanas? Un no-humano no hubiera podido ser humano, por lo menos si la propiedad de ser no-humano es constitutiva:

No-Humano( $z$ )  
-----  
 $z = x \equiv_x \text{No-Humano}(x)$

---

<sup>67</sup> Discuto el teorema en el Capítulo 5 de *The Construction of Logical Space*, y lo pruebo en el Apéndice C.

Esto significa que los individuos que hubieran podido ser humanos son simplemente los individuos que son humanos. Por tanto, que el enunciado original tiene las mismas condiciones de verdad que el enunciado siguiente:

$$\exists x (\text{Mamífero}(x) \ \& \ \text{Humano}(x))$$

(Hay algo que es un mamífero y humano)

De hecho, alguien que cree que parte de lo que es ser humano es ser mamífero cree que el enunciado que acabamos de considerar tiene las mismas condiciones de verdad que este:

$$\exists x (\text{Humano}(x))$$

(Hay algo que es humano)

El enunciado ' $\exists x (\text{Humano}(x))$ ' nos da, por tanto, una manera de especificar las condiciones de verdad de ' $\exists x (\text{Mamífero}(x) \ \& \ \Diamond \text{Humano}(x))$ ' sin utilizar operadores modales.

Lo que muestra el teorema es que este punto puede generalizarse a todo enunciado de  $L\Diamond$ . ¿Significa esto que tenemos una reducción de lo modal a lo no modal? Es tentador pensar sí. El teorema nos dice, después de todo, que existe una manera de expresar las condiciones de verdad de cualquier enunciado de  $L\Diamond$  sin utilizar operadores modales. Me parece, sin embargo, que la respuesta correcta es 'no'. El problema es que para que un enunciado cuente como no modal, no es suficiente que prescinda de operadores modales. Uno podría pensar, por ejemplo que el predicado ' $\text{Humano}(x)$ ' expresa una noción modal, porque uno podría pensar que si alguien es humano, entonces parte de lo que es ser ese individuo es ser humano.

\* \* \*

Me gustaría concluir diciéndoles brevemente qué es lo que es lo que he tratado de hacer a lo largo de estas seis sesiones.

Mi objetivo principal ha sido desarrollar una cierta herramienta teórica: los enunciados de identidad. Traté de mostrar que esa herramienta es bien útil. Argumenté que se puede utilizar para desarrollar una buena filosofía de las matemáticas, y para iluminar los logros cognitivos de María.

Lo que hicimos en la sesión de hoy fue vincular nuestra herramienta con una idea fundamental en la filosofía analítica contemporánea: la noción de posibilidad metafísica. Una manera de pensar en estas conferencias, por tanto, es como un esfuerzo por clarificar la noción de posibilidad metafísica, y mostrar lo útil que puede ser para atacar ciertos problemas filosóficos.