

CONTRAFÁCTICOS Y GRADOS DE SIMILARIDAD*

José A. Díez Calzada

«'If kangaroos had no tails, they would topple over' seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over».

(D. Lewis, *Counterfactuals*.)

Abstract

The object of this work is to offer a modification of the possible worlds semantics for counterfactuals, specifically of the notion of truth value of a counterfactual in a world. I shall use for reference Lewis' semantics, for the other main one, that of Stalnaker, can be considered a special case of Lewis' (in the sense of Lewis (1973b, p. 78)) and, although they both have different foundations and consequences, they both agree in considering enough to relativize the truth conditions of a counterfactual to a similarity relation, just the issue here concerned. The modification proposed consists in introducing an additional parameter in these truth conditions. First we briefly sketch Lewis' semantics; then we present the emendation by means of an example; finally, we discuss the exact nature of the proposal.

1. Los condicionales contrafácticos, expresiones del tipo «si ocurriese entonces ocurriría β », y que simbolizamos mediante $\alpha \square \rightarrow \beta^1$, son esencialmente contexto-dependientes, sus condiciones de verdad dependen de ciertos parámetros presentes en el contexto de uso. Una semántica para contrafácticos debe establecer cuáles son esos parámetros y el modo exacto en que de ellos dependen las condiciones de verdad de los contrafácticos, y, haciendo esto, explica su comportamiento semántico.

* El presente trabajo es una ampliación y revisión sustanciales de una contribución presentada al «VI Congreso de Lenguajes Naturales y Lenguajes Formales» (Tarragona, sept. 1990). Agradezco a los profesores M. García-Carpintero y T. Mormann sus comentarios, críticas y la corrección de algunos errores.

¹ Cuando, como en este caso, no produzca confusión prescindiré de los corchets para nombrar esquemas de enunciados.

Así, p. e., si determinado contrafáctico es vago en determinado contexto, será porque en el contexto en cuestión no han quedado plenamente fijados sin vaguedad talés parámetros.

El valor de verdad de un contrafáctico tiene obviamente que ver con lo que *ocurriría* de ser las cosas diferentes a como son, al menos lo *suficientemente* diferentes como para que ocurra el antecedente (aunque tampoco pueden ser *demasiado* diferentes, como veremos); esto es, tiene que ver con lo que *ocurre* en otro(s) estado(s) de cosas diferente(s) a aquél en el que deseamos evaluar el contrafáctico². Para quien no tenga reparo en usar en el análisis conceptos como los de circunstancia, estado de cosas o mundo posible³, lo anterior sugiere la siguiente vía de análisis: el valor de verdad de un contrafáctico en un mundo depende, de un modo que hay que especificar, de los valores de verdad de antecedente y consecuente en ciertos otros mundos; en la especificación del modo de dependencia es donde se introducirán los parámetros del contexto respecto de los que se relativizan las condiciones de verdad del contrafáctico. Esta es la vía seguida por las semánticas de mundos posibles. Las dos más conocidas, la de Stalnaker y la de Lewis, introducen (o postulan) como parámetro cierta relación de similaridad entre mundos. La idea, que se inspira en Ramsey⁴, es que $\alpha \square \rightarrow \beta$ es verdadero (p. e. en nuestro mundo) si en los mundos en que ocurre α , y que se parecen al nuestro tanto como permita el que en ellos ocurra α , ocurre también β . El valor de verdad de un contrafáctico en un mundo es pues función de los valores de verdad de antecedente y consecuente en ciertos otros mundos «semejantes» al primero. La semejanza o similaridad entre mundos es el parámetro contexto-dependiente (y generalmente vago) respecto del que se relativiza el análisis o definición de las condiciones de verdad del contrafáctico⁵.

Aunque la relación de similaridad entre mundos depende del contexto y se determina (o «indetermina») pragmáticamente en cada caso, está sometida a ciertas constricciones formales, constricciones que debe satisfacer toda relación entre mundos para poder ser considerada de similaridad. Para referirme a tales condiciones,

² Esto es así para los contra-fácticos en sentido propio, e. e., cuando el antecedente no ocurre en el estado de cosas en que se evalúa.

³ Tanto si se consideran conceptos primitivos legítimos, como si se piensa que son a su vez analizables mediante otros. Como es sabido, Lewis defiende la existencia de mundos posibles no actuales (cf. p. e. 1973b 4.1 y 1986), pero muchos de sus ideas semánticas no dependen de este supuesto. Se pueden seguir por tanto sus desarrollos, al menos parcialmente, sin tomar los mundos posibles como primitivos, considerándolos, por ejemplo, como modelos de discurso. Una reinterpretación tal de las doctrinas de Lewis se encuentra, en filosofía de la ciencia, en algunos representantes de las llamadas 'concepciones semánticas de las teorías'; cf. p.e. Balzer 1984 (cap. V) y van Fraassen 1989 (p. 45 y c. 4).

⁴ La sugerencia, que guió el análisis de Stalnaker (cf. Stalnaker 1968 pp. 43 y ss.), se encuentra en Ramsey 1950 p. 248, aunque ahí se formula para conjuntos de creencias, no con mundos posibles. La intuición que hay tras esta idea es simple y parece adecuada: $\alpha \square \rightarrow \beta$ es verdadero si, añadiendo al actual estado de cosas y dejando todo lo demás igual, el resultado es tal que ocurre β . La cuestión, claro está, es que casi nunca es posible añadir α y dejar todo lo demás igual, pues la ocurrencia de α hará que algunas cosas cambien, obliga a ciertos «reajustes», y ciertos reajustes que dependen de cómo son las cosas actualmente.

⁵ Como recuerda Lewis (1973b 4.2), analizar un concepto vago mediante otro también vago, no convierte en vago (impreciso) el análisis.

y también más adelante, utilizaré principalmente el instrumental de sistemas de esferas introducido por Lewis, pues, aunque algo más engorroso, es equivalente al de las relaciones de similaridad y con él resulta más sencillo expresar la modificación que propondré. La idea es que un sistema de esferas (s. e.) para cierto mundo x es una colección «anidada» o nidificada (como unas muñecas rusas) de conjuntos de mundos progresivamente más disimilares a x ; cada conjunto-esfera de mundos está incluido en los «posteriores» y determina un *grado de semejanza* respecto de x . La definición es la siguiente (nótese que (3) y (4) sólo añaden algo cuando $\$$ es infinito)⁶:

- D1 Sea $x \in M$. $\$x$ es un sistema de esferas para x syss $\$x \subseteq \emptyset (M)$ y:
- (1) $\forall S \in \$x \ x \in S$
 - (2) $\forall S, S' \in \$x \ (S \subset S' \vee S' \subset S)$
 - (3) $\forall T \subset \$x \ (T \neq \emptyset \rightarrow \cup T \in \$x)$
 - (4) $\forall T \subset \$x \ (T \neq \emptyset \rightarrow \cap T \in \$x)$

Estos s. e. no son exactamente los de Lewis⁷, pero los aspectos en los que difieren son irrelevantes para lo que se va a discutir. Primero, los L.s.e. tienen siempre a \emptyset como elemento, una consecuencia antiintuitiva que evita D1 y que se puede probar que no altera las condiciones de verdad de los contrafácticos. Segundo, los L. s. e. son *centrados*, tienen a $\{x\}$ como elemento (la «primera» esfera, si prescindimos de \emptyset), mientras que los s. e. de D1 son sólo *débilmente centrados*, la primera esfera puede tener otros mundos además de x ⁸. Ello supone que, en ciertos sistemas de similaridad, puede haber mundos diferentes de x tan semejantes a x como x mismo. Lewis se opone a ello y, como en seguida veremos, exige que los s. e. respecto de los que se relativiza el valor de verdad de un contrafáctico sean siempre centrados. Ahora bien, ello no va a afectar directamente a lo que se desea discutir, pues que se exija o no que sean centrados sólo tiene consecuencias para la evaluación de contrafácticos con antecedente verdadero⁹.

⁶ Cf. Lewis 1973b, p. 14, aunque, como se comentará más adelante, esta def. no coincide exactamente con la de Lewis. Por otro lado, uso aquí ' $x \in M$ ' y ' $\$x \subseteq \emptyset (M)$ ' como meras abreviaturas, respectivamente, de ' x es un mundo posible' y ' $\$x$ es un conjunto de conjuntos de mundos posibles', sin presuponer que exista una entidad como el conjunto de todos los mundos posibles.

⁷ En los de Lewis (1) es « $\{x\} \in \$x$ » y (3) es « $\forall T \subseteq \$x \ (\cup T \in \$x)$ ».

⁸ El propio Lewis define unos s.e. prácticamente iguales, salvo por la presencia de \emptyset , a los de D1 (cf. 1973b p. 29).

⁹ Hasta qué punto se trata entonces de *contra*-fácticos es difícil de decir. En cualquier caso, es cierto que la aceptación de s.e. no centrados tiene consecuencias aparentemente extrañas para los contrafácticos *con antecedente verdadero*: si α y β son verdaderos en x , $\alpha \rightarrow \beta$ es verdadero en x pero (si la primera esfera contiene mundos diferentes de x) $\alpha \square \rightarrow \beta$ podría ser falso.

Por otro lado, parecería en principio que la exigencia de centralidad fuese intuitivamente obvia pues, como dice Lewis, «Es seguro que cada mundo es tan similar a sí mismo como cualquier otro [...] Casi tan seguro es que ningún otro mundo es tan similar a un mundo x como x mismo» (1973b, pp. 14-15). Si aquí, sin embargo, no la exigimos es porque esa (casi)seguridad se desvanece un tanto cuando nos apercebimos de que la similaridad respecto de la que se relativiza la verdad de un contrafáctico es similaridad en los aspectos *relevantes* para el contrafáctico y el mundo base puede *coincidir* con otro(s) en tales aspectos (sobre esta cuestión, cf. más adelante nota 26 y texto).

En un s. e. para un mundo x , los mundos involucrados en el sistema, los $y \in \cup \$x$, son los mundos «suficientemente similares» a x (en diversos grados), o, como se acostumbra a decir, los mundos *accesibles desde* x (relativamente a $\$x$). Así, $\cup \$x$, que es la última esfera, marca los «límites» de la similaridad, y dentro de esos límites las restantes esferas marcan los diversos grados de similaridad respecto del mundo base. Los mundos que quedan fuera de esos límites, los inaccesibles, no «cuentan» a la hora de evaluar el contrafáctico¹⁰. Los mundos inaccesibles son demasiado disimilares al base como para que importe, a la hora de saber si $\alpha \square \rightarrow \beta$ es verdadero en el base, lo que en ellos pase entre α y β ; en ellos no es relevante para la verdad de $\alpha \square \rightarrow \beta$ lo que ocurre con α y β (de donde, por contra y como insistiremos más adelante, en los accesibles –en todos ellos, en principio– sí importa).

Por otro lado, los s. e. inducen de forma natural una relación comparativa de similaridad sobre los mundos posibles:

D2 Sea $\$x$ un s. e. para x .

- (1) $y <_{\$x} z$ (y es tan o más similar a x , rel. a $\$x$, que z) syss
y pertenece a todas las esferas a las que pertenece z .
- (2) $y <_{\$x} z$ syss $y <_{\$x} z$ y no $z <_{\$x} y$.
- (3) $y \approx_{\$x} z$ syss $y <_{\$x} z$ y $z <_{\$x} y$.

Es inmediato ver que la relación de similaridad así inducida tiene las propiedades esperadas: $<_{\$x}$ es un orden débil (reflexivo, transitivo y conexo) sobre todos los mundos posibles, incluidos los inaccesibles; el mundo base es minimal (aunque quizá no es el único, si el s. e. no es centrado); los mundos inaccesibles son maximales; $\approx_{\$x}$ es una relación de equivalencia entre todos los mundos, las clases de equivalencia están formadas por los mundos que pertenecen a las mismas esferas (los inaccesibles forman también una clase).

Los s. e. pueden tener propiedades formales más fuertes que las expresadas por D1. Ya mencionamos una de ellas, la centralidad. Según D1 la primera esfera puede contener otros mundos además del base. Un s. e. para x es centrado, satisface la Suposición de Centralidad, si lo anterior no ocurre, esto es, si $\{x\}$ pertenece la sistema, o lo que es igual, si ningún otro mundo es tan similar a x como x . En los s.e. centrados, sin embargo, puede ocurrir que en otras esferas que no sean la primera sí haya mundos diferentes igual de similares al base. Cuando ello no ocurre decimos que el sistema satisface la Suposición de Unicidad: un s. e. satisface SU si en ninguna esfera hay mundos diferentes igual de similares al base. Por último, de D1 se sigue que \subseteq ordena linealmente los s. e., pero no que los bienordene, esto es, entre la primera esfera $\cap \$x$ y la última $\cup \$x$ puede haber cadenas infinitas ascendentes y/o descendentes. Cuando en un s. e. no hay cadenas infinitas descendentes diremos que satisface la Suposición del Límite. Es inmediato que SU implica SC y que SU y SL son independientes¹¹.

¹⁰ « $\alpha \square \rightarrow \beta$ es verdadero en un mundo si y sólo si β ocurre en ciertos α –mundos; pero ciertamente no todos los α –mundos importan» (Lewis 1973b, p.8).

¹¹ Son independientes en la versión que aquí hemos dado de ellos. SL es equivalente a la que da Lewis (cf. p.e. 1973b, pp. 19-20). SU no es equivalente a, sino más débil que, la que recibe el mismo

- D3 Sea $\$x$ un s. e. para x . $\$x$ satisface:
- (1) SC syss $\{x\} \in \$x$
 - (2) SU syss $\forall y, z \in U\$x (y \approx_{\$x} z \rightarrow y = z)$
 - (3) SL syss \subseteq bienordena $\$x$.

Estos supuestos separan a Lewis y Stalnaker a la hora de dar las condiciones de verdad de los contrafácticos. Para Lewis, los s. e. respecto de los que se relativiza la definición de verdad de un contrafáctico deben satisfacer sólo SC, mientras que para Stalnaker deben satisfacer SU y SL¹². Según Lewis, dado un s. e. (centrado) $\$x$ para x , $\alpha \Box \rightarrow \beta$ es verdadero en x relativamente a $\$x$ syss (i) (caso vacuo:) a no ocurre en ningún mundo accesible desde x , o (ii) (caso interesante:) ocurre en algún mundo z accesible desde x y en todos los mundos tan o más semejantes a x que z en los ocurre a ocurre también β . Abreviaré ' $\alpha \Box \rightarrow \beta$ es verdadero en x relativamente a $\$x$ ' mediante ' $\alpha \Box \rightarrow \beta$ es $V_{\$x}$ ' y ' α es verdadero en y ' mediante ' α es Vy ' (de un mundo en el que es verdadero diremos que es un « α -mundo»).

- D4 Sea $\$x$ un s. e. para x . $\alpha \Box \rightarrow \beta$ es $V_{\$x}$ syss (i) $\neg \exists y \in U\$x$ es Vy , o (ii) $\exists S \in \$x (\exists y \in S \alpha$ es $Vy \wedge \forall y \in S (\alpha$ es $Vy \rightarrow \beta$ es $Vy))$.

Así, en el caso interesante (no vacuo) $\alpha \Box \rightarrow \beta$ es verdadero en x si hay alguna esfera con algún α -mundo en la que todos los α -mundos son β -mundos. En D4 no hemos introducido la exigencia de Lewis de que el s. e. sea centrado, pero como ello sólo afecta al valor de verdad de contrafácticos en mundos en los que el antecedente es verdadero, y no es esa la cuestión que se quiere discutir, no tomaremos de momento en cuenta SC. Si el s. e. satisface SL, entonces (en el caso no vacuo) existe «la esfera más cercana a x con α -mundos», y si satisface además SU, entonces existe «el α -mundo más próximo a x ». De este modo, si se satisfacen ambos supuestos, $\alpha \Box \rightarrow \beta$ es (no vacuamente) verdadero en x syss el α -mundo más próximo a x es también un β -mundo; ésta es justamente la definición de verdad que da Stalnaker, que presupone por tanto SU y SL. Las diferencias entre las semánticas de Lewis y Stalnaker se pueden ver entonces como diferencias sobre las constricciones formales que debe cumplir un s. e. para que la relativización de las condiciones de verdad a un s. e. exprese adecuadamente la semántica de los contrafácticos. Estas diferencias tienen determinadas consecuencias, en las que no podemos detenernos ahora, sobre el comportamiento lógico de estas expresiones¹³.

nombre en Lewis, y que a veces llama también Suposición de Stalnaker (cf. 1973a p. 60 y 1973b p.78); el supuesto de Unicidad o de Stalnaker a que se refiere Lewis es equivalente a la conjunción de nuestro SU con SL (conjunción que equivale, a su vez, a que $\langle \$x$ bienordene $U\$x$).

¹² En realidad Stalnaker no presenta así las cosas, sino que presenta un aparato (funciones de selección) con cuya ayuda da una definición de verdad para contrafácticos que, como veremos, equivale a la de Lewis si en ella se exige (contra lo que hace Lewis) que los s.e. satisfagan SU y SL; en este sentido, como señala el propio Lewis (1973b p.78), la teoría de Stalnaker es un caso especial de la de Lewis.

¹³ Cf. p.e., Lewis 1973b cap. 3 y Stalnaker 1981 y 1987 cap. 7. La principal consecuencia es que la teoría de Stalnaker, contrariamente a la de Lewis, hace válida la ley del tercio excluso condicional «($\alpha \Box \rightarrow \beta$) \vee ($\alpha \Box \rightarrow \neg \beta$)» y, con ello, hace equivalentes a $\langle \neg(\alpha \Box \rightarrow \beta) \# y \langle \alpha \Box \rightarrow \beta \rangle$ a Stalnaker le parecen ambas cosas deseables, a Lewis no.

2. Las semánticas de Stalnaker y Lewis tienen pues diferentes consecuencias (y fundamentos). Ahora bien, a pesar de sus diferencias, comparten sin embargo un supuesto básico, que es el que aquí nos interesa: para ambos, las condiciones de verdad de un contrafáctico se relativizan exclusivamente a un sistema de similaridad¹⁴, y, supuestamente fijado por el contexto tal sistema, $\alpha \square \rightarrow \beta$ es verdadero en un mundo x si en los mundos en que ocurre α , y que se parecen (según el sistema) a x tanto como permita el que en ellos ocurra α , ocurre también β . Dado un sistema de similaridad, el valor de verdad del contrafáctico queda plenamente determinado. Ello no quiere decir, por supuesto, que después de todo los contrafácticos no son vagos, que en todo contexto está claro si son verdaderos o falsos. Lo que se quiere decir es que en la medida en que el contexto determina bien un sistema de similaridad, en esa misma medida las condiciones de verdad quedan bien determinadas: «el límite de vaguedad de la similaridad da cuenta adecuadamente del límite de vaguedad de los contrafácticos»¹⁵. El concepto de *sistema de similaridad* es un concepto teórico que pretende explicar el comportamiento semántico de los contrafácticos, concepto que indica el parámetro (contexto-dependiente y generalmente vago) del que dependen las condiciones de verdad del contrafáctico. Según Lewis (y Stalnaker), si un contrafáctico es vago o indeterminado en un contexto, es porque la similaridad también lo es; si dos hablantes competentes ideales, (o uno en dos ocasiones diferentes), difieren sobre el valor de verdad de un contrafáctico, es porque «manejan» similaridades diferentes. De otro modo: es imposible que en un contexto la similaridad esté plenamente determinada sin vaguedad alguna y que el contrafáctico gobernado por ella en ese contexto sea vago; es imposible que dos hablantes competentes compartan una misma similaridad y difieran sobre el valor de verdad del contrafáctico gobernado por ella. (Aunque la conversa no es cierta, pues un mismo contrafáctico puede ser verdadero relativamente a similaridades diferentes. Por eso, algunos contrafácticos pueden tener un valor determinado incluso si el contexto no determina unívocamente, sino sólo dentro de cierto rango, la similaridad¹⁶.)

Bien, la cuestión que, a mi juicio, sugiere la conveniencia de una modificación en la teoría se deriva de que, en mi opinión, en *algunas* circunstancias sí es *concebible* que un contexto determine completamente la similaridad y, a pesar de ello, permanezca vago el contrafáctico, o que dos hablantes compartan la misma similaridad y difieran en el valor de verdad del contrafáctico¹⁷. Ello hace que, aunque el modo en que Lewis explica la vaguedad de los contrafácticos es en general satisfactorio, si estoy en lo cierto habría *algunos* casos que requerirían otro tipo de explicación. Recuérdese que para Lewis (y Stalnaker), supuestamente fijado un sistema de si-

¹⁴ Uso aquí 'sistema de similaridad' para referirme indistintamente a sistemas de esferas y relaciones de similaridad, pues como hemos visto son expedientes equivalentes.

¹⁵ Lewis 1973b, p.94; cf. también Lewis 1979, p. 52: «la relación de similaridad [que gobierna las condiciones de verdad de los contrafácticos] es a la vez vaga y dependiente del contexto».

¹⁶ Cf. Lewis 1973b, p. 94.

¹⁷ Es obvio que si, en mi opinión, tal situación es concebible, alguno de mis conceptos «sistema de similaridad» y «verdad de un contrafáctico» (incluyendo en ambos su relación con el otro) *debe* diferir de los de Lewis (cf. más adelante el texto).

milaridad, el contrafáctico es verdadero o falso *simpliciter*, pues la similaridad es lo único respecto de lo que se relativiza el valor de verdad: $\alpha \square \rightarrow \beta$ es verdadero (prescindiremos del caso vacuo) si hay alguna esfera con algún α -mundo en la que todos los α -mundos son β -mundos, *aunque haya esferas más lejanas con mundos en los que α es verdadero y β falso*.

Los casos a los que me refiero, en los que el contrafáctico sería vago sin serlo la similaridad, requerirían para su explicación la relativización de las condiciones de verdad a otro parámetro además del sistema de similaridad. En esos casos su vaguedad se explicaría porque el contexto, que ha determinado completamente la similaridad, deja sin determinar ese segundo parámetro (relativo, como veremos, a cuán lejos se permite ir, según esa similaridad, para que todos los α -mundos sean β -mundos). Insisto, antes de ver el tipo de caso al que me refiero, en que todo esto parte del supuesto, imposible para Lewis, de que sea *concebible* una situación en el que el contrafáctico es vago y la similaridad relevante para él no lo es.

Bien, el tipo de caso al que me refiero viene ilustrado por el siguiente ejemplo¹⁸. Supongamos que hay un mundo x (el nuestro, si se quiere) con una barra A de longitud exactamente igual a $2m$ y otra B de longitud exactamente $1m$ (prescindiré en adelante de indicar la escala). Consideremos ahora el siguiente contrafáctico:

- (*) Si la longitud de A fuese mayor que 2 , la suma de las longitudes de A y de B sería menor que 3 .

Contemplemos un contexto en el que (*) no sea patentemente falso. Un contexto en el que esté implícito que al crecer A decrece B y que A no puede llegar a 3 . Supongamos, para hacerlo más plausible, que esa relación se da según una ley física aceptada en el contexto. Supongamos, por ejemplo, que la ley dice que si A crece $\delta \in [0,1]$, B decrece $4\delta^3 - 6\delta^2 + 3\delta$ (por tanto, no sólo si A mide más de 2 B mide menos de 1 , sino que, al ser siempre creciente la función que expresa la ley, cuanto más se aleja A de 2 , más se aleja B de 1). Supongamos además que se acepta que esa ley es lo único que afecta a las longitudes de A y B .

En mi opinión, es *concebible* que en un contexto tal quede completamente determinada la similaridad relevante para nuestro contrafáctico. Quizás, p. e., la siguiente: un mundo es más similar al nuestro que otro si en el primero A excede a 2 menos que en el segundo (y en cada caso B disminuye según la ley), y dos mundos son igual de similares al nuestro si en ambos A excede a 2 en la misma cantidad (y B disminuye según la ley). Esta *supuesta* similaridad (según la cual lo único relevante para (*) es qué ocurre con las longitudes de A y B) es la que correspondería al siguiente sistema de esferas $\$x$ (' L_yA ' abrevia a 'longitud de A en el mundo y '):

- (#) $S \in \$x$ syss hay un real $\delta \in [0,1]$ tal que:
 $\forall y (y \in S \Leftrightarrow 2 \leq L_yA \leq 2 + \delta \wedge L_yB = 1 - [4(L_yA - 2)^3 - 6(L_yA - 2)^2 + 3(L_yA - 2)])$

¹⁸ Quizás no es intuitivamente el mejor, pero tiene la ventaja de ser muy manejable formalmente; creo que, una vez señalado el punto, no sería difícil encontrar otros más apropiados.

Es fácil ver que $\$x$ es un s. e. (débilmente centrado) para nuestro mundo. En tal s. e., la primera esfera contiene a x y a todos los mundos en que, como en él, A mide 2 y B mide 1¹⁹. A partir de ahí, cada real δ entre 0 y 1 marca una esfera en cuyos mundos A mide entre 2 y $2+\delta$ y B mide lo que marca la ley (el resto de mundos son inaccesibles). Si $\$x$ *fuese* el s. e. que determina el contexto, (*) sería, según la teoría de Lewis, verdadero o falso *simpliciter*. En este caso verdadero, pues hay una esfera con mundos en los que $LA > 2$ y en todos esos mundos $LA + LB < 3$, por ejemplo, la esfera correspondiente a $\delta = 0,1$. En realidad, es sencillo mostrar que cualquier esfera con $\delta < 0,5$ en cuyos mundos es verdadero el antecedente también lo es el consecuente, mientras que en las esferas con $\delta \geq 0,5$ hay mundos en los que el antecedente es verdadero y el consecuente falso.

¿Es verdadero (*) en el contexto descrito? Creo que estaríamos de acuerdo en que la respuesta no es clara, que (*) es vago en ese contexto (si no se conviene en ello, habría que buscar otros ejemplos mejores, pero para lo que quiero expresar supongamos que convenimos aquí en ello). Bien, no pretendo que este ejemplo sea un contraejemplo a, un supuesto «fallo empírico» de, la teoría de Lewis. Los sistemas de similaridad no están «por ahí» para tomarlos (construirlos) a nuestro antojo e *imponer* después valores de verdad a los contrafácticos: construyo un s. e. para (*) en el contexto mencionado y, como en ese contexto (*) nos parece vago, y según la teoría de Lewis dado ese s. e. (*) no es vago, estamos entonces ante una anomalía empírica de la teoría de Lewis. Bien, obviamente las cosas no se pueden presentar así. Simplemente, desde la teoría de Lewis, si ciertamente (*) es vago en el contexto descrito entonces $\$x$ no es (ni él ni ningún otro determinado) el s. e. que lo gobierna. No sostengo pues haber dado con ningún contraejemplo empírico de la teoría. Lo que sostengo es que para mí (contrariamente a Lewis, para quien lo que sigue es simplemente una contradicción), en casos como el descrito es *concebible* que la similaridad esté determinada y el contrafáctico no, que, p. e., dos hablantes competentes compartan la misma relación de similaridad, coincidan en sus ideas sobre cuan semejantes son los mundos respecto del nuestro en lo relevante para el contrafáctico y, sin embargo, disientan sobre su valor de verdad. Si tal cosa fuese efectivamente concebible, entonces en estos casos haría falta otra explicación de la vaguedad del contrafáctico.

Lo que propongo considerar como alternativa es explicar este tipo de casos (si los hubiere, o simplemente, si en efecto son concebibles) diciendo que, aunque en ellos el contexto determina completamente la similaridad, el contrafáctico permanece vago porque, en general, para determinar las condiciones de verdad de un contrafáctico no basta determinar sólo la similaridad, sino que el contexto debe determinar otro parámetro, el «grado» de semejanza hasta el que se permite alejarnos de nuestro mundo, y ese parámetro no ha sido determinado en este caso. Puede que dos hablantes disientan sobre su verdad, aunque consideren la misma similaridad, si contemplan grados diferentes; o que se considere indeterminada su

¹⁹ La primera esfera no es pues necesariamente $\{x\}$ y por eso no es centrado. Pero ello no afecta por ahora directamente a lo que se discute pues el antecedente de (*) es falso en x (por otro lado, sería sencillo, si así se quisiera, dar esencialmente el mismo sistema pero centrado).

verdad, aunque no la similaridad, si no se considera ningún grado concreto. Esta alternativa supone pues que se debe modificar la definición de verdad de un contrafáctico relativizándola, no sólo a un sistema de similaridad, sino a un sistema tal y una esfera (grado) del mismo. Un contrafáctico es verdadero en x , relativamente a un sistema y una esfera del mismo, si todos los α -mundos hasta ese grado son también β -mundos. La definición de las condiciones de verdad sería como sigue, donde ' $\alpha \square \rightarrow \beta$ es $V_{\$x, S^*}$ ' es una abreviatura de ' $\alpha \square \rightarrow \beta$ es verdadero en x , relativamente a $\$x$, hasta grado S^* exclusive'²⁰.

D5 Sea $\$x$ un s. e. para x y $S^* \in \$x$. $\alpha \square \rightarrow \beta$ es $V_{\$x, S^*}$ syss (i) [como D4(i)] o
(ii) $\exists S \in \$x (\exists y \in S \alpha \text{ es } V_y \wedge \forall y \in S (\alpha \text{ es } V_y \rightarrow \beta \text{ es } V_y)) \wedge$
 $\forall S \in \$x (S \subset S^* \rightarrow \forall y \in S (\alpha \text{ es } V_y \rightarrow \beta \text{ es } V_y))$.

3. Antes de discutir otros aspectos de esta propuesta, veamos primero si se trata en realidad de una alternativa diferente a la de Lewis. Ya he advertido que mi pretensión no era presentar un contraejemplo empírico de la teoría de Lewis, sino ofrecer una explicación alternativa de algunos casos de vaguedad de contrafácticos, sin pretender, insisto, que ella tiene actualmente más rendimiento empírico que la de Lewis. Los supuestos casos de vaguedad como el del ejemplo, que yo explicaría diciendo que en ese contexto la similaridad está bien determinada, pero no el grado, Lewis los explica (como el resto) afirmando que la similaridad está indeterminada; o, la disensión entre dos hablantes competentes que yo atribuiría a la determinación de dos grados diferentes en una misma similaridad, Lewis la atribuye a la determinación de dos similaridades diferentes. Por tanto, las cosas se pueden ver (teorizar) de una forma o de otra. Cierto. Es más, se puede continuar, en el fondo mi propuesta no es diferente a la de Lewis, pues no sólo se pueden interpretar las cosas de ambos modos, sino que en el fondo no son sino dos modos de decir lo mismo (sencillo el de Lewis, gratuitamente complicado el mío). En efecto, «marcar» un s. e. hasta un grado (esfera) es *lo mismo* que dar otro s. e.: si $\$x$ es un s. e. y S^* es una esfera suya, es fácil mostrar que el conjunto $\{S \cap S^* / S \in \$x\}$ es también un s. e. para x (el s. e. que resulta de «recortar» de $\$x$ las esferas posteriores a S^* . Llamemos $\$x^*$ al s. e. inducido por el $\$x$ y la esfera $S^* \in \$x$. Parece entonces que es lo mismo relativizar las condiciones de verdad a un s. e. $\$x$ y un grado suyo $S^* \in \$x$ que relativizar tales condiciones exclusivamente a otro s. e., $\$x^*$. Bien, eso no es exacto. Es cierto que un s.e. $\$x$ y una $S^* \in \$x$ determinan otro²¹ s. e. $\$x^*$. Pero no lo es que D5, dados ciertos $\$x$ y S^* , sea equivalente a D4 para el $\$x^*$ generado por aquellos: (1) « $\alpha \square \rightarrow \beta$ es $V_{\$x, S^*}$ » y (2) « $\alpha \square \rightarrow \beta$ es $V_{\$x^*}$ » no son equivalentes. (1) implica (2) pero no al revés: (*) no es $V_{\$x, S^*}$ con, p. e., el $\$x$ de (#) y la S^* con $\delta = 0,7$, aunque sí es

²⁰ Defino «... hasta grado S^* exclusive ...» para dar cuenta de casos como el nuestro, en que podemos aproximarnos infinitamente a la esfera con $\delta = 0,5$ (siempre que no la alcancemos) sin hacer falso el consecuente. Para otros casos en los que la esfera límite (es la última que) hace verdadero el consecuente (como ocurriría con (*) si en el consecuente escribimos 'menor o igual' en lugar de 'menor') sería necesario definir «... hasta grado S^* inclusive ». Para ello basta sustituir en (ii) 'C' por ' \subseteq ' o, lo que es igual, cambiar la parte derecha de la conjunción por ' $\forall y \in S^* (\alpha \text{ es } V_y \rightarrow \beta \text{ es } V_y)$ '.

²¹ Otro, si S^* no es $\cup \$x$, en caso contrario obviamente $\$x^* = \x .

$V_{\$x^*}$ para el $\$x^*$ que se obtiene con esos $\$x$ y S^* . Por tanto, esta propuesta no es lógicamente equivalente a la de Lewis, es realmente una propuesta de explicación alternativa de los mismos fenómenos.

Alternativa sí, se dirá, pero no del tipo que se ha pretendido: se ha pretendido que lo esencial de esta propuesta es, manteniendo aproximadamente la misma idea de las condiciones de verdad de un contrafáctico, relativizarlas no a un único parámetro sino a dos, pero el párrafo anterior muestra que lo que se propone en el fondo es modificar esencialmente tales condiciones, manteniendo la relativización a un único parámetro. En efecto, D5 exige que hasta el grado $S^* \in \$x$ (exclusive) todos los α -mundos sean β -mundos, pero eso es lo mismo que exigir que $\alpha \square \rightarrow \beta$ sea «verdadero en x relativamente al s. e. Sx^* » según otra definición de verdad, relativizada también exclusivamente a un sistema de similaridad (sin referencia alguna al grado). Digamos que $\alpha \square \rightarrow \beta$ es verdadero en x , relativamente a un s. e. $\$x$ para x , cuando todos los α -mundos (salvo quizás los de la última esfera²²) son β -mundos:

D6 Sea $\$x$ un s. e. para x . $\alpha \square \rightarrow \beta$ es $W_{\$x}$ syss
 $\forall S \in \$x (S \neq \cup \$x \rightarrow \forall y \in S (\alpha \text{ es } Vy \rightarrow \beta \text{ es } Vy))$.

Ahora sí ocurre que (1) « $\alpha \square \rightarrow \beta$ es $V_{\$x, S^*}$ » y (3) « $\alpha \square \rightarrow \beta$ es $W_{\$x^*}$ » son equivalentes, de modo que no se trata en el fondo de introducir otro parámetro en la idea de Lewis, sino de modificar sustancialmente las condiciones de verdad, condiciones que después de todo seguirían estando relativizadas exclusivamente a un parámetro, el sistema de similaridad.

Bien, la propuesta podría verse así si se pretendiera que todos los casos de vaguedad requieren la explicación mencionada (vaguedad del grado dentro de una similaridad no vaga). Pero no es esa la pretensión. Ya hemos dicho que en la mayoría de los casos la explicación de Lewis nos parece la correcta, por lo que no es adecuado presentar la alternativa como D6, sino como D5, pues los casos «normales» (quizás la mayoría) serían entonces aquellos en los que resultaría superflua la referencia al grado, esto es, en los que serviría un grado cualquiera. En efecto, es inmediato que a partir de D5 se puede obtener D4 como un caso «especial» (aunque quizás mayoritario), pues (4) « $\alpha \square \rightarrow \beta$ es $V_{\$x}$ » es equivalente a (5) « $\exists S \in \$x \alpha \square \rightarrow \beta$ es $V_{\$x, S}$ ». Los casos normales serían pues los insensibles al grado.

Ahora bien, es cierto que eso sólo muestra que se pueden reubicar los casos normales dentro de nuestro esquema, pero no los «explica». En el plano teórico, esta alternativa debería dar cuenta de los casos normales en los que la relativización a un grado parece totalmente superflua, casos en los que parece no afectar al comportamiento semántico del contrafáctico el que en algún lugar los consecuentes empiecen a ser falsos (siendo los antecedentes todavía verdaderos). Ella debería decir, como hemos visto, que en esos casos el contexto es insensible al grado, que permite implícitamente un grado cualquiera (que basta con que sea verdadero para

²² Esta excepción se debe a que en D5 he definido «... hasta grado S^* exclusive»; si hubiera definido «... hasta grado S^* inclusive» (del modo indicado en la nota), no sería necesaria esta excepción para lo que sigue.

algún grado). Pero una cosa es decirlo y otra diferente explicarlo o justificarlo teóricamente. En este sentido, nos hemos buscado una complicación adicional, al exigir nuestra alternativa una *explicación* de la diferencia entre los casos normales, para los que bastaría el análisis de Lewis, y los supuestamente peculiares que han motivado la modificación²³.

4. Nuestra propuesta, por tanto, no sólo no es más adecuada empíricamente, sino que es teóricamente más compleja e inacabada. ¿Por qué entonces defenderla? Porque, como ya señalamos, nos parece *concebible* una situación en la que el contrafáctico sea vago y, sin embargo, el contexto determine completamente la relación de similaridad; nos parece concebible que, en un caso quizás como el de nuestro ejemplo, dos hablantes compartan exactamente la misma idea sobre la semejanza de dos mundos respecto del nuestro y, sin embargo, disientan sobre el valor de verdad del contrafáctico. Como una situación tal es simplemente contradictoria desde el análisis de Lewis, tras esta propuesta y tras la de Lewis debe haber conceptos diferentes de «similaridad», de «verdad de un contrafáctico» y de la relación entre ambos. El motivo último para defender esta propuesta debe radicar pues en una noción diferente de estos conceptos.

Comenzando por el concepto de similaridad, parecería en un primer momento que la diferencia básica entre nuestra noción y la de Lewis es que la suya es «global», mientras que la nuestra sería «parcial». Para Lewis, la similaridad respecto de la que se relativizan las condiciones de verdad de los contrafácticos es una «similaridad global (*overall similarity*) que consiste en innumerables parecidos y diferencias en innumerables aspectos de comparación, sopesados (*balanced*) unos con otros de acuerdo con la importancia relativa que concedemos a tales aspectos de comparación»²⁴. Frente a ella, parecería que la nuestra no es «tan global», que no toma en cuenta todos los innumerables aspectos de comparación en cada caso, sino sólo aquellos relevantes para el contrafáctico en cuestión (sean esos aspectos los que sean y pudiendo variar en diferentes contextos). Por eso estaríamos dispuestos a pensar que, en algunos contextos suficientemente específicos en relación a un contrafáctico (como el del ejemplo), es posible que los hablantes coincidan en cuán semejantes son los mundos en los aspectos relevantes para el contrafáctico, y que lo hagan independientemente de si también coinciden, o no, sobre el valor de verdad.

²³ Lo que debería explicar exactamente nuestra propuesta es por qué a veces basta la referencia a la similaridad para determinar el valor del contrafáctico, esto es, por qué a veces el contrafáctico no es vago cuando el contexto sólo ha determinado la similaridad y no el grado. La respuesta debe ser, como hemos visto, que en esos casos el contexto permite implícitamente un grado cualquiera, pero eso es algo que también debe ser explicado. Parecería que Lewis se encuentra con una situación semejante cuando afirma que en ciertos contextos un contrafáctico puede no ser vago aun siendo vaga, o determinada sólo dentro de cierto rango, la similaridad: «otros muchos [contrafácticos] tienen bien definidos valores de verdad, siendo insensibles a pequeños cambios en nuestros estándares de similaridad comparativa» (1973b p. 94). Pero él dispone de una explicación que no requiere otra ulterior, a saber, que en esos casos el contrafáctico es verdadero relativamente a *cada una* de las similaridades del rango.

²⁴ Lewis 1973, p. 91; el texto continúa: «En la medida en que estas importancias relativas difieren de una persona a otra, o de una ocasión a otra, o están indeterminadas incluso para una única persona en una única ocasión, en esa medida la similaridad comparativa es indeterminada».

Bien, aunque la diferencia tiene que ver con esto, no se puede formular exactamente en estos términos, pues la similaridad de Lewis es también «parcial» (en el sentido indicado), o si se prefiere, la nuestra también es global. En efecto, Lewis afirma también que la similaridad corresponde a cada contrafáctico y contexto; por tanto, aunque los respectos de comparación sean innumerables, al sopesarlos globalmente puede que a algunos, si en ese contexto se les considera irrelevantes para el contrafáctico, no se les conceda *ninguna* importancia y, en consecuencia, dos mundos que difieran sólo en (alguno de) esos aspectos serían equivalentes según la relación de similaridad. Dicho brevemente, aunque la similaridad de Lewis es global, ello no impide que dos mundos que difieren en algunos aspectos sean, en cierto contexto para cierto contrafáctico, igual de similares al base²⁵. La diferencia, respecto de Lewis, no es pues que nuestra similaridad no sea global, que tome en cuenta, en cierto contexto, sólo los aspectos relevantes para el contrafáctico en cuestión, pues lo mismo hace en el fondo la de Lewis.

La diferencia radica más bien, por expresarlo de algún modo, en que nuestra similaridad es algo más «rígida» que la de Lewis, esto es, en determinados contextos para ciertos contrafácticos, es posible que relativamente pocos factores sean los relevantes para el contrafáctico y los que determinan propiamente la similaridad (en el sentido especificado en el párrafo anterior). En casos como el del ejemplo me parece plausible que, en lo relativo o relevante para el contrafáctico, el contexto descrito (en ausencia de otros factores) determine suficientemente la similaridad, e.e., la semejanza relativa de dos mundos respecto del nuestro en *los aspectos relevantes para el contrafáctico*. Quizás haya hablantes para los que no quede plenamente determinada, hay muchos factores que intervienen en un contexto, sobre los que no hemos dicho nada en nuestro ejemplo, que podrían dejar vaga la similaridad o producir desacuerdos en los hablantes respecto de ella. Pero me parece posible que en un contexto como el nuestro no intervengan (o no produzcan efecto sobre la similaridad, o produzcan en todos los hablantes el mismo) otros factores que los indicados y que en él dos hablantes compartan (el conocimiento de) la misma similaridad relevante al contrafáctico. Si en una situación tal (y en ausencia de otros factores) el contexto deja algún cabo suelto en lo referente a la similaridad relevante para nuestro contrafáctico, me es difícil entender entonces en qué consiste la idea de similaridad.

²⁵ La negación de esto es precisamente el Supuesto de Unicidad, que como sabemos rechaza Lewis. Por cierto que, si la similaridad de Lewis a pesar de ser global permite que dos mundos coincidan en su parecido a otro, no se comprende muy bien su insistencia en el Supuesto de Centralidad, en que ningún otro mundo es tan similar al base como él mismo. Como sabemos, SC es SU aplicado sólo al base, pero no parece haber motivos para distinguir un caso de los otros: si y, z pueden coincidir en su semejanza respecto de x en los aspectos relevantes para el contrafáctico, ¿por qué no puede haber un mundo diferente de x tan semejante a x como x mismo en tales aspectos? Lewis aduce (1973b p. 15) que aunque eventualmente haya mundos distintos de x cualitativamente indistinguibles de x, x se parece a x en algo en lo que los otros no, a saber, en ser idéntico a x. Pero si la identidad respecto de sí mismo cuenta para distinguir en semejanza al base de otros, ¿por qué no cuenta para distinguir en semejanza a y, z? Porque, se dirá, se trata de semejanza respecto del base x. Pero no creo que esta cuestión técnica sea lo importante; después de todo se podría argüir que y se parece a x en algo en lo que z no (y viceversa), a saber, en ser diferente de x *siendo* y.

Bien, se dirá, nada de ello se opone a Lewis, pues él no dice que en un caso como el visto la similaridad no pueda ser determinada. Lo que dice es que no puede ocurrir a la vez que la similaridad sea determinada y el valor de verdad del contrafáctico no, que dos hablantes compartan la misma similaridad y disientan sobre la verdad de (*): si, como hemos convenido, (*) nos parece indeterminado en el contexto descrito, es porque la similaridad también nos lo parece; es posible que ese contexto nos determinara completamente la similaridad, pero en ese caso no podríamos convenir en que el contrafáctico es vago. La (supuesta) diferencia entre ambos conceptos de similaridad no justifica por sí sola la defensa de la propuesta, en cuyo favor se aduce, no que es concebible que en casos como el del ejemplo dos hablantes compartan la similaridad, sino que es concebible que compartiéndola, estando de acuerdo sobre la semejanza de los mundos respecto del nuestro en los aspectos relevantes para (*), disientan sin embargo sobre el valor de verdad de (*). Por tanto, tras esta propuesta debe haber además un concepto diferente de «verdad de un contrafáctico». Efectivamente así es.

La idea que inspira el análisis de Lewis (y Stalnaker), y que se remonta a Ramsey, es que *un contrafáctico es verdadero si en los mundos en que ocurre el antecedente y que se parecen al nuestro tanto como permita el que en ellos ocurra el antecedente, ocurre también el consecuente*; si, como Stalnaker, se exige SL, entonces existe «la mínima»²⁶ variación que permite que se cumpla el antecedente (en términos de Ramsey, la mínima modificación de nuestras creencias) y el contrafáctico es verdadero si en esa mínima modificación también se cumple el consecuente. Pues bien, la diferencia mencionada sobre el concepto de verdad de un contrafáctico tiene que ver con esa idea. En mi opinión, aunque ella es apropiada en la mayoría de los casos, no lo es siempre, sobre todo puede no serlo en algunos casos en los que, como el de nuestro ejemplo, se incumple SL. Lo que quiero decir es lo siguiente.

Supongamos que cierto contexto determina completamente la similaridad relevante para cierto contrafáctico y que ella es tal que se inicia, salvo por la primera esfera, con una cadena infinita descendente de esferas, e. e., no existe «la segunda» esfera; supongamos además que algunas de esas esferas contienen α -mundos y todos esos α -mundos son también β -mundos, pero que esas esferas son, respecto del total de las que determinan los mundos accesibles, *extremadamente* escasas y *extremadamente* próximas a la primera, es decir, que «en seguida» (en la cadena infinita) aparecen α -mundos que no son β -mundos. Esto es, supongamos que, dentro del conjunto de variaciones mundanas que se consideran aceptables o relevantes para la evaluación del contrafáctico (mundos accesibles), sólo en algunas pocas que difieren extremadamente poco de nuestro mundo es verdadero el consecuente cuando lo es el antecedente, mientras que en las restantes (la «inmensa mayoría») es verdadero el antecedente pero no el consecuente; variaciones restantes que, este es el punto, deben en algún sentido ser *también relevantes para la evaluación del contrafáctico pues quedan dentro de los límites de accesibilidad*. Por ejemplo, conside-

²⁶ En sentido genérico, esto es, la clase de mundos con la mínima variación etc.; como mencionamos anteriormente, para que exista propiamente «el» mundo más cercano en que etc., se ha de cumplir SU además de SL.

remos un caso para nuestro (*) en el que, *supongamos*, la similaridad de hecho determinada por el contexto es tal que sólo cuando es menor que, digamos, 10^{-500} los α -mundos son también β -mundos, en todos los demás mundos accesibles, la abrumadora mayoría, ocurre α y no ocurre β . Pues bien, en casos como éste (supuesta esa similaridad), yo no diría que el contrafáctico es verdadero *simpliciter*. En realidad, si se me obligara a afirmar algo *simpliciter*, me inclinaría quizás (en casos tan extremos como éste) antes por la falsedad, aunque, como vengo defendiendo, no creo que en estos casos se pueda afirmar algo *simpliciter*: dada la similaridad, el contrafáctico es verdadero hasta cierto grado²⁷.

Se dirá quizás que estas consideraciones parecen olvidar que lo que interesa determinar es el valor de verdad del contrafáctico en nuestro mundo (o en el mundo base) y no en alguno de los mundos accesibles. Si recordamos eso, se dirá, tiene razón Lewis al afirmar que no importa qué pase después con el resto de los accesibles si, aunque sean extremadamente pocos y extremadamente semejantes al nuestro, hay α -mundos que son β -mundos y tales que en ningún α -mundo anterior a ellos β es falso; dicho de otro modo: como nos interesa el valor de verdad del contrafáctico en nuestro mundo, nos vamos alejando de él tan poco a poco como sea posible, y en cuanto encontremos mundos en los que α y β son verdaderos sin haber encontrado antes otros en los que lo sea α pero no β , «paramos de contar» sin importar qué pase después.

Bien, en mi opinión, que lo que nos interese sea el valor de verdad en nuestro mundo, no invalida aquellas consideraciones. La cuestión es que para el valor de verdad *en nuestro mundo* deben ser relevantes todos los mundos accesibles, pues *los mundos accesibles son justamente aquellos que se parecen (todos ellos, aunque no todos de igual modo) al nuestro lo suficiente como para que lo que ocurra en ellos se considere relevante para evaluar el contrafáctico en nuestro mundo*. Aunque haya una esfera en la que todos los α -mundos son β -mundos, hay que seguir viendo qué pasa con los α -mundos posteriores que sean accesibles, e. e., que se consideran (de acuerdo con la definición de sistema de similaridad) todavía *suficientemente semejantes* al base. Dicho de otra forma, la teoría de Lewis no da cuenta de una diferencia que, en mi opinión, debe ser tomada en consideración (al menos para algunos casos). La diferencia es la siguiente. Supongamos que, para cierto contrafáctico $\alpha \square \rightarrow \beta$ y mundo x , dos contextos diferentes determinan dos similaridades diferentes $\$x_1$ y $\$x_2$ tales que en ambas hay al menos una esfera (que no es $\cap \$$) con α -mundos en la que todos los α -mundos son β -mundos, pero mientras en $\$x_1$ son extremadamente pocas y extremadamente próximas a x (p. e., algunas pocas que vienen inmediata-

²⁷ Desde luego que si el contrafáctico se considerase claramente falso, Lewis tendría la respuesta: la similaridad en cuestión no es la que se ha pretendido sino otra en la que todos esos α -mundos (aquellos en los que β es verdadero y al menos algunos de aquellos en los que β es falso) son equivalentes entre sí. Pero yo no estoy considerando ahora la diferente explicación de la falsedad de un contrafáctico, sino qué dice la teoría de Lewis en un *hipotético* caso en que la similaridad fuese *efectivamente* como la indicada. Y está claro que si las similaridades no son comodines que explican *por definición*, sino que tienen ciertas condiciones de identidad *propias*, por así decir, entonces *puede* realmente haber una similaridad relevante para un contrafáctico como la indicada, y podemos preguntarnos si lo que en esos casos afirma la teoría de Lewis es siempre conforme a nuestras intuiciones.

mente después de \cap), en Sx_2 son todas las esferas (salvo \cap), o si se prefiere, todas salvo algunas pocas de las últimas). Es decir, en Sx_1 la mayoría de los α -mundos no son β -mundos, mientras que en Sx_2 todos (o la mayoría) lo son. Pues bien, según Lewis no hay diferencia semántica alguna para el contrafáctico en ambos contextos, en ambos es simplemente verdadero. En mi opinión, sí hay alguna diferencia o, dicho de otro modo, si no la hay²⁸ entonces la noción de mundo accesible, de *estado de cosas lo suficientemente semejante al nuestro como para que importe lo que ocurre en él para evaluar el contrafáctico en el nuestro*, me resulta incomprensible²⁹. La alternativa que hemos presentado se puede entender como destinada a recoger esta (presunta) diferencia, pues ambas similaridades harían verdadero el contrafáctico en *diferente grado*.

Hasta aquí la (vaga y tentativa) explicación del diferente uso de los conceptos de «similaridad» y «verdad de un contrafáctico» que hace nuestra propuesta en relación a la de Lewis. Esta diferencia conceptual es la responsable de lo que ha motivado la propuesta, a saber, que nos parece concebible que dos hablantes competentes coincidan sobre la semejanza de otros mundos respecto del nuestro en los aspectos relevantes para cierto contrafáctico y disientan sin embargo sobre el valor de verdad del contrafáctico. De ser efectivamente posible, una situación tal se explicaría por la diferencia en el grado dentro de una misma similaridad; como dijimos, no se afirma que todo caso de desacuerdo o vaguedad se deba explicar así, pues creemos que en la mayoría de los casos se debe seguramente a la diferencia o vaguedad de la similaridad. Por otro lado, esta alternativa debe explicar, y no lo hace de momento, por qué a veces el contexto es insensible al grado, esto es, por qué a veces (como parece que ocurre en la mayoría de ocasiones) se permite un grado *cualquiera*.

La insistencia en que la alternativa presentada se basa en conceptos diferentes no es una estrategia elusiva, no pretende eludir la posible crítica aduciendo: «bien, en cualquier caso, estos son mis conceptos de similaridad y de verdad de un contrafáctico». No pretendo, desde luego, refugiarme en semejante solipsismo filosófico. Cuando me refiero a los conceptos de similaridad y de verdad de un contrafáctico pretendo capturar las mismas cosas que las que pretende capturar Lewis, y puesto que ambas pretensiones tienen diferencias sobre lo que es concebible, alguno está

²⁸ Si no la hay nunca, o mejor, si no puede haberla, esto es, si queda excluida, como en Lewis, *a priori*. Como ya dijimos, no defendemos que siempre la haya, sino que puede haberla, y (como hipótesis empírica) que probablemente algunas veces la hay.

²⁹ De nuevo, si las dos supuestas similaridades fuesen efectivamente como las que se indican. De no ser así, las diferencias semánticas explicarían en la teoría de Lewis perfectamente mediante similaridades que hicieran el contrafáctico en un caso verdadero y en el otro falso (si es que esa es la diferencia semántica) o que lo dejaran indeterminado en un caso y no en otro (si es que es esta la diferencia observada). Lo que yo propongo es *suponer*, sí, pero suponer que *efectivamente* los contextos determinan para un contrafáctico esas similaridades, y ver qué afirma la teoría de Lewis en ese caso y si ello se adecua a nuestras intuiciones. La función de este experimento mental es por tanto sólo «persuasiva», no decisiva. La única situación decisiva sobre toda esta cuestión (ver penúltimo párrafo del texto) sería una en que se hubiera identificado independientemente la similaridad relevante para un contrafáctico y ella no «encajara» con lo que la teoría de Lewis predice sobre el comportamiento semántico observado del contrafáctico. Obviamente, es difícil que se llegue alguna vez a poder identificar de modo tal los sistemas de similaridad, por lo que es difícil poder dilucidar *empíricamente* entre las dos alternativas.

equivocado. Eso, claro está, en el supuesto de que mis conceptos sean siquiera sensatos, pues podría ocurrir que una reflexión más penetrante muestre que son sin más «totalmente extravagantes», que ni siquiera quien dice concebir lo que se sigue de ellos lo concibe efectivamente, esto es, que son meras palabras más que conceptos.

Pues bien, si se trata en verdad de un conflicto entre conceptos mínimamente sensatos, debe ser posible establecer la situación que lo dirimiría. La similitud entre mundos en un contexto no es un comodín que de cuenta *por definición* del comportamiento semántico de los contrafácticos. Es algo cuyo conocimiento (generalmente implícito) hay que atribuir al hablante, que el hablante tiene *in mente* y que puede o no compartir con otros; algo que debe tener ciertas condiciones de identidad, con ciertos elementos relevantes que intervengan en tales condiciones, ciertos «respectos de comparación» conforme a los cuales se establece (en un sentido no arbitrario ni subjetivo de ‘establecer’). El propio Lewis ha apuntado alguno de ellos, como la coincidencia en historias parciales de los mundos o en sus leyes³⁰. Supongamos por un momento que es posible acabar ese trabajo e «identificar» la similitud entre mundos en un contexto. Ello permitiría «juzgar» cuál de las alternativas es la correcta: si se *encontrara* un contexto en el que un contrafáctico fuese vago pero la similitud de la que depende no lo fuera, ello sí constituiría entonces un contraejemplo a la teoría de Lewis, un caso de algo que según la teoría es imposible, contraejemplo para el que mi alternativa ofrecería una explicación. Sin embargo, este modo de decidir la cuestión tiene menos importancia práctica que teórica, pues es muy improbable que investigaciones futuras capturen completamente la escurridiza noción de similitud, y sólo sirve para confirmar que ciertamente se trata de alternativas diferentes.

En el mejor de los casos, pues, sólo puedo aducir en favor de mi alternativa la confianza en que, si llega a ser posible «identificar» la similitud entre mundos (relevante para un contrafáctico) en un contexto, se hallarán casos de contrafácticos vagos cuya similitud asociada no lo sea. En el peor, si la alternativa se basa en una noción de similitud totalmente extravagante, lo expuesto es sólo una explicación de cómo he venido a dar con ella. Como ninguna de ambas cosas es desde luego una razón, se debe considerar la modificación presentada, más que como una alternativa a defender, como una propuesta a explorar.

Bibliografía

- Balzer, W. (1984), *Theorie und Messung*, Springer, Berlin.
Harper, W. L. y Stalnaker, R. (eds.) (1981), *IFS*, Reidel, Dordrecht.
Lewis, D. (1973a), ‘Counterfactuals and Comparative Possibility’, en W. L. Harper y R. Stalnaker (eds.) 1981, pp. 57-85.
– (1973b), *Counterfactuals*, Harvard U.P., Cambridge.
– (1979), ‘Counterfactual Dependence and Time’s Arrow’, *Philosophical Papers II*, Oxford U.P., Oxford 1986, pp. 32-66.

³⁰ Cf. Lewis 1979 y su «postscript».

- (1983), 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 343-377, pp. 61.
- (1986), *On the Plurality of Worlds*, Blackwell, Oxford.
- Ramsey, F. P. (1950), *Foundations of Mathematics and other Logical Essays*, New York.
- Stalnaker, R. (1968), 'A Theory of Conditionals', en W. L. Harper y R. Stalnaker (eds.) 1981, pp. 41-55.
- (1981), 'A Defense of Conditional Excluded Middle', en W. L. Harper y R. Stalnaker (eds.) 1981, pp. 87-103.
- (1987), *Inquiry*, M. I. T., Cambridge.
- van Fraassen, B. (1989), *Laws and Symmetry*, Clarendon, Oxford.

José A. DIEZ CALZADA
Universitat Rovira i Virgili