

# CREENCIAS, LÓGICA Y DECISIONES RACIONALES: A PROPÓSITO DE LA PARADOJA DE NEWCOMB

## BELIEFS, LOGIC AND RATIONAL CHOICE: ON THE NEWCOMB PARADOX

Francisco Díaz Montilla

Universidad de Panamá, Panamá

Email: francisco.diazm@up.ac.pa

*\*Recibido: 15 de abril de 2021; aprobado: 4 de mayo de 2021*

**Abstract:** The work explores some issues related to beliefs, logic, and rational decisions, taking Newcomb's paradox as a reference. The article explores the ways in which this paradox has traditionally been seen, pointing out that it is not possible to speak of an incorrect position, given the conceptual basis that both argumentative strategies use, although one of them stands out due to the practical consequences it has.

The problem of determinism is also dealt with, and it is argued - following Dennett - that determinism is not necessarily incompatible with freedom.

**Key words:** Newcomb's paradox, expected utility, dominance, (in)determinism, freedom.

**Resumen:** El trabajo explora algunas cuestiones relacionadas con las creencias, la lógica y las decisiones racionales teniendo como referente la paradoja de Newcomb. El artículo explora las maneras en que se ha visto tradicionalmente esta paradoja, señalando que no se puede hablar de una postura incorrecta, dado los parámetros que ambas estrategias argumentativas usan, aunque una de ellas sobresale por las consecuencias prácticas que tiene.

También se trata el problema del determinismo y se argumenta -a partir de Dennett- que el determinismo no necesariamente es incompatible con la libertad.

**Palabras clave:** Paradoja de Newcomb, utilidad esperada, dominancia, (in)determinismo, libertad.

## 0. Introducción

Si se le pidiese que elabore una lista de  $n$  pensadores (científicos) reconocidos en la historia de la física, es muy probable que Newton, Laplace, Maxwell, Einstein, Bohr, Schrödinger, Heisenberg, Feynman estén en ella; y es muy poco probable, para no decir imposible, que en ella se encuentre William Newcomb.

De hecho, tal vez no haya mucho que decir, salvo que se desempeñó en los laboratorios de radiación en Livermore, California. Incluso Wikipedia presenta pocas referencias sobre su vida, y la resume en ¡78 palabras!

Sin embargo, Newcomb es famoso por la paradoja que lleva su nombre, aunque ha sido Robert Nozick quien la diera a conocer en el mundo académico, tras publicar un enjundioso artículo sobre el tema al cual tituló, *Newcomb`s Problem and two principles of choice* (*El problema de Newcomb y dos principios de elección*), el cual apareció inicialmente en (Rescher, 1969). Para este trabajo nos apoyaremos en la versión en español que aparece en (Nozick, 1999, pp. 73-109).

La paradoja sería popularizada por Martin Gardner, el célebre divulgador científico y filósofo (de la ciencia) en su columna *Juegos Matemáticos* de la revista *Scientific American*, lo que obligó a Nozick a volver sobre el tema en un segundo artículo aparecido en 1974 para responder a más de un centenar de cartas enviadas por los lectores a la revista [véase (Nozick, 1999, pp. 111-112)].

En este artículo haré una descripción de la paradoja, presentaré y comentaré las maneras en que tradicionalmente ha sido tratado el problema describiendo el núcleo argumentativo de las dos posiciones que se debaten; procuraré responder a la pregunta de si alguna de ellas es ilegítima y comentaré algunas repercusiones filosóficas, específicamente en el debate del (in)determinismo. Luego presentaré algunas conclusiones generales.

En términos generales, se argumenta que los dos modos de ver la paradoja se fundamentan en supuestos racionales que generan soluciones divergentes, lo cual no implica que alguno de ellos sea ilegítimo; aunque considerando las consecuencias prácticas de ambos, sería preferible aplicar el criterio de utilidad esperada sobre el criterio de dominancia. Actuar de acuerdo con dicho criterio, no necesariamente implica incompatibilidad con la libertad, como algunos defensores del criterio de dominancia postulan.

## 1. La paradoja

La situación va de la siguiente manera (Nozick, 1999, p. 74):

Supongamos que existe un ser en cuyo poder para pronosticar lo que uno elige tenemos una gran confianza<sup>1</sup> (...). Sabemos que en el pasado este ser ha hecho frecuentes predicciones correctas acerca de lo que elegimos (y que nunca, que sepamos, ha hecho una predicción incorrecta acerca de lo que elegimos), y además sabemos que este ser ha hecho frecuentes predicciones correctas acerca de lo elegido por otras personas...

Tenemos dos cajas, B<sub>1</sub> y B<sub>2</sub>. B<sub>1</sub> contiene 1000 dólares. B<sub>2</sub> contiene o bien 1000000 de dólares o bien nada...

Podemos elegir entre dos acciones:

1. Tomar lo que hay en ambas cajas.
2. Tomar lo que hay en la segunda caja.

Además, y **esto es algo que sabemos, el ser en cuestión sabe que lo sabemos, y así sucesivamente:**

- i. Si el ser vaticina que cogeremos lo que hay en ambas cajas, no mete los M dólares en la segunda caja.
- ii. Si el ser vaticina que cogeremos sólo lo que hay en la segunda caja, mete los M dólares en la segunda caja. (Las negritas son nuestras).

La pregunta, entonces, es: ¿qué hacemos?

## 2. Dos maneras de ver el problema

Quienes se enfrentan al problema se posicionan en dos bandos. Aunque Nozick no los califica, diremos que quien opta por la segunda caja es -para propósitos descriptivos- un *saiyajin*; en cambio, quien opta por ambas cajas es un *maestro jedi*.

El *saiyajin* razona más o menos de la siguiente manera: *Si elijo ambas cajas, entonces -dado que de seguro el Profesor X lo ha predicho- no pone el millón de dólares en la segunda, y obtendré solamente mil dólares. Por el contrario, si elijo solamente la segunda caja, entonces -dado que de seguro Profesor X lo ha predicho- pone el millón de dólares en la segunda caja, y obtengo el millón de dólares. Dado que prefiero un millón de dólares a mil, elijo solamente la segunda caja.*

---

<sup>1</sup> Asumiremos que se trata de Profesor X, Charles Xavier.

Por su parte, el *maestro jedi* razona del siguiente modo: *Independientemente de lo que yo decida, el Profesor X ha puesto o no ha puesto el millón de dólares en la segunda caja, y no puede cambiar su decisión; como se trata de un estado determinado, si no están y elijo ambas cajas, habré obtenido mil dólares. En cambio, si están y elijo ambas, habré obtenido un millón mil dólares. Dado que prefiero un millón mil dólares a mil, elijo ambas cajas.*

Señala Nozick que, al publicarse el problema en *Juegos Matemáticos*, se recibió una avalancha de cartas de los lectores, emitiendo sus opiniones sobre la perspectiva *correcta* del problema. Y aunque algunas de ellas suponían la introducción de elementos que tal vez nada tienen que ver con el *núcleo* del problema, las posturas eran las siguientes: De 126 lectores, el 70.6% indicaba que elegiría solamente la segunda caja, y el 29.4% restante indicaba que elegiría ambas cajas.

Para ver qué perspectiva se puede tener del problema en el presente y en nuestro contexto, envié a diversas personas una encuesta donde preguntábamos qué elegirían, aunque sin entrar en detalles sobre cómo conceptualizaban sus respuestas.

Participaron 52 personas, de las cuales el 67.3% indicó que elegiría solamente la segunda caja, mientras que el 32.7% indicó que elegiría ambas cajas. Pareciera, por tanto, que hay más *saiyajin* que *maestros jedi*, lo cual coincide de un modo bastante aproximado con los resultados reportados por Nozick inicialmente.

Del total de participantes en nuestra encuesta, el 59.6% eran mujeres y el 40.4% eran hombres. Por los resultados, no habría razones para suponer que hay formas distintas de abordar la situación entre uno y otro grupo. De hecho, la probabilidad de que una mujer eligiera ambas cajas fue de 0.323 vs 0.333 de los hombres; de igual manera, la probabilidad de que una mujer eligiera solo la segunda caja fue de 0.677 vs 0.667 en los hombres. Desde luego, estos resultados no pretenden ser científicos (cualquier cosa que ello signifique) ni mucho menos, y solamente ilustran que hay tantos *saiyajin* como *maestros jedi* en ambos grupos.

### 3. El núcleo de las estrategias de los argumentos

Cuando se trata de elegir entre acciones alternativas, salen a relucir dos principios: el de utilidad esperada y el de dominancia. De acuerdo con el primero:

Entre las acciones que le son asequibles, una persona realiza la acción que promete tener un grado máximo de utilidad (Nozick, 1999, p. 77).

La utilidad esperada es la suma de los productos de los resultados exclusivos con las probabilidades respectivas. Por ejemplo, invierto un dólar en un billete de lotería, ¿cuál es la esperanza de ganar? Dado que los premios son tres, y cada uno de ellos paga una cantidad distinta: 2000, 600 y 300 dólares respectivamente, y la probabilidad de ganar un premio es de  $1/10000$ , entonces la esperanza matemática de ganar es:  $E(G) = \frac{1}{10000} \times 2000 + \frac{1}{10000} \times 600 + \frac{1}{10000} \times 300 = 0.29$ .

Claramente la estrategia del *saiyajin* asume este criterio:  $E(1000000) > E(1000)$ , dado que en ambos casos las probabilidades de que Profesor X prediga correctamente son altas.

De acuerdo con el segundo criterio,

Si se da una participación de estados del mundo tal que, en lo relativo a la misma, la acción A ejerce un dominio total sobre la acción B, entonces debe ejecutarse A más bien que B (Nozick, 1999, p. 78).

Claramente, la estrategia del *maestro jedi* asume este criterio: 1001000 dólares dominan el resto de las opciones, y por ello esta es la que asume, dado que los estados han sido ya determinados por Profesor X.

La situación de la utilidad esperada puede expresarse de la siguiente manera:

		Profesor X	
		Predice correctamente	No predice correctamente
Usted	Elige ambas cajas	1000	1001000
	Elige solo B <sub>2</sub>	1000000	0

Es obvio que, de acuerdo con el criterio de utilidad esperada, la esperanza involucrada en elegir sola la segunda caja es mayor que la esperanza involucrada en ambas cajas. Supongamos que no se tiene idea del registro de éxitos de Profesor X, y que guiándonos por el criterio de racionalidad laplaciano, las probabilidades son iguales: 0.5. En ese caso, es obvio que hay que elegir ambas cajas. Nótese que, de esta perspectiva, no se trata únicamente de cuánto se obtenga o se deje de obtener, sino de las probabilidades de que ello ocurra.

De modo similar, la situación de dominancia puede expresarse de la siguiente manera:

		Profesor X	
		Puso el millón en B <sub>2</sub>	No puso el millón en B <sub>2</sub>
Usted	Elige ambas cajas	1001000	1000
	Elige solo B <sub>2</sub>	1000000	0

Como se puede constatar, la estrategia *elegir ambas cajas* domina estrictamente a la estrategia *elegir solo B<sub>2</sub>*, pues ésta en el peor escenario devuelve un pago de 1000 y en el mejor escenario, 1001000.

Desde el punto de vista intuitivo, a pesar de que desembocan en acciones distintas, ambas «argumentaciones parecen correctas», sin embargo -agrega Nozick- «El problema consiste en explicar por qué una de ellas no puede legítimamente aplicarse a una situación como ésta, en la que hay que elegir» (Nozick, 1999, p. 74).

¿Cuál es, entonces, la postura *ilegítima*?:

Los que se decantaron a favor de coger sólo la segunda caja intentaron por diversos medios socavar la fuerza del argumento de dominancia. Muchos señalaron que, si uno pensó en ese argumento y le convenció, el ser (casi con toda seguridad) lo habría vaticinado y habríamos acabado con 1000 dólares solamente. Interpretaron el argumento de dominancia como intento de burlar al adivino. Esta postura vuelve las cosas demasiado simples. El defensor del argumento de dominancia cree que acabará con 1000 dólares solamente, y sin embargo piensa que es mejor coger ambas cajas. Varios defensores del principio de dominancia se lamentan del hecho de que obtendrían peores resultados los individuos racionales que los irracionales, pero eso no los hizo cambiar (Nozick, 1999, p. 113).

Del pasaje citado, cabría hacer los siguientes comentarios: (i) que una postura simplifique las cosas (demasiado), no implica que la postura sea ilegítima; podría deberse a que no es necesario complejizar innecesariamente la situación, de modo que el *saijyayin* está adoptando una postura práctica; (ii) no por elegir la segunda caja optando por el principio de utilidad esperada, se es irracional (sobre esto volveremos más adelante); (iii) de igual modo, se podría sostener que el *maestro jedi* es irracional si se acepta que obtendría (tal vez en promedio) peores resultados e insistiera en elegir esa alternativa.

#### 4. ¿Hay *una* postura ilegítima?

Tal vez no tiene mucho sentido plantearse esta pregunta, pues entran en juego *dos* principios de elección distintos. Esto no tiene por qué ser un problema, si se tiene en mente que cuando se trata de decisión bajo incertidumbre, hay más de un criterio para analizar problemas de decisiones (Taha, 2012) que -al aplicarse- no implican idénticos resultados.

¿Qué podría jugar en contra del *saiyajin* sobre su elección? Tal vez el desconocimiento de las probabilidades. En efecto, el hecho de que el *saiyajin* desconozca las probabilidades para determinar la utilidad esperada, induce a pensar que -después de todo- no tiene un criterio que permita fundamentar su decisión, de modo que, ante esa limitante, lo que cabe hacer racionalmente, es optar por la estrategia de dominancia, lo cual implica elegir  $B_1 \wedge B_2$ , ya que dicha elección maximizaría la utilidad obtenida.

Pero en la práctica, las probabilidades no tienen por qué existir objetivamente. No se puede negar que -después de todo- la decisión se toma en un contexto de creencias que el decisor posee, y desde una perspectiva bayesiana se puede interpretar dichas creencias en términos probabilísticos (probabilidad doxástica).

Ciertamente, cuando el *saiyajin* decide, sabe (o cree que sabe) lo suficiente para tomar su decisión considerando:

- i. Que Profesor X ha hecho con frecuencia predicciones correctas (sobre lo que nosotros u otras personas han elegido);
- ii. Se desconoce que Profesor X haya hecho predicciones falsas;
- iii. Profesor X sabe que sabemos que sabemos...que si vaticina que cogeremos solo lo que hay en la segunda caja, mete los M dólares en dicha caja.

Nótese que (iii) puede interpretarse como expresando la idea de conocimiento público: *Todos saben que los demás (incluido Profesor X) saben que Profesor X mete el millón de dólares si vaticina que...* Llamemos a esta condición, la condición de conocimiento público: **CCP**.

Bajo (i) y (ii) el *saiyajin* no tiene razones para pensar que, en una situación cualquiera, Profesor X actuará violando **CCP**, es decir, el *saiyajin* estaría justificado en un sentido no deontológico a creer que ello es así. Por lo cual, puede perfectamente asumir que antes de decidir, Profesor X ha vaticinado qué elegirá, con una alta (tal vez absoluta) probabilidad de acierto. Esa probabilidad -como

indicamos- no tiene por qué existir *objetivamente*, basta con que el *saiyajin* crea en ella en un grado para actuar de acuerdo con esa creencia, y esto le garantiza que obtenga un beneficio de 1000000 de dólares.

Se puede objetar que en esas condiciones el *saiyajin* ha obtenido dicho premio por pura suerte. Pero esto es falso. De hecho, (i) y (ii) posibilitan que se pueda asumir que la probabilidad de predicción no es simplemente correcta desde el punto de vista inductivo, sino que es inductivamente fuerte; y en este escenario, mientras más fuerte sea la probabilidad (subjetiva), tanto más a favor del *saiyajin* para elegir la segunda caja. Este escenario se desdibuja cuando no hay razones para pensar que (i), (ii) y **CCP** ocurren; si este fuese el caso, entonces el *saiyajin* no tendría qué argumentar a favor de su decisión, pero esto no es lo que de hecho ocurre.

Con respecto al *maestro jedi*, la situación es distinta, porque pareciera estar actuando en un escenario donde (i)-(ii) y **CCP** no juegan rol alguno, sino que se focalizan en el hecho de que en la situación de elección en la que se encuentran, los estados posibles *ya* están determinados, es decir: *el millón de dólares está o no está en la caja*. En otras palabras, en su decisión el *maestro jedi* pareciera presuponer que ella es independiente de la acción realizada por Profesor X.

Es importante tener presente que la situación descrita en la paradoja discurre en al menos dos instantes; el primero, digamos  $t_i$ , en el cual Profesor X ha vaticinado y decidido si pone o no el millón de dólares; el segundo, digamos  $t_j$ , en el cual el *maestro jedi* toma la decisión. Nada de lo que haga el *maestro jedi* en  $t_j$  incidirá sobre la determinación de los posibles estados en  $t_i$ . Así, pues, habiéndose determinado los estados, una evaluación de la matriz de pago envuelta indica que la elección de las dos cajas domina la elección de solo la segunda, ya que en el mejor escenario supone una utilidad de 1001000 vs 1000000; y en el peor escenario, supone una utilidad de 1000 vs 0.

Debiéramos, por tanto, reconciliarnos con la idea de que el problema se puede abordar desde más de una perspectiva, y que ambas son correctas a partir del compromiso con ciertos presupuestos. Por ejemplo, se puede notar que, si nuestra confianza en que Profesor X acierta en sus predicciones es alta, entonces se debe optar solamente por la segunda caja; sin embargo, cuando ese nivel de confianza es bajo, entonces es conveniente elegir ambas cajas.

Por lo anterior, tal vez la cuestión no es tanto si hay una postura que es incorrecta. Como hemos indicado, las situaciones de elección pueden abordarse de más de una manera, y ello no implica que los resultados sean los mismos.



## 5. Relevancia filosófica de la paradoja

No estoy seguro de si toda paradoja es filosóficamente relevante, pero la de Newcomb sí que lo es, sobre todo por las interrogantes que plantea, al menos en el ámbito metafísico y epistemológico. Desde el punto de vista metafísico, por ejemplo, emerge el añejo problema determinismo-indeterminismo. Nozick cita las posturas de Mohan S. Kalehar, físico de los Laboratorios Nevis de la Universidad de Columbia, y de Isaac Asimov.

Según Kalehar,

Quizás sea falso decir que el ser ha elegido categóricamente una u otra opción, del mismo modo que es falso decir que el electrón [en el experimento de las dos ranuras] pasó por una ranura o la otra. Quizás lo único que podemos afirmar es que hay cierta amplitud de que B2 [la segunda caja] tenga el millón de dólares y otra cierta amplitud de que esté vacía. Estas amplitudes se interfieren a no ser que y hasta que decidamos actuar y abramos la caja...Afirmar que 'o bien B2 contiene un millón de dólares o si no es así, está vacía' es un argumento intuitivo que no puede probarse hasta que no abramos la caja. Hay que reconocer que el argumento intuitivo es poderoso, pero como en el caso de la difracción del electrón de ranura doble, nuestras intuiciones resultan a veces incorrectas. (Nozick, 1999, pp. 113-114).

Por su parte Asimov escribió:

Sin duda alguna, me llevaría ambas cajas...Soy determinista, pero me resulta perfectamente claro que cualquier ser humano merecedor de que se le considere tal (incluyéndome a mí, por supuesto) preferiría el libre albedrío, si tal cosa pudiese existir...Ahora bien, supongamos que uno se lleva ambas cajas y al final resulta (como muy probablemente así sea) que Dios lo ha previsto y no ha colocado nada adentro de la segunda caja. Uno habrá entonces, al menos, expresado su voluntad de apostar por la falta de omnisciencia de Dios y por el libre albedrío de uno mismo, y habría renunciado de buen grado a un millón, y habría renunciado a un millón de dólares a cambio de tal voluntad -que es en sí misma como un chasquido de los dedos en la cara del Todopoderoso, y un voto más que vano a favor del libre albedrío...Y, por supuesto, si Dios se ha embozado y ha dejado un millón de dólares dentro de la caja, entonces no solo habremos ganado ese millón sino, algo mucho más importante, habremos demostrado la no-omnisciencia de Dios. Sin embargo, si cogemos sólo la segunda caja conseguimos el dichoso millón; pero no solo somos esclavos, sino que también hemos demostrado que estamos dispuestos a ser unos esclavos

a cambio de ese millón, y no somos reconocibles como humanos. (Nozick, 1999, p. 119).

La observación de Kalehar puede interpretarse como una objeción al argumento del *maestro jedi*. Sin embargo, tal vez la analogía - independientemente de la sobresimplificación a la que alude Nozick- que se establece entre el experimento de la doble ranura y la caja no sea del todo adecuada; posiblemente se puede tener un mayor control de lo que ocurre en la caja que con lo que ocurre con el electrón en la ranura; por ejemplo, la caja puede ser transparente en una de sus caras de modo que sea visible a todos, menos a quien debe tomar la decisión. Al ver el contenido de la caja, estas personas darían fe, como si fuesen notarios, de si Profesor X ha colocado o no el millón en la caja. De esta manera, aunque el sujeto no pueda afirmar si el millón de dólares está en la caja o no, tampoco se podría afirmar que los estados *haber puesto el millón de dólares* o *no haber puesto el millón de dólares* no están determinados. Desde este punto de vista la especulación cuántica y la analogía con el gato de Schrödinger sobraría. De hecho, la situación sería la misma si en el caso del famoso gato, se introdujese la figura del observador notario, para el cual la conjetura de que el gato está vivo y muerto a la vez, no tendría sentido.

Por otro lado, Kalehar señala que nuestras intuiciones resultan a veces incorrectas, es decir: no siempre son correctas; la pregunta sería si en este caso las intuiciones del *maestro jedi* son efectivamente incorrectas. Aunque el *maestro jedi* no pueda especificar el estado decidido por Profesor X, desde una perspectiva superevaluacionista está claro que no hay alternativa a *se ha puesto un millón de dólares en la caja* o *no se ha puesto un millón de dólares en la caja*, y a partir de ahí aplicar la estrategia de dominancia.

Mientras que Kalehar presenta un argumento indeterminista para sugerir que la elección de las dos cajas podría cuestionarse, Asimov -a pesar de que admite ser determinista- indica que elegiría ambas cajas, pues cree que dicha elección es compatible con la libertad humana, lo cual podría implicar que en el mejor escenario la persona obtenga el millón de dólares y haya burlado la omnisciencia de Dios; o -en el peor escenario- que obtenga solo **mil** dólares, reafirmando su libertad.

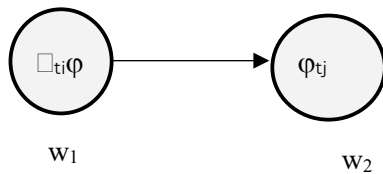
Asimov asume que, si se quiere defender la libertad, entonces la opción sería la alternativa no determinista, representada por la elección de ambas cajas. No está claro, sin embargo, que, al elegir solamente la segunda caja, la persona sea un esclavo, como afirma. Es decir, asumir una estrategia determinista no implicaría necesariamente la renuncia a la libertad, como si determinismo implicase inevitabilidad, y el sujeto no tuviese opciones reales al elegir [véase (Dennett, 2003)].

Esto, desde luego, no supone que el problema determinismo-indeterminismo no haya de considerarse, aunque -como afirmé- ello no tiene por qué incidir sobre el tema de la libertad.

La paradoja de Newcomb tiene que ver con la posibilidad de la causalidad invertida: «una decisión no causada causa una predicción anterior» (Nozick, 1999, p. 120).

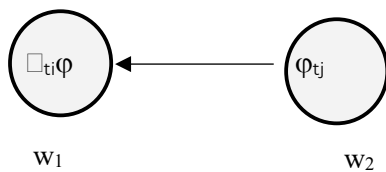
La situación es simple. Tenemos, por un lado, un instante  $t_i$  donde Profesor X vaticina (porque lo sabe) qué hará el sujeto en el instante  $t_j$ . Denote  $\varphi$  la decisión del sujeto, obtenemos, por un lado, que Profesor X vaticina en  $t_i$  (porque cree o sabe) que  $\varphi$ , denotado por  $\Box_{t_i}\varphi$ <sup>1</sup>, y -por el otro- que el sujeto actúa de acuerdo con  $\varphi$  en  $t_j$ , de modo que  $\varphi$  es verdadero; denotemos esto simplemente como:  $\varphi_{t_j}$ .

Supongamos que cada mundo posible  $w_k$  está en una relación de correspondencia con un instante  $t_k$ . Pensemos que entre cada mundo posible se da una relación de accesibilidad epistémica que indica las condiciones que permiten conocer. A partir de esta idea podemos obtener algo como sigue:



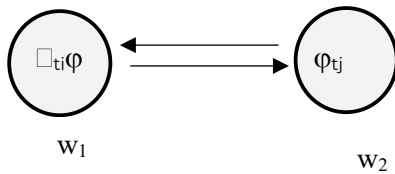
Este diagrama indica que al ser  $w_2$  accesible desde  $w_1$  no es posible que la predicción realizada sea verdadera en  $w_1$  y que el sujeto actúe contrario a ella.

Adicional, hay que considerar este escenario:



En este caso, al ser  $w_1$  accesible desde  $w_2$ , significa que la predicción realizada es verdadera. Lo anterior quiere decir que:

<sup>1</sup> Usamos « $\Box_{t_i}\varphi$ » con base en PCC, aunque nada impide que se use « $\Diamond_{t_i}\varphi$ », que lo cree.



En otras palabras, la relación entre  $w_1$  y  $w_2$  es simétrica. La simetría en cuestión quiere decir que, interpretado en el espacio modal adecuado, por ejemplo, en  $D$  o  $S_5$  se obtiene:  $\phi_{t_j} \rightarrow \Box_{t_i}\phi_{t_i}$ , es decir que si  $\phi$  es verdadero en  $t_j$ , entonces se sabe en  $t_i$  que  $\phi$  es creíble, y eso es lo que mínimamente Profesor X necesita para hacer su predicción.

Si lo anterior es correcto, entonces la idea de que la realidad es determinista resulta inevitable. Ahora, ¿supondría este hecho que la libertad es imposible? No necesariamente.

De acuerdo con Dennett [ (1995), (2003)], para entender adecuadamente el problema de la libertad es necesario cuidar nuestro lenguaje, y en ese sentido hay que considerar tres perspectivas: (i) la física (del físico o del químico), (ii) la del diseño (del biólogo o ingeniero genético) y (iii) la intencional (propia de los seres humanos).

Dennett -siguiendo a Van Inwagen- resume al determinismo en la siguiente tesis: «en cada momento dado hay exactamente un futuro posible» (Dennett, 2003), mientras que el indeterminismo sostiene que hay más de un futuro posible. El determinismo requiere la existencia de leyes universales, el indeterminismo no. De modo que el determinismo y el indeterminismo son conceptos físicos que nada dicen de la libertad; así las cosas, el determinismo y la libertad son compatibles ya que cada concepto pertenece a un lenguaje distinto, físico el primero, intencional el segundo.

Relacionado con dicho tópico, Dennett analiza también el problema del futuro. ¿Implica el determinismo suscribir que el carácter personal es en gran medida inmutable? Al respecto, distingue entre tener un futuro personal prefijado como algo distinto de tener una naturaleza personal prefijada. Nuevamente es asunto de perspectiva, desde el punto de vista de los microeventos los átomos de los que estoy formado son los mismos; en este sentido mi futuro es fijo. No obstante, en el nivel del diseño (e incluso intencional) el futuro puede ser muy variado, incluso para quien ha optado por elegir solamente la segunda caja.

## 6. Conclusión

Una pregunta crucial que surge a partir de la paradoja de Newcomb es: *¿qué hacemos con la información que tenemos?* Esta pregunta vista desde una

perspectiva bayesiana se responde de una manera muy sencilla. En un instante  $t_i$  poseemos una creencia  $C$ ; luego, en un instante  $t_j$  advertimos que nuestra creencia original  $C$  está condicionada por nuevos datos (o evidencia), así que dada la condicionalización en cuestión, se ajustan los grados de creencias con respecto a  $C$ .

Esto es importante para entender la conducta del *saiyajin* y del *maestro jedi*. En el primer caso, bajo el supuesto de PCC, supone para él una ganancia de un millón cada vez que *apuesta* contra Profesor X; de modo que, bajo ese supuesto, no tendría razón para cambiar su estrategia: elegir solo la segunda caja.

La situación del *maestro jedi* es distinta. Si tras  $n$  apuestas los resultados obtenidos no superan -probablemente- los mil dólares, entonces debería replantearse la estrategia: elegir solo la segunda caja. Es libre, en realidad, tanto como el *saiyajin* para hacerlo.

Finalmente, si elegir ambas cajas es la decisión racional, entonces no siempre una decisión racional implica los mejores (preferibles) resultados.

## **Bibliografía:**

Dennett, D., 1995. *La conciencia explicada*. Madrid: Paidós.

Dennett, D., 2003. *Freedom Evolves*. s.l.:Penguin Books.

Nozick, R., 1999. El problema de Newcomb y dos principios de elección. En: *Puzzles socráticos*. Madrid: Cátedra, pp. 73-109.

Nozick, R., 1999. Reflexiones sobre el problema de Newcomb. En: *Puzzles socráticos*. Madrid: Cátedra, pp. 111-122.

Rescher, N., 1969. *Essays in Honor of Carl G. Hempel*. s.l.:Reidel.

Taha, H. A., 2012. *Investigación de operaciones*. México: Prentice Hall.