

LA TEORÍA DE LA DECISIÓN DE DAVID LEWIS Y LA PARADOJA DE NEWCOMB

THE DECISION THEORY OF DAVID LEWIS AND THE NEWCOMB PARADOX

Esteban Céspedes*

Pontificia Universidad Católica de Valparaíso
Valparaíso-Chile

*Recibido 15 de noviembre 2008/Received november 15, 2008
Aceptado 20 de julio 2009/Accepted july 20, 2009*

RESUMEN

En el presente trabajo se realiza una exposición breve del punto de vista filosófico de la teoría de la decisión, profundizando específicamente en el trabajo que propone David Lewis al respecto. Este autor sustenta que la teoría de la decisión puede sostenerse sobre bases causales o no causales, siendo la Teoría Causal de la Decisión un modelo necesario para el análisis de este tipo de problemas. En casos particulares como en la Paradoja de Newcomb, las teorías no causales de la decisión serían insuficientes. A continuación se presentarán brevemente las nociones básicas del análisis de decisiones, para en segundo lugar explicar en detalle algunos aspectos de la Paradoja de Newcomb. Finalmente se presentan las tesis principales de la teoría de Lewis y algunas posibles objeciones de la misma.

Palabras Clave: Racionalidad, Decisión, Agente, Grado de Creencia, Valor Esperado.

* Salvador Vergara 274, Viña del Mar. Chile. E-mail: esteban.cespedes.b@mail.ucv.cl

ABSTRACT

In this work a brief account of the philosophical point of view on decision theory is shown by studying David Lewis' work in particular. This author states that a theory of decision can have either causal or non causal grounds and that causal decision theory is a necessary model for the analysis of this kind of problems. In particular, cases like Newcomb's Paradox, non causal decision theories would be insufficient. I will show some basic notions of decision analysis, in order to explain in detail some aspects of Newcomb's Paradox. Finally, the fundamental theses of Lewis's Theory and some possible objections are presented.

Key Words: *Rationality, Decision, Agent, Credence, Expected Value.*

1. INTRODUCCIÓN A LA TEORÍA DE LA DECISIÓN DE LEWIS

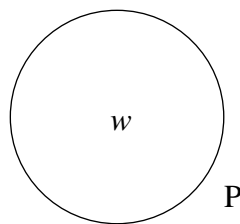
Considerada como una disciplina dentro de la teoría de juegos, la teoría de la decisión consiste en un tipo de análisis aplicable a diversas áreas como la sociología, la economía, las ciencias en general (especialmente en el método científico) y, por supuesto, la filosofía. Sin embargo, en esta última la teoría de la decisión no es solamente algo aplicable a la disciplina, sino que es también objeto de análisis y examen. Como primera definición habría que señalar que la teoría de la decisión tiene como finalidad analizar una situación determinada de modo que su resultado sea la acción más racional. El problema que ésta arroja es que hay cierta equivocidad en el término “racional”. Por esto es que existen definiciones que destacan los aspectos pragmáticos de la teoría, como el principio de Bayes, según el cual la mejor decisión (o la más racional) es “elegir el acto que se estime máximamente deseable” (p. 1).¹ Lo que sí es un hecho es que si una decisión es más racional que otra, depende

¹ Jeffrey, 1983, p. 1.

del modelo que se ocupe. Existen ciertos elementos importantes dentro de la teoría de la decisión que conviene tener presentes para comprender la estructura de estos modelos.

Un elemento muy importante en una decisión es el sujeto que decide, que será denominado “agente” de aquí en adelante. Otras dos nociones son fundamentales para que un análisis de decisión pueda considerarse completo: el grado de creencia² y el valor esperado.

- a) *El grado de creencia:* Todo agente posee ciertas convicciones con respecto a los hechos que considera relevantes para su decisión. De esta manera, sea w un mundo posible cualquiera. El grado de creencia de un sujeto acerca de que el mundo posible w es el mundo actual es una función proposicional de probabilidad de creencia y se escribe “ $C(w)$ ”. El grado de creencia, entonces, expresa a la vez el grado de convicción que posee un agente de que w es el mundo en que él habita. Ya que una proposición puede tomarse como un conjunto de mundos posibles, el grado de creencia se puede representar gráficamente mediante el diagrama de Euler, donde un conjunto es una proposición P cualquiera y w un elemento de esa proposición:



² Se hablará de aquí en adelante de “grado de creencia” como la traducción del vocablo inglés “credence”, que no tiene relación directa con la creencia (*belief*) acerca de una proposición. El uso habitual de la palabra “credence” se ocupa al “dar crédito a algo”.

De la misma manera, el grado de creencia es una función de probabilidad³ en un espacio cuyos puntos son mundos posibles y cuyas regiones son proposiciones.⁴ Además, la relación llamada “grado de creencia condicional” entre dos funciones quedará expresada de la siguiente manera. Sean X e Y dos proposiciones. El grado de creencia condicional de X dado que Y se anota “ $C(X | Y)$ ” y es igual a:

$$(i) C(X | Y) = C(X \cap Y) / C(Y)$$

$$(ii) C(X | Y) \cdot C(Y) = C(X \cap Y)$$

- b) *El valor esperado*: El grado de creencia estará siempre relacionado a un valor esperado, ya que si un agente otorga un grado de probabilidad al hecho de que w es el mundo actual, también podría asignar un valor a ese hecho. En base a la definición de David Lewis,⁵ podemos expresar esta noción de la siguiente manera. Sea $V(X)$ el valor esperado de una proposición X y $V(w)$ el valor esperado de un mundo posible. El valor esperado $V(X)$ es la suma de los grados de creencia condicional de cada w dado que X , multiplicado por el valor esperado de w :

$$(iii) V(X) = \sum_w C(w | X) \cdot V(w)$$

³ Una función de probabilidad es entendida como la relación entre las veces que ocurre un evento y un número real. Este número real indica la probabilidad de que dicho evento ocurra y debe satisfacer tres axiomas:

- 1) Para cualquier evento e , la probabilidad $P(e)$ debe ser mayor o igual que cero.
- 2) La probabilidad del total Ω , que representa todos los casos posibles, es igual a 1.
- 3) Si un evento está formado por varias alternativas que se excluyen entre sí, entonces la probabilidad de ese evento compuesto es igual a la suma de las probabilidades de las alternativas.

⁴ Lewis, 1986b, p. 267.

⁵ *Ibidem*, p. 306.

De (i) tenemos que:

$$(iv) V(X) = \sum_w C(w \cap X) / C(X) \cdot V(w)$$

Entonces el valor esperado de una proposición es igual a la suma de cada grado de creencia de un mundo posible y de X, dividido por el grado de creencia de X y multiplicado por el valor esperado de w .

Estas nociones fundamentales de la teoría de la decisión permitirán más adelante exponer la Teoría Causal de la Decisión de David Lewis. Según este autor, existen ciertos casos, como la Paradoja de Newcomb, en los cuales, si no se adopta una posición causal de la teoría, las soluciones del problema resultan ser irracionales. Para comprender este punto, se realizará ahora una explicación de la paradoja mencionada.

2. LA PARADOJA DE NEWCOMB

Como el objetivo fundamental aquí será realizar un examen de la Teoría Causal de la Decisión de Lewis, será necesario explicar de qué se trata la Paradoja de Newcomb, puesto que dependiendo de la interpretación que se adopte con respecto a la misma, la teoría deberá ser causal o no. Este autor afirma claramente que existe una distinción entre una teoría causal y una teoría no causal de la decisión, siendo esta última necesaria para el análisis decisonal.⁶ De tal forma, su tesis consiste en que una teoría de la decisión sería insostenible de no considerar los factores causales de un agente. El argumento principal que será utilizado para demostrar esta tesis es que en situaciones de decisión tales como la paradoja en cuestión, una teoría no causal tendría resultados irracionales.

⁶ Lewis 1986a, 305.

A) La formulación de Robert Nozick

La versión original nunca fue publicada por su creador, William Newcomb,⁷ sino por Robert Nozick.⁸ Ésta es presentada de la siguiente manera: Un agente deposita gran parte de su confianza en una especie de ente o programa de computador,⁹ cuyas capacidades predictivas son casi infalibles en cuanto a este tipo de situaciones. En casos de decisiones anteriores y similares al que será descrito, este programa no tiene ningún registro conocido de error. Si es en sí mismo infalible o si su infalibilidad es solamente una consideración del sujeto es un detalle en el que no se podrá profundizar ahora.

Además de esta máquina y del agente, existen dos cajas, una abierta y otra cerrada. La caja abierta contiene \$ 1.000 y la caja cerrada o bien tiene \$ 1.000.000 o bien nada. El contenido de ésta dependerá de la predicción del computador acerca de la decisión del sujeto. El agente tiene a su vez dos opciones: puede escoger quedarse con el contenido de las dos cajas o quedarse solamente con el contenido de la caja cerrada. En el siguiente cuadro, donde B1 simboliza a la caja abierta y B2 a la caja cerrada, se muestran las dos opciones del agente.

Caja	Valor
• B1	• 1.000
• B2	• 1.000.000
	• 0

⁷ El creador de la paradoja es profesor y físico teórico en el laboratorio Lawrence Livermore de la Universidad de California.

⁸ Nozick, 1969, pp. 114-116.

⁹ Aunque en la formulación clásica se habla de un ente (*being*), aquí se ejemplificará pensando en un programa. Lo que realmente importa es la definición de este ente y su capacidad casi infalible de predecir.

Si, por un lado, el agente se queda con el contenido de ambas cajas y su acción es prevista por el computador, entonces modificará el contenido de la caja cerrada y la dejará vacía. Si, por otro lado, el agente escoge solamente la caja cerrada, entonces el programa modificará su contenido para dejar \$ 1.000.000. Si esta información es sabida previamente por el sujeto, ¿qué opción debería tomar? En la descripción de Robert Nozick, el procedimiento ocurre en el siguiente orden:

- i) El computador realiza la predicción.
- ii) El computador determina el contenido de la caja cerrada.
- iii) El agente toma la decisión.

Para que esta situación constituya una paradoja, tendría que haber dos soluciones contrarias y mutuamente excluyentes, que sean igualmente racionales. De esta manera, hay dos argumentos, uno en favor de cada opción, que prueban que la opción favorecida es más racional que la otra.

Primer argumento: Si el programa predijo correctamente que el agente elegirá las dos cajas, la caja cerrada está vacía y el agente obtendrá sólo \$ 1.000. Pero si el programa predijo correctamente que el agente elegirá sólo la caja cerrada, entonces la caja cerrada contendrá \$ 1.000.000. Se concluye entonces que, bajo el supuesto de que el programa no puede errar en sus predicciones, lo más racional es que el agente opte por la sola caja cerrada.

Segundo argumento: Suponiendo que el programa ya ha predicho la acción del agente y también ha determinado el contenido de la caja cerrada, el millón está ahí o bien no lo está. Si, por una parte, el programa llenó la caja cerrada y el agente escoge ambas cajas, entonces el resultado es de \$ 1.001.000. Pero si se queda solamente con la caja cerrada, entonces su ganancia es de \$ 1.000.000. Por otra parte, si el programa dejó vacía la caja cerrada, entonces el sujeto obtendría \$ 0 y \$ 1.000 en cada caso, respectivamente. Haga lo que haga el sujeto, siempre obtendrá

\$ 1.000 más si escoge las dos cajas. Por consiguiente, lo más racional es quedarse con las dos.

B) La formulación de David Lewis

Este autor toma como base la formulación original de la paradoja, aunque ocupando solamente los elementos más generales y dejando aparte lo accidental que poseen otras formulaciones.¹⁰ El planteamiento de Lewis es así: Está el agente (elemento fundamental en la decisión), pero esta vez no juega el ente de gran capacidad predictiva. También se encuentra frente a él un beneficio de pequeña magnitud y sus opciones son tomarlo o dejarlo. Además, el agente podría sufrir un perjuicio de gran magnitud en comparación con el beneficio ya mencionado, lo cual dependería de su decisión y de un estado inicial. Dicho estado condicionaría tanto la acción del sujeto de elegir el beneficio como la consecuencia de que sufra el perjuicio. Según Lewis, no existe ninguna razón para no tomar el beneficio y, según él, tomarlo sería la decisión más racional. La razón de esto es que no existe nada que el sujeto pueda hacer para evitar que los hechos sucedan de esa manera. Según una teoría no causal de la decisión, la mejor opción es no tomar el beneficio. Se pueden expresar los valores expresados de cada una de las opciones de la siguiente manera. Sea “G” la opción de escoger el beneficio, “-G” la de no tomarlo y sea “E” el evento en que el agente padece el perjuicio, entonces se tiene que los valores esperados de no escoger el beneficio y de sí hacerlo son:

$$\begin{aligned} \text{i) } V(-G) &= C(E \mid -G) \cdot V(E-G) + C(-E \mid -G) \cdot V(-E-G) \\ \text{ii) } V(G) &= C(E \mid G) \cdot V(EG) + C(-E \mid G) \cdot V(-EG) \end{aligned}$$

Análisis no causal: Se pueden reemplazar los valores de las distintas funciones para resolver el problema. Si $C(E \mid G)$ posee un valor cercano a 1 (tomando en cuenta la definición del problema), entonces $C(-E \mid G)$ tendría que tener un valor cercano a

¹⁰ Lewis, 1986a, p. 309.

0, en el caso de (ii). De esta manera el valor esperado de tomar el beneficio será cercano al de $V(EG)$ y que obviamente no es algo muy positivo, ya que el perjuicio es mucho mayor en magnitud que el beneficio. En el caso de (i), $C(E | -G)$ tendría que tener un valor cercano a 0, debido a que no es muy probable que el perjuicio ocurra dado que el beneficio no fue tomado. Esto haría que el valor esperado de $V(-G)$ fuese similar al de $V(-E-G)$, lo cual no es ninguna ganancia, pero tampoco es una pérdida y es un resultado obviamente mucho mejor que el de $V(G)$. Se concluye que, en tal caso, la mejor opción es no tomar el beneficio.

Análisis causal: Sería bueno considerar el problema para valores intermedios de $C(E | G)$ y $C(-E | G)$, es decir, tomando un valor de 0,5 para cada uno. Un caso así es quizás el que consideraría la teoría causal de la decisión, ya que el grado de creencia de padecer el perjuicio es igual que el de no padecerlo, pues no es algo que depende causalmente de la acción del sujeto. En tal caso, $V(G)$ sería igual a la mitad de $V(EG)$ más la mitad de $V(-EG)$, lo cual entrega ganancias. Por otra parte, el valor esperado de $V(-G)$ sería la mitad de $V(E-G)$ más la mitad de $V(-E-G)$, cuyo valor no es positivo. La conclusión en este caso es que lo más racional sería optar por el beneficio de todas formas.

C) Valor esperado de la formulación clásica de la Paradoja de Newcomb

Se analizará la Paradoja de Newcomb en su versión original a partir de las fórmulas del valor esperado. Sean “B1” y “B2” la caja abierta y la caja vacía respectivamente:

$$\begin{aligned} \text{i.1)} \quad V_{(\text{caja cerrada})} &= C_{(\text{B2 llena})} \cdot V_{(\text{B2})} + C_{(\text{B2 vacía})} \cdot V_{(-\text{B2})} \\ \text{i.2)} \quad V_{(\text{caja cerrada})} &= C_{(\text{B2 llena})} \cdot 1.000.000 + C_{(\text{B2 vacía})} \cdot 0 \end{aligned}$$

$$\begin{aligned} \text{ii.1)} \quad V_{(\text{ambas cajas})} &= C_{(\text{B2 llena} | \text{tomar B1})} \cdot V_{(\text{B2B1})} + C_{(\text{B2 vacía} | \text{tomar B1})} \cdot V_{(-\text{B2B1})} \\ \text{ii.2)} \quad V_{(\text{ambas cajas})} &= C_{(\text{B2 llena} | \text{tomar B1})} \cdot 1.001.000 + C_{(\text{B2 vacía} | \text{tomar B1})} \cdot 1.000 \end{aligned}$$

Análisis no causal: En el caso de (i), por una parte, el grado de creencia de que la caja cerrada esté llena, es decir, $C_{(B2 \text{ llena})}$, debería ser de un valor cercano a 1, ya que existe *ex hypothesi* gran confianza en la predicción que haya realizado el programa. Si esto es así, entonces el valor de $V_{(caja \text{ cerrada})}$ debería ser muy cercano al valor de $V_{(B2)}$, lo cual es un muy buen resultado. En el caso de (ii), por otra parte, el grado de creencia de que la caja cerrada esté llena dado que se ha tomado también la caja abierta debería ser cercano a 0. En una situación así, el valor de $V_{(ambas \text{ cajas})}$ sería muy cercano al de $V_{(-B2B1)}$. Se concluye entonces que, como $V_{(caja \text{ cerrada})}$ tiene un valor aproximado de 1.000.000 y $V_{(ambas \text{ cajas})}$ es de un valor aproximado de 1.000, la opción de máximo valor esperado y, por ende, la más racional es la de tomar solamente la caja cerrada.

Análisis causal: Como se ha visto anteriormente, lo que hace la teoría causal de la decisión es equilibrar los grados de creencia, lo que tiene como consecuencia ignorar los factores que no dependen causalmente de las acciones del agente. En este caso tenemos que:

- i) $V_{(caja \text{ cerrada})} \approx 0,5 (1.000.000)$
- ii) $V_{(ambas \text{ cajas})} \approx 0,5 (1.002.000)$

Entonces, la opción más racional es tomar ambas cajas. Sin embargo, el análisis causal hace caso omiso de las hipótesis y condiciones que exige la Paradoja de Newcomb. En otras palabras, no está analizando una Paradoja de Newcomb. Este punto será examinado al final de este artículo. La Paradoja es analizada con mayor profundidad en los artículos de Isaac Levi (1978, pp. 369-383) y Doris Olin (1978, pp. 385-398).

D) Otras formulaciones de la Paradoja

Existen otras formulaciones similares a la Paradoja de Newcomb que servirán para comprender mejor cuál es el rasgo principal de los problemas de este tipo. En todos los casos similares a la Paradoja de Newcomb, el sujeto otorga un grado de creencia considerable para un factor que es determinante de dos cosas: a) que el sujeto se incline por un beneficio y b) que ese sujeto padezca un perjuicio de gran magnitud en comparación con el beneficio.

El fumador: En este caso, hay un agente que tiene dos opciones: dejar de fumar o seguir fumando. Al mismo tiempo, este fumador sufre de una lesión (arterial, por ejemplo) que provoca sus deseos de fumar y que a la vez podría ocasionarle una gran enfermedad. Al igual que los casos analizados anteriormente, un análisis no causal dirá finalmente que lo más racional es dejar de fumar. Sin embargo, para Lewis no existe una razón suficiente para que el agente deje de fumar (1986a, p. 311).

El dilema del prisionero: Ésta es explicada en detalle por Nozick (1993, p. 53) y Nigel Howard (1971, p. 75) y es una de las mejores formulaciones alternativas de la paradoja de la decisión. Eso sí, es un poco más compleja, ya que es un juego entre dos personas. Básicamente, el problema consiste en que hay dos prisioneros y cada uno tendrá dos opciones: confesar el delito o no confesar. Por un lado, si ninguno confiesa, entonces ambos reciben una pena de dos años, pero si ambos confiesan, reciben diez años. Por otro lado, si uno confiesa y el otro no, entonces el primero queda en libertad y el otro recibe doce años. Según un análisis formal, se obtendría que lo más racional es no confesar, ya que se asume que el otro prisionero tiene procesos mentales similares. No obstante, la teoría causal de la decisión diría tal vez que lo más racional es confesar, ya que el mejor resultado es quedar en libertad y no es necesario dar un grado de creencia

mayor a los factores que son independientes de las acciones de cada prisionero.

prisionero1/prisionero2	no confesar	confesar
• no confesar	• 2-2	• 12-0
• confesar	• 0-10	• 10-10

3. LAS DOS TESIS DE LA TEORÍA CAUSAL DE LA DECISIÓN

Hasta ahora se ha explicado que para cierto tipo de problemas de decisión existen diferentes respuestas posibles, lo que depende de si se adopta una teoría causal o no causal de la decisión. A continuación se verá por qué y mediante qué elementos es que la Teoría Causal de la Decisión de Lewis llega a dichos resultados. Existen dos tesis fundamentales de esta teoría; una de ellas es la que el autor llama “tesis principal” y la otra es la “tesis secundaria”.¹¹

Tesis principal: Una decisión es racional si y sólo si es aquella en la cual se maximiza la utilidad esperada de una acción calculada a partir de su hipótesis de dependencia.

Una *hipótesis de dependencia* está definida como un conocimiento completo de un agente con respecto a las cosas que dependen de sus acciones.¹² Es una proposición que describe exactamente cómo las cosas dependen causalmente de las acciones de un sujeto. Por una parte, esta tesis ocupará una nueva formulación, que reemplaza el valor esperado por la utilidad esperada y el grado de creencia condicional por el grado de creencia en una hipótesis de dependencia expresada como un condicional

¹¹ Lewis, 1986a, p. 333.

¹² *Ibidem*, p. 312.

contrafáctico. Estos últimos son proposiciones condicionales del tipo “si ocurriera p, entonces ocurriría q”, cuyo valor de verdad dependerá de los valores de verdad de su antecedente y su consecuente en determinados mundos posibles.¹³ Allan Gibbard y William Harper afirman, además, que las decisiones racionales pueden considerarse sin problemas a partir de un análisis lógico de proposiciones condicionales de este tipo.¹⁴ De hecho, estos autores han servido de base para la Teoría Causal de la Decisión de Lewis. La fórmula de la utilidad esperada se expresará de la siguiente manera. Sea “U(A)” la utilidad esperada de una proposición y “ $A \square \rightarrow S$ ” un condicional contrafáctico:

$$U(A) = \sum_s C(A \square \rightarrow S) \cdot V(AS)$$

Esta fórmula puede ocuparse solamente considerando las hipótesis de dependencia y la dependencia causal entre las acciones del agente y sus consecuencias. Debido a que no es exactamente una acción del agente, la influencia del computador es algo que deberá ser apartado del análisis que propone la teoría causal de la decisión. Ésta pone énfasis únicamente en el grado de creencia de las proposiciones que expresan relaciones de dependencia causal entre las acciones del sujeto y sus efectos. De esta manera, el teórico causal tomará su decisión otorgando un alto grado de creencia al condicional “si el agente tomara ambas cajas, entonces la segunda caja estaría llena”, lo cual parece disentir con los supuestos iniciales de la Paradoja de Newcomb. Al reducir la confianza en el programa de computador, la teoría causal de la decisión debe concluir que tomar ambas cajas es la decisión más racional.

Tesis secundaria: Una hipótesis de dependencia es exactamente la conjunción de patrones probabilísticos completos.

¹³ Lewis, 1973, p. 1.

¹⁴ Gibbard & Harper, 1978, p. 125.

Un *patrón probabilístico completo* se define como un conjunto que contiene una proposición contrafáctica por cada opción.¹⁵ De esta manera, una hipótesis de dependencia es en otras palabras un conjunto de condicionales contrafácticos. Queda claro que la tesis secundaria se deduce de la tesis principal.

4. DEBILIDADES DE LA TEORÍA CAUSAL DE LA DECISIÓN

Existen, al menos, dos clases de inconsistencias en esta teoría, que serán explicadas a continuación. Una de las objeciones es más bien un contraejemplo en el que se intenta demostrar, por medio de un problema tipo Paradoja de Newcomb, que la teoría causal no es suficiente o no es mejor que otros modelos de toma de decisiones. La otra consecuencia contraintuitiva es considerada como un error en la hipótesis, lo cual llevaría a una contradicción en las bases de la teoría.

El contraejemplo: Andy Egan presenta un problema parecido a la Paradoja de Newcomb en que el agente se encuentra en la siguiente situación.¹⁶ Entra un ladrón a su casa y tiene dos opciones: dispararle con su arma o no dispararle. Existe, al mismo tiempo, una condición (una descompensación química del cerebro, quizás) que lo motiva a disparar y que a la vez lo hace temblar y perder precisión en sus disparos. Ahora bien, si su disparo no falla, entonces el resultado es muy bueno, pero si llegase a fallar, entonces el resultado sería mucho peor que si no hubiese disparado. La teoría causal de Lewis obviaría el grado de creencia en la condición mencionada y consideraría que lo más racional es disparar. Egan afirma que claramente tal decisión no es racional, con lo cual reduce al absurdo la hipótesis de Lewis, suponiendo que es cierta y mostrando que

¹⁵ Lewis, 1986a, p. 332.

¹⁶ Egan, 2007, p. 97.

implica una contradicción. Por lo tanto, si la teoría causal de la decisión tiene como principio que en todos los problemas del tipo Paradoja de Newcomb la única decisión racional proviene únicamente de un análisis causal y, al mismo tiempo, existe un contraejemplo en el que la solución de un análisis causal no es la opción más racional en el contexto de la Paradoja de Newcomb, entonces existe claramente un error en la teoría. O bien, la teoría causal debería admitir que no funciona para toda situación de tipo Newcomb, o bien debería admitir que no es la mejor teoría de la decisión para ninguna Paradoja de Newcomb. Esto último podría tener la consecuencia de que la teoría causal no es sostenible para ningún tipo de decisión.

El error en la hipótesis: Esta objeción es quizás más simple que la anterior, pero al mismo tiempo va dirigida al fundamento mismo de la teoría. Básicamente, Lewis postula que *existe un caso y dicho caso es la Paradoja de Newcomb, en que la Teoría Causal de la Decisión es la única solución racional*. En segundo lugar, se toma un caso de la Paradoja de Newcomb y se resuelve quitándole lo que es propio de un problema de ese tipo, es decir, equiparando los grados de creencia, dejando a un lado los factores determinantes señalados anteriormente.¹⁷ No existen las paradojas de Newcomb para el defensor de la teoría causal, ya que no se consideran los factores que son causalmente independientes de las acciones del sujeto, lo cual significa equiparar todas las probabilidades. De hecho, para todo tipo de paradojas de Newcomb, las soluciones racionales estarían relacionadas con los análisis no causales. El problema de Newcomb original incluye en su definición un alto grado de creencia en la predicción del programa; de ahí la inestabilidad que existe en la teoría de Lewis, cuyos fundamentos corresponden a una hipótesis que no se respeta en la conclusión.

¹⁷ El hecho de que una teoría de la decisión no tome en consideración todos los factores que están en juego la convierte además en una teoría incompleta.

Lo que realmente postula la Teoría Causal de la Decisión es que existe un caso p , tal que p es una Paradoja de Newcomb y p es una solución racional si y sólo si p no es una Paradoja de Newcomb. Esto es una contradicción, si se quiere que tal caso posea una solución racional dentro de la teoría.

REFERENCIAS

- Egan, A. (2007). Some Counterexamples to Causal Decision Theory. *Philosophical Review* (116), 93-114.
- Gibbard, A. & Harper, W. (1978). Counterfactuals and Two Kinds of Expected Utility. En C. A., Hooker, J. J. Leach & E. F. McClennen (eds.), *Foundations and Applications of Decision Theory, I*. (pp. 125-162). Dordrecht, Holland: D. Reidel.
- Howard, N. (1971). *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. Massachusetts: MIT Press.
- Jeffrey, R. (1983). *The Logic of Decision*. Chicago: University of Chicago Press.
- Levi, I. (1978). Newcomb's Many Problems. En C. A. Hooker, J. J. Leach & E. F. McClennen (eds.), *Foundations and Applications of Decision Theory, I*. (pp. 369-383). Dordrecht, Holland: D. Reidel.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.
- Lewis, D. (1986a). Causal Decision Theory. En *Philosophical Papers (II)*. Oxford: Oxford University Press.
- Lewis, D. (1986b). A Subjectivist's Guide to Objective Chance. En *Philosophical Papers, II*. (pp. 83-132). Oxford: Oxford University Press.
- Nozick, R. (1969). Newcomb's Problem and Two Principles of Choice. En N. Rescher (ed.), *Essays in Honor of C. G. Hempel*. Dordrecht, Holland: D. Reidel.
- Nozick, R. (1993). *The Nature of Rationality*. Princeton: Princeton University Press.
- Olin, D. (1978). Newcomb's Problems, Dominance and Expected Utility. En C. A. Hooker, J. J. Leach & E. F. McClennen (eds.), *Foundations and Applications of Decision Theory (I)*. Dordrecht, Holland: D. Reidel.