

teorema

Vol. XXXIII/1, 2014, pp. 95-109

ISSN: 0210-1602

[BIBLID 0210-1602 (2014) 33:1; pp. 95-109]

EL ESTADO DE LA CUESTIÓN/STATE OF THE ART

Indeterminismo e inferencia causal

Iñaki San Pedro y Mauricio Suárez

RESUMEN

El indeterminismo supone un cambio conceptual en relación a nuestras intuiciones más robustas y mejor establecidas en torno a la noción de causalidad, relacionadas, en la mayor parte de los casos, con concepciones deterministas de la realidad, tal y como se refleja en teorías como la mecánica clásica. Los métodos de inferencia causal basados en la idea de causalidad probabilista, tales como el Principio de Causa Común (PCC), proporcionan un marco conceptual capaz de dar sentido y adaptarse a tales cambios. Aun así, persisten un buen número de problemas. Este ensayo ofrece una revisión de las principales implicaciones filosóficas en torno al PCC, así como de los argumentos más influyentes a favor y en contra de éste –tanto a nivel metafísico, como desde un punto de vista metodológico– en contextos indeterministas.

PALABRAS CLAVE: *causalidad; indeterminismo; principio de causa común; horquilla interactiva; extensibilidad.*

ABSTRACT

Indeterminism forces a conceptual shift concerning our most rooted intuitions about causation, which are mostly related to deterministic world-views, as reflected in scientific theories such classical physics. Methods of causal inference based on the idea of probabilistic causality, such as the Principle of the Common Cause (PCC), provide a suitable framework to cope with that shift. A number of problems however still remain. We offer here an overview of the main philosophical implications as well as the diverse arguments for and against the PCC – both at the metaphysical and the methodological level – in indeterministic contexts.

KEYWORDS: *Causation; Indeterminism; Principle of the Common Cause; Interactive Fork; Extensibility.*

I. INTRODUCCIÓN

La consideración de escenarios indeterministas introduce cambios sustanciales respecto a nuestras intuiciones más robustas y mejor estableci-

das sobre la causalidad, cuyo origen se encuentra, en la mayor parte de los casos, asociado a concepciones deterministas de la realidad, implícitas, por ejemplo, en teorías como la mecánica clásica. Los métodos de inferencia causal basados en la idea de causalidad probabilista, tales como el Principio de Causa Común (PCC), proporcionan un nuevo marco conceptual capaz de dar sentido y adaptarse a tales cambios. Aun así, persisten un buen número de problemas.

Este ensayo tiene como objeto ofrecer una revisión de algunas de las principales cuestiones filosóficas a que este tipo herramientas de inferencia causal, y en concreto el Principio de Causa Común (PCC), se enfrenta, así como de los argumentos más influyentes a favor y en contra de éste último en contextos indeterministas. Hacemos hincapié aquí, además, en la importancia de la distinción entre las componentes metafísica, por un lado, y metodológica, por otro, del PCC que, a nuestro modo de ver, no siempre ha sido lo suficientemente apreciada. Algunos de los problemas planteados en relación con la validez o aplicabilidad del PCC, de hecho, pierden fuerza a la luz de dicha distinción. Así mismo, el presente ensayo pretende poner en perspectiva la extensa literatura generada en torno al estatus filosófico del PCC a la luz de algunos avances recientes en inferencia causal que incluyen, por ejemplo, el tratamiento matemático-formal del PCC (por parte de la llamada *Escuela de Budapest*).

El ensayo comienza con una breve revisión del impacto que la consideración de contextos indeterministas ha tenido sobre nuestras intuiciones en relación con la causalidad y, por tanto, sobre las diferentes propuestas causales basadas en éstas. (Se describen aquí, por ejemplo, las diferentes nociones de indeterminismo que es posible considerar y las implicaciones de cada una.) El Principio de Causa Común (PCC) surge, precisamente, en un intento de reconciliar las nociones causales que podríamos llamar más “tradicionales”, ligadas a una concepción determinista de la naturaleza, con las variaciones conceptuales, básicamente en forma de consideraciones probabilísticas, que el indeterminismo demanda. El ensayo dedica su segunda parte a este tipo de consideraciones. Señalamos, así mismo en esta segunda parte, la importancia de la ya mencionada distinción entre las componentes metafísica y metodológica del PCC. Esta distinción es fundamental, desde nuestro punto de vista, a la hora de evaluar el estatus filosófico del principio, cuestión que abordamos en la tercera y última parte del ensayo, en que se exponen los argumentos más influyentes a este respecto, así como algunos de los avances más recientes.

II. DETERMINISMO, INDETERMINISMO Y CAUSALIDAD

Teorías como la mecánica clásica de Newton, o la teoría general de la relatividad de Einstein, proporcionan una visión completa y exacta de cómo los diferentes eventos se suceden unos a otros (de una manera secuencial). En con-

creto, según establecen estas teorías, ciertos eventos se siguen a partir de otros, de manera invariable e ineludible. Resulta posible entonces, e incluso adecuado, asociar a este tipo de secuencias, detalladas y continuas, una estructura en forma de cadenas causales que representen, a modo de “hilo conductor” por así decirlo, las relaciones mismas entre unos eventos y otros. Una estructura capaz de dar cuenta incluso de qué hay tras estas sucesiones, de por qué se dan unas secuencias y no otras, etc. El resultado es, obviamente, una visión del mundo articulada en torno al orden causal que subyace en la teoría, formado por una suerte de “red” de cadenas causales, cada una de las cuales es posible identificar, al menos en principio, de forma precisa y detallada¹. El establecimiento de cadenas causales de este tipo proporciona además un modelo explicativo preciso y fiable, en tanto que capaz de identificar inequívocamente el origen de cada evento observado haciendo referencia a su causa anterior².

La estructura causal que acabamos de describir es, obviamente, determinista. La pregunta surge, por tanto, si es posible en este contexto hablar de *indeterminismo*. Y si lo es, ¿en qué sentido es así? Es decir, ¿cómo puede tener cabida dentro de esta idea de causalidad determinista un comportamiento indeterminista? Para poder responder a estas preguntas es necesario distinguir, antes que nada, entre dos tipos de indeterminismo fundamentalmente diferentes.

En primer lugar, es fácil ver que, a pesar de disponer de una teoría que, como en el caso de la mecánica clásica, proporciona una visión causal determinista en la que los eventos se suceden unos a otros de forma precisa e invariable —a través de las ecuaciones dinámicas correspondientes—, esto no es suficiente para poder establecer una determinada cadena causal asociada a una secuencia de eventos específicos. Para ello es necesario además proporcionar las condiciones iniciales (o condiciones de contorno) a partir de las cuales se aplican las ecuaciones en cuestión. En otras palabras, si bien son las ecuaciones (deterministas) de una teoría como la mecánica clásica las que permiten una interpretación causal de ésta (en términos de cadenas causales formadas por sucesiones invariables de eventos), su aplicación a casos particulares requiere establecer un conjunto de condiciones iniciales de partida que, por decirlo de alguna manera, permitan poner en marcha el “algoritmo determinista” que proporcionan dichas ecuaciones. Es en nuestra (in)capacidad de especificar de forma más o menos exacta este conjunto de condiciones iniciales donde podemos identificar el primer tipo de *indeterminismo*. En concreto, la imposibilidad de conocer *con total exactitud* el estado de un sistema en el momento de poner a funcionar la teoría (mediante la aplicación de las ecuaciones correspondientes) introduce una cierta indeterminación en las soluciones que ésta ofrece. Un ejemplo habitual de este tipo de indeterminismo es el asociado al lanzamiento de dados, o monedas al aire. Dadas las ecuaciones de la mecánica clásica que gobiernan el movimiento de un sólido rígido, la incertidumbre en el resultado del lanzamiento de una moneda (cara o cruz al 50%) se debe simplemente al hecho de que no somos capaces de especificar

con total precisión el estado de la moneda en el momento del lanzamiento (ángulo de lanzamiento e impulso, masa de la moneda y su forma exacta, altura desde la que se lanza la moneda y altura a la que cae, etc.).

Este primer tipo de indeterminismo está, por tanto, asociado a una falta de información –a una ignorancia– acerca del estado inicial de un sistema. Esta postura presupone claramente que de ser capaces de tener acceso epistémico a todos los detalles de un sistema físico en un momento determinado, el indeterminismo desaparecería. Este sería, por ejemplo, el caso del famoso *demonio de Laplace*, cuya capacidad de acceso al estado de un sistema en todo detalle le permite predecir (o “retrodecir”) cualquier evento que pudiera suceder (o que haya sucedido). Dicho de otro modo, bajo este planteamiento, el mundo sigue siendo fundamentalmente determinista –como por cierto muestran las teorías a través de sus ecuaciones– a pesar de no poder predecir el resultado exacto de algunos eventos, siempre debido a nuestra incapacidad para acceder a *toda* la información asociada al estado de un sistema en un momento determinado. Este tipo de indeterminismo es por tanto de origen completamente epistémico, reducible (al menos teóricamente), y en absoluto fundamental. Y es del todo compatible, evidentemente, con una concepción determinista de la causalidad.

Existe, sin embargo, otra noción de indeterminismo bastante más problemática para la concepción determinista. Surge del hecho de que parece imposible en algunos casos predecir con exactitud la evolución de un sistema –o lo que es lo mismo, el resultado de un experimento–, al margen de nuestras capacidades para el acceso epistémico a la información asociada a éste. En estos casos no sería posible por tanto establecer una secuencia de eventos, a modo de cadena causal, exacta o bien determinada, a pesar de que dispusiéramos de *toda* la información sobre el sistema en cuestión. Este tipo de indeterminismo es el que habitualmente se tiene por característico de la mecánica cuántica. En su interpretación más extendida –la llamada “interpretación ortodoxa”– la mecánica cuántica asocia a todo sistema físico un indeterminismo fundamental inherente al sistema en cuestión. Un ejemplo típico de este comportamiento indeterminista es la desintegración de isótopos radiactivos. En estos casos la teoría no permite predecir con exactitud el momento en que el núcleo de un átomo inestable (un isótopo), pongamos de Tritio (^3H), acabará por desintegrarse como consecuencia de emisiones radiactivas. Lo que sí proporciona la mecánica cuántica es una estimación probabilística de la desintegración del isótopo, mediante su llamada “vida media” (unos 12,3 años en el caso del Tritio). En otras palabras el proceso de desintegración de un átomo por emisión de radiación es un proceso fundamentalmente estocástico, o indeterminista.

Este tipo de indeterminismo no tiene, por tanto, su origen en nuestras limitaciones epistémicas, la falta de información con relación a las condiciones iniciales, o las imprecisiones de nuestros aparatos y métodos de medición. Podríamos decir que es éste un *indeterminismo genuino*, en ningún modo “reducible” en el sentido anterior. Más bien al contrario, el hecho de

que algunos sistemas físicos presenten este tipo de indeterminismo irreducible apunta a una visión del mundo radicalmente diferente a la descrita anteriormente y, hasta cierto punto, irreconciliable³.

Se hace necesaria por tanto una revisión del concepto de causalidad de forma que sea posible adecuarlo a contextos genuinamente indeterministas como el que acabamos de describir. Desde luego, no parece apropiado ya apelar a cadenas causales, o a secuencias de eventos que se siguen unos a otros de forma invariable. Necesitamos, más bien, replantearnos en la medida de lo posible la relación entre causalidad y determinismo, y revisar nuestras intuiciones con el fin de reformular la noción de causalidad de forma que dé cabida a eventos de genuino carácter indeterminista. En concreto, este nuevo concepto de causalidad debe permitir situaciones como que la presencia de una causa X no garantice la ocurrencia de su efecto Y . Dicho de otra manera, con indeterminismo de por medio, las relaciones causales no pueden demandar la presencia de la causa para la ocurrencia del efecto. Y deben permitir, así mismo, que de la ocurrencia de una causa no se siga invariablemente la ocurrencia del efecto correspondiente. La presencia de una causa por tanto no es necesaria y, a diferencia del caso determinista, tampoco es suficiente para que se dé el efecto correspondiente⁴. Consideraciones como las anteriores dan lugar a la idea de *causalidad probabilista*. Muy sucintamente, en el marco de la causalidad probabilista, la relación causal entre un evento X y otro evento Y se refleja en una dependencia probabilística entre ambos, es decir la probabilidad de que ocurra el evento X es básicamente diferente en función de si el evento Y también ocurre: $p(X|Y) \neq p(X)$. (Donde $p(X|Y)$ representa la probabilidad condicional de que el evento X ocurra dado que el evento Y también lo hace.)

A la luz de este nuevo concepto de causalidad aparecen diversas propuestas de inferencia y descubrimiento causal en las que el concepto de probabilidad –y el cálculo de probabilidades– desempeñan un papel fundamental. El Principio de Causa Común (PCC), propuesto originalmente por Hans Reichenbach, es una de las más significativas.

III. EL PRINCIPIO DE CAUSA COMÚN DE REICHENBACH

Hans Reichenbach introduce el Principio de Causa Común (PCC) por vez primera en su libro *The Direction of Time*, en 1956. En origen, la formulación del PCC no estaba dirigida a la inferencia causal *per se*. Reichenbach trataba de derivar más bien, apelando al orden causal que el principio propone, la dirección del tiempo. El cometido del principio, en otras palabras, era resolver el problema de la flecha del tiempo, apelando, eso sí, a consideraciones causales. Fue este “contenido causal” del principio, sin embargo, lo que acabó perdurando, quizás debido a la evidente fuerza intuitiva de éste, quizá por su pretendido poder explicativo. De forma muy resumida el Princi-

pio de Causa Común establece que *toda correlación tiene una explicación causal*, ya sea en términos de una relación causal directa entre los eventos que se hallan correlacionados, o en términos de una *causa común* a ambos. El conocido ejemplo del barómetro, cuyas lecturas están en correlación con la aparición o no de diferentes fenómenos meteorológicos (cielos claros y despejados, con nubes, viento, lluvia, etc.) es útil para ilustrar el significado de la idea de causa común. Evidentemente, y a pesar de la correlación observada, no es el barómetro el que, con su funcionamiento, causa la aparición de los diferentes fenómenos atmosféricos. Diríamos más bien que, tanto las lecturas del barómetro como la aparición de dichos fenómenos se deben a una *causa común a ambos*, como es el valor de la presión atmosférica. De otro modo, la correlación entre las lecturas de un barómetro y la aparición de ciertos fenómenos atmosféricos se explica a partir de las variaciones de la presión atmosférica, que actúa como causa común.

Por tanto, según el PCC, hallada una correlación entre dos eventos — entre dos tipos de evento más bien—, si ésta no responde a una relación causal (directa) entre ellos, entonces debemos suponer que *existe* una causa común a ambos (que además explicaría la correlación misma). En este sentido, el PCC muestra un claro compromiso de carácter metafísico (por el cual se postulan las causas comunes *inicialmente no observadas*) que, de acuerdo con la terminología de Suárez, (2007), denominamos *Postulado de Causa Común* (PosCC). Reichenbach, sin embargo, no sólo apunta a consideraciones de carácter metafísico como las expresadas por el PosCC, sino que proporciona además un criterio específico para identificar, o caracterizar, las causas comunes postuladas mediante las siguientes relaciones probabilísticas:

$$(1) p(A \& B|C) = p(A|C) p(B|C),$$

$$(2) p(A \& B|-C) = p(A|-C) p(B|-C),$$

$$(3) p(A|C) > p(A|-C),$$

$$(4) p(B|C) > p(B|-C),$$

donde A , B y C son *eventos tipo*, “ $A \& B$ ” representa el evento tipo “el evento A y el evento B tienen lugar de forma conjunta”, y “ $-C$ ” es el complemento de C , que representa la no ocurrencia del evento C . De nuevo, $p(X|Y)$ representa la probabilidad condicional del evento X dado Y .

Estas cuatro relaciones constituyen por tanto un *Criterio de Causa Común* (CritCC), que establece las condiciones bajo las que un evento se considera una causa común de otros. En concreto, según el CritCC, las causas comunes deben, antes que nada, poder ser consideradas causas (en algún sentido) de los eventos correlacionados. A este respecto el criterio opera bajo una noción de causalidad débil por la cual un evento X puede ser considerado

causa de otro Y si la probabilidad de este último aumenta con la presencia del primero, cuando se mantienen fijas todas las otras posibles causas de Y . La relevancia causal de la causa común C con respecto a los eventos A y B se refleja obviamente en las dos últimas relaciones, (3) y (4): la presencia (ausencia) de C es relevante para la ocurrencia tanto de A como de B , es decir, la presencia de C aumenta la probabilidad de que A ocurra (y de igual modo para B). Por otro lado, y siempre según CritCC, la causa común debe ser un evento tal que, una vez tomado en consideración, la correlación original desaparece. (Decimos, en concreto, que la causa común “apantalla” la correlación.) Esto es de hecho lo que expresan las dos primeras ecuaciones, conocidas como *condiciones de apantallamiento* (*screening-off conditions*): tanto en (1) como en (2) la presencia de la causa común C hace que la correlación entre A y B desaparezca, es decir, dada la causa común C , los eventos A y B son probabilísticamente independientes.

El concepto de causalidad asociado al PCC se puede encontrar en una serie de programas de inferencia causal que persiguen, todos ellos, establecer relaciones causales a partir de ciertas relaciones probabilísticas. Es decir, persiguen descifrar el contenido causal que supuestamente asociamos a la observación de correlaciones.

Desde el punto de vista filosófico, la importancia de estos métodos de inferencia causal es tanto a un nivel conceptual como desde una perspectiva metodológica o de aplicación. Por un lado, el contenido conceptual de métodos de inferencia causal como la Condición Causal de Markov –que consiste en una generalización del PCC bajo unas condiciones determinadas–, aquellos que hacen uso de las llamadas redes Bayesianas, o el mismo PCC, permiten profundizar en los análisis de la noción de causalidad y su posible aplicación a contextos genuinamente indeterministas⁵. Por otro lado, a un nivel más práctico, el amplio rango de aplicación que de hecho encontramos de estas técnicas de inferencia en diferentes ámbitos científicos da una idea de su importancia. No sólo la física hace uso de modelos basados en tales métodos de inferencia causal –como puede ser el uso, tan extendido, de simulaciones Monte Carlo, por ejemplo, para la modelización, diseño o calibración de ciertos experimentos–. Otras disciplinas como la medicina, la sociología, la psicología, la economía o la biología necesitan recurrir a ellos con objeto de evaluar el impacto de éste o aquel componente químico de un medicamento sobre una enfermedad, la influencia de ésta o aquella condición social en los rasgos o características observados de una población, la relación entre ciertos estados emocionales y mentales del individuo, las consecuencias evolutivas sobre una especie determinada a partir de la presencia de una cierta carga genética, etc.

Sin embargo, y a pesar de su impacto en los programas de inferencia causal, el PCC sufre de varios problemas en relación con su estatus filosófico. En concreto, y aunque por diferentes motivos, tanto el Postulado de Causa Común (PosCC) como el CritCC están expuestos a contraejemplos que

siembran dudas sobre su aplicabilidad, e incluso su validez. Por un lado, el compromiso ontológico que expresa el PosCC quedaría en entredicho si observáramos correlaciones entre dos eventos A y B causalmente no relacionados de forma directa –es decir, casos en que ni A es causa de B , ni éste lo es del primero– cuya explicación no necesitara de causas comunes. De existir este tipo de casos, por tanto, en que una correlación entre dos eventos no respondiera a una relación causal (directa) entre ellos, ni a la existencia de una causa común a ambos, estaríamos ante claros contraejemplos del PosCC (y por tanto, más generalmente, del PCC). Esta es la situación que explota un famoso ejemplo propuesto por Elliott Sober en que los “niveles de mar en Venecia” y los “precios de pan en Gran Bretaña” se encuentran correlacionados, sin presentar relación causal alguna entre ellos, ni directa ni en forma de causa común [Sober (1997),(2001)].

En cuanto al CritCC los problemas surgen del hecho de que éste no parece proporcionar condiciones ni necesarias ni suficientes para poder hablar de causas comunes. Debemos preguntarnos, por tanto, bajo qué condiciones es posible su aplicación. El caso de la adecuación o aplicabilidad del CritCC para la inferencia causal ha sido ampliamente debatido, lo que ha dado lugar a una extensa literatura en la que el consenso más generalizado es que el CritCC no puede ser tomado como condición *suficiente* (ni tampoco necesaria, por cierto) para la identificación de causas comunes. Los argumentos más extendidos en este sentido apelan a casos concretos en los que ciertos eventos, a pesar de satisfacer las condiciones (1)-(4) impuestas por el CritCC, no pueden ser considerados causas comunes. Un ejemplo sencillo de este tipo de situaciones es la observación de un efecto común a dos eventos correlacionados⁶.

A un lado la posibilidad de aplicar el CritCC como condición suficiente para la identificación de causas comunes, nos queda la opción de proceder a la inversa, por eliminación, descartando como causa común todo aquel evento que no satisfaga las condiciones impuestas por el criterio. Para ello es necesario mostrar que el CritCC constituye una condición *necesaria* para la noción de causa común. Dicho de otra forma, sólo si consideramos que una causa común C debe satisfacer necesariamente las relaciones (1)-(4), podemos concluir, de la violación de alguna de ellas por parte de un evento X , que el evento en cuestión *no es una causa común*. La fuerza operativa del principio claramente se debilita en este caso aunque a pesar de ello, por otro lado, podríamos seguir apelando al PCC como un método útil para la inferencia causal. La pregunta es, por tanto, ¿puede considerarse el CritCC una condición necesaria para la caracterización de causas comunes?

Las diferentes respuestas que se han dado en la literatura a esta pregunta proporcionan un interesante mapa de las diferentes actitudes en torno al significado, fiabilidad metodológica o capacidad predictiva de una variedad de programas de inferencia causal que hacen uso, ya sea directa o indirectamente, de las intuiciones contenidas en el PCC. Y es precisamente en contex-

tos de indeterminismo genuino donde se dirimen estas cuestiones, y donde las diferentes posturas muestran sus diferencias más notables⁷.

IV. EL PCC EN CONTEXTOS INDETERMINISTAS

Consideremos el ejemplo, propuesto originalmente por Bas van Fraassen, en que un átomo inestable se desintegra dando lugar a dos subproductos como resultado. Si suponemos la conservación del momento, los dos subproductos resultantes de la desintegración se dispersarán de modo que el momento lineal resultante iguale al momento inicial del sistema (obviamente nulo si el átomo se encuentra inicialmente en reposo). En concreto, debido a la conservación del momento, los dos subproductos seguirán trayectorias opuestas⁸. La ley de conservación del momento nos permite, por tanto, a partir del ángulo de dispersión a de uno de los subproductos, inferir de forma *inequívoca* el ángulo de dispersión b del otro.

Supongamos además, que son varios los posibles ángulos de dispersión de cada uno de los subproductos, y que conocemos únicamente su probabilidad. Pongamos por simplicidad que son sólo dos las trayectorias posibles que cada uno de los subproductos seguirá en cada caso –a lo largo de los ángulos a y a' para el primer subproducto, y a lo largo de $b = a + \pi$ y $b' = a' + \pi$ para el segundo de ellos–, y que cada una de estas trayectorias se da con probabilidad del 50%. Sabemos, por tanto, que $p(a) = p(a') = 1/2 = p(b) = p(b')$.

Bajo estas condiciones es posible observar una correlación entre los ángulos de dispersión de los dos subproductos. Es decir $p(a \& b) - p(a)p(b) \neq 0$. Es más, gracias a la ley de conservación, la correlación es *perfecta*, ya que $p(a \& b) = p(a) = p(b)$. (Esto es de hecho lo que quiere decir que podemos determinar el ángulo de dispersión de uno de los subproductos una vez conocido el del otro.) La pregunta es, por tanto, dado que el hecho de que uno de los subproductos se disperse con un ángulo a no parece ser la causa de que el otro lo haga con ángulo $b = a + \pi$, ¿podemos explicar la correlación apelando a una causa común C , tal y como propone el PCC? Y, en concreto, suponiendo que exista una causa común C ¿satisfará ésta los requisitos impuestos por el CritCC?

Esta pregunta tiene una respuesta obviamente positiva en caso de que la causa común postulada, siguiendo el PCC, sea *determinista*. Es decir, si C es tal que su presencia *determina completamente* (con probabilidad 1) uno de los ángulos de dispersión de los subproductos: $p(a|C) = p(b|C) = 1$. En tal caso, como es fácil ver, las condiciones (1)-(4) impuestas por el CritCC quedarían satisfechas.

Si postulamos, en cambio, la causa común C como un evento puramente (o genuinamente) indeterminista, el asunto se complica. Pongamos, por ejemplo, que la presencia de C fija en r la probabilidad de que el primer subproducto se disperse con ángulo a : $p(a|C) = r$. Si esto es así podemos saber

con toda certeza, dada la ley de conservación del momento, que la probabilidad de que el ángulo de dispersión b del segundo subproducto es también r : $p(b|C) = r$. Esto implica, por tanto que, dada la presencia de C , el evento conjunto “ a & b ” tiene también probabilidad r , es decir $p(a \& b|C) = r$. Volviendo al CritCC, vemos que la situación que acabamos de describir viola las relaciones de “apantallamiento” (1) y (2). Es decir, la *causa común* postulada en este caso *no satisface* el Criterio de Causa Común (CritCC).

Ejemplos como éste se hallan en el origen de los argumentos más convincentes en contra de la idea de que el CritCC deba considerarse como una condición necesaria para la identificación, o caracterización, de la noción de causa común, al menos en lo que respecta a contextos genuinamente indeterministas. Encontramos, sin embargo, diferentes reacciones ante este tipo de casos, asociadas a las diversas posturas en relación al significado filosófico del PCC, y más generalmente a la idea de causa común.

En primer lugar, la conclusión quizá más radical que se puede extraer de tales ejemplos pasa por rechazar por completo el Principio de Causa Común, tanto en su propuesta metafísica –el PosCC–, como a nivel metodológico –el CritCC–. Esta es de hecho la postura que el mismo van Fraassen defiende. El a la postre muy influyente argumento que encontramos en su “Charybdis of Realism...” [van Fraassen (1982a)] es un buen ejemplo de las implicaciones sobre el estatus del PCC que, según van Fraassen, tiene el considerar sistemas genuinamente indeterministas. Muy brevemente, en este ensayo se identifica por vez primera la idea de causa común contenida en el PCC con la de *variable oculta*, que aparece en la derivación de las desigualdades de Bell en el contexto de las correlaciones Einstein-Podolsky-Rosen (EPR). Como es bien sabido, el teorema de Bell muestra que no es posible construir teorías de variables ocultas locales –como las que en opinión de van Fraassen sugieren las causas comunes postuladas por el PCC– capaces de dar cuenta de las correlaciones entre ciertas propiedades (como el espín, o el estado de polarización) propias de partículas distantes que forman parte de un estado cuántico entrelazado –estas son precisamente las llamadas correlaciones EPR–. Correlaciones que, por cierto, no sólo predice correctamente la mecánica cuántica, sino que han sido testadas experimentalmente con cierto éxito. La conclusión de van Fraassen es que estas correlaciones no responden a estructura causal alguna y que el PCC, por tanto, debe ser abandonado por completo. En resumen, a partir del hecho de que las causas comunes *deban* satisfacer el criterio sugerido por Reichenbach –es decir, aceptando el CritCC–, van Fraassen concluye que la idea de causa común no es adecuada para explicar ciertas correlaciones. Esto es sin duda suficiente para rechazar el compromiso metafísico que supone el Postulado de Causa Común (PosCC).

En segundo lugar encontramos propuestas menos radicales que, si bien sugieren que el PCC, tal y como lo define Reichenbach, no es aplicable a contextos indeterministas, sí retienen la idea de causa común como válida y

útil para el descubrimiento causal. Este tipo de propuestas, por lo tanto, centrarán sus esfuerzos en reformular en la medida de lo posible el criterio para la caracterización de causas comunes –rechazarán que éstas deban ser caracterizadas mediante el CritCC–, si bien mantendrán intacto el PosCC. El objeto de tal reformulación es claramente el ofrecer cabida a casos como el del ejemplo anterior, lo que se conseguirá al relajar los requisitos (necesarios) impuestos sobre la idea de causa común. Es decir, se trata de debilitar el CritCC con el fin de ampliar la aplicación del criterio resultante a contextos genuinamente indeterministas. Los ejemplos más relevantes de esta segunda vía son las llamadas *horquillas interactivas* (*interactive forks*) de Wesley Salmon por una parte, y las *horquillas generalizadas* (*Generalised Fork Criterion*) de Nancy Cartwright por otra⁹. Así, tanto Salmon como Cartwright no ven mayor dificultad en ejemplo del átomo que hemos visto anteriormente. En efecto, dado un sistema (como el átomo desintegrándose) que responde a un comportamiento causal genuinamente indeterminista, y que sugiere la existencia de una causa común *C* (así mismo indeterminista), nada hace suponer que debamos exigir que ésta satisfaga las condiciones impuestas por el CritCC. Más bien al contrario, ya que, según Salmon y Cartwright, el CritCC resulta demasiado restrictivo y no es, por tanto, adecuado en estos casos.

No es necesario entrar aquí en los detalles de las propuestas de Salmon y Cartwright. Bastará con insistir en que la idea fundamental para ambos consiste en que, de hecho, y tal y como postula el PCC, toda correlación –incluidas aquellas correlaciones que aparecen en contextos indeterministas genuinos– tiene una explicación causal (bien en términos de causa directa entre los eventos correlacionados, bien en términos de una causa común a ambos). Por otro lado, la principal diferencia entre las propuestas de Salmon y Cartwright se halla en el grado de generalización que alcanzan los nuevos criterios propuestos (o del grado de debilitación a que cada uno de ellos somete al CritCC). En este sentido, las *horquillas interactivas* de Salmon constituyen quizá un criterio demasiado débil y amplio que, además de dar cabida a los casos específicos en los que estamos interesados (de causas comunes genuinamente indeterministas), incluye estructuras causales que no debiéramos considerar como de causa común. Más preciso es el criterio generalizado de Nancy Cartwright.

Pero, ¿es absolutamente necesario reformular el CritCC, tal y como proponen Salmon y Cartwright, para poder superar los problemas que a éste se le plantean en contextos de indeterminismo genuino? O, de otro modo, ¿necesitamos dejar a un lado la caracterización de la idea de causa común original de Reichenbach? La respuesta a estas preguntas no tiene por qué ser afirmativa, siempre que estemos preparados a optar por una vía alternativa, que consistiría en insistir, tras un análisis detallado de las presuposiciones y compromisos, tanto conceptuales como formales del PCC, en que éste al completo –es decir, la conjunción del PosCC y el CritCC– es válido y aplicable, incluso en contextos indeterministas [San Pedro and Suárez (2009)].

Esta tercera posibilidad es en esencia lo que se desprende del trabajo sobre el significado y alcance formal de las relaciones de “apantallamiento” del CritCC llevado a cabo por la Escuela de Budapest (a partir de Hofer-Szabó *et al.* (1999)). En concreto, el trabajo de la Escuela de Budapest proporciona un análisis detallado desde el punto de vista formal (matemático) del concepto de causa común a través de las condiciones de “apantallamiento”. Los principales resultados se refieren a sendos teoremas por los que queda (formalmente) demostrada la posibilidad teórica de la existencia de eventos de causa común, caracterizados mediante el CritCC, para cualquier correlación dada. Estos teoremas, que podemos llamar de “extensibilidad” (*extendibility*) y “compleción causal” (*common cause completeness*), tienen en su origen la intuición de que siempre es posible ampliar nuestro campo inicial de eventos con la intención de hallar nuevos eventos que, una vez considerados, nos puedan proporcionar una mejor explicación. En otras palabras, los teoremas de la Escuela de Budapest reflejan el hecho de que nuestra descripción inicial de unos hechos (la observación de una correlación, en este caso) bien puede ser incompleta, y carecer de la información relevante para su explicación. No resultaría extraño, en absoluto, que fuéramos incapaces de encontrar una explicación de causa común en estas condiciones. Sería, de hecho, bastante esperable. A pesar de ello, la intuición es que sí parece posible, en todo caso, ampliar nuestras consideraciones, recabar una mayor información, y todo ello en medida suficiente para dar con una explicación del tipo que buscamos (una causa común). Por tanto, y partiendo de estas intuiciones, los teoremas de la Escuela de Budapest muestran, a nivel formal, que nunca podemos descartar la existencia de una causa común para una correlación cualquiera –sea ésta de origen genuinamente indeterminista o no–, que satisfaga el CritCC. Todo esto es, por supuesto, aplicable a nuestro anterior ejemplo del átomo. En ese caso, sin embargo, es necesario observar que todo evento causa común C que cumpla las relaciones de “apantallamiento” deberá forzosamente operar de forma completamente determinista. En otras palabras, lo que los resultados de la Escuela de Budapest implican es que, en casos como el del átomo que se desintegra en que las correlaciones observadas son *perfectas*, todos los candidatos a causa común que encontremos al aplicar las correspondientes herramientas formales –eventos, todos ellos, que cumplirán las relaciones impuestas por el CritCC– se comportarán de forma determinista.

Parecemos estar, por lo tanto, ante una herramienta formal que permite una defensa aparentemente definitiva del PCC. A la luz de estos resultados, no solo parece posible rebatir los argumentos más radicales en contra del PCC, como en el caso de van Fraassen, sino que ni tan siquiera parece ya necesario dejar a un lado la propuesta original de Reichenbach para la caracterización de las causas comunes, como creen Salmon y Cartwright. El asunto resulta algo más complejo, sin embargo, dados los problemas a que se enfrenta el programa “Reichenbachiano” de la Escuela de Budapest. Las dificulta-

des vienen de la mano de la alta sofisticación matemática y formal en que se apoyan los resultados mencionados. En pocas palabras, los teoremas de “extensibilidad” y “compleción causal” están expuestos a severas carencias de carácter interpretativo, en tanto que los eventos que involucran, más allá de la información ya disponible inicialmente, no dejan de ser puro artificio formal en muchos casos. Dicho de otro modo, se hace difícil en muchos casos –quizá en demasiados casos– interpretar físicamente y de forma adecuada los eventos de causa común que resultan como consecuencia de la aplicación de los teoremas de “extensibilidad” y “compleción causal”. La aplicabilidad y utilidad práctica de los resultados de la Escuela de Budapest puede verse, por lo tanto, severamente reducida. Y a causa de ello, tampoco parece ya tan claro que podamos seguir manteniendo el CritCC como condición necesaria (al menos) para la caracterización de causas comunes.

Esto nos lleva a una cuarta y última opción que consistiría, siguiendo a Jon Williamson, en tomar el PCC simplemente como un principio guía metodológico para la inferencia causal en casos concretos [Williamson (2005)]. Esta es, de algún modo, la forma en que buena parte de las prácticas científicas que aplican métodos de inferencia causal tales como el PCC o similares se acercan a ellos. Se trata por tanto de insistir en el rango específico de aplicación de la vertiente metodológica del PCC a ciertos casos, experimentos, etc., cuyas peculiaridades permiten realizar la correspondiente inferencia causal de forma fiable. El conocimiento causal así obtenido constituirá a su vez la base sobre la que poder aplicar de nuevo el PCC con ciertas garantías. Esta forma de proceder está, por supuesto, bastante más expuesta a errores en las inferencias –como ocurre, por ejemplo, ante la carencia de suficiente información causal–. Sin embargo, el que debamos apoyarnos en un método de prueba-error no deja de permitir avances en el conocimiento causal, como de hecho ocurre en la mayor parte de los casos.

Departamento de Lógica y Filosofía de la Ciencia

Universidad del País Vasco, UPV/EHU

Carlos Santamaría Zentroa, D10

Pl. Elhuyar, 2, 20018 Donostia-San Sebastián

E-mail: inaki.sanpedro@ehu.es

Institute of Philosophy, School of Advanced Study, London University/

Senate House, Malet Street,

London WC1H 7EU, Reino Unido

E-mail: mauricio.suarez@sas.ac.uk

y Departamento de Lógica y Filosofía de la Ciencia

Universidad Complutense, 28040 Madrid.

E-mail: msuarez@filos.ucm.es

AGRADECIMIENTOS

Agradecemos los comentarios y sugerencias recibidas en los seminarios celebrados en la Universidad Complutense de Madrid y en la London School of Economics, así como el apoyo económico del proyecto “Inferencia, Causalidad y Ciencia” (FI12011-29834-C03-01) del Ministerio de Economía y Competitividad del Gobierno de España.

NOTAS

¹ Es preciso hacer una pequeña apreciación con respecto a la analogía con cadenas causales. La referencia a cadenas causales como base de una estructura causal debe tomarse aquí en un sentido amplio. La analogía, en concreto, no pretende enfatizar tanto la “forma” en que estarían estructuradas las relaciones entre eventos (causas y efectos) sino, más bien, a las relaciones mismas. Dicho de otro modo, la referencia a cadena causal no implica, ni debe sugerir al lector, que los eventos deban concebirse formando una única cadena o línea causal, i.e. causa-efecto/causa-efecto/causa-efecto.... Se trata más bien de enfatizar el hecho de que las causas y efectos en cuestión se conciben, en teorías deterministas, como eventos contiguos, que se encuentran relacionados de una forma robusta, aun dejando abierta la posibilidad de que un efecto, por ejemplo, sea producto de una o más causas. En este sentido, la idea de “red de cadenas causales” es quizás algo más precisa.

² No vamos a entrar aquí en los diferentes problemas asociados a tales modelos de explicación causal, que son numerosos e interesantes.

³ Este es, en buena parte, el origen, de la incompatibilidad entre nuestras dos mejores teorías del momento: la relatividad general y la mecánica cuántica.

⁴ Recordemos que en el caso determinista, y según la analogía de la “red de cadenas causales” que hemos utilizado con anterioridad, la ocurrencia de una causa es suficiente para que se dé el efecto. La ocurrencia del efecto, sin embargo, no implica la ocurrencia de una causa en concreto. (Este sería el caso, por ejemplo, si el efecto en cuestión tuviera diferentes causas posibles). En otras palabras, incluso en un contexto determinista, la presencia de una causa no es necesaria para la ocurrencia del efecto correspondiente.

⁵ Veremos más adelante cómo el indeterminismo en sí ya es un potencial problema para esta concepción de la causalidad.

⁶ El mismo Reichenbach era consciente de que este tipo de caso era ciertamente habitual.

⁷ Resulta, sin embargo, un tanto paradójico que la piedra de toque para la resolución de este tipo de cuestiones se encuentre precisamente en el indeterminismo genuino ya que, como hemos apuntado anteriormente, la noción de causalidad que el PCC incorpora está diseñada, de hecho, para poder dar cuenta de situaciones que involucran indeterminismo. La paradoja no es tal, en cualquier caso, si atendemos a los diferentes argumentos que apuntan a que el PCC, y métodos de inferencia similares, buscan establecer en última instancia una estructura causal profunda determinista. Es decir, indeterminismo sí, pero sólo a nivel superficial o aparente, en tanto que relacionado exclusivamente con las correlaciones que pretendemos explicar, pero no con las causas comunes que las explican.

⁸ Es decir, si uno de los subproductos se dispersa con ángulo a , el otro deberá hacerlo con ángulo $b = a + \pi$.

⁹ Tanto la *horquilla interactiva* de Salmon como la *horquilla generalizada* de Cartwright se definen en contraste con la idea de *horquilla conjuntiva*, como también se conoce a las relaciones de “apantallamiento” (1) y (2) que aparecen en el CritCC.

REFERENCIAS BIBLIOGRÁFICAS

- CARTWRIGHT, N. (1987), “How To Tell a Common Cause: Generalizations of the Conjunctive Fork Criterion”; en Fetzer, J. H. (ed.), *Probability and Causality*, Dordrecht, Reidel Pub. Co, pp. 181-88.
- HOFER-SZABÓ, G., RÉDEI, M. y SZABÓ, L. E. (1999), “On Reichenbach’s Common Cause Principle and Reichenbach’s Notion of Common Cause”, *The British Journal for the Philosophy of Science*, 50, pp. 377–99.
- REICHENBACH, H. (1956), *The Direction of Time*; editado por Maria Reichenbach. Edición completa del original publicado en 1956 por University of California Press, Nueva York, Dover, 1999.
- SALMON, W. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton, Princeton University Press.
- SAN PEDRO, I. y SUÁREZ, M. (2009), “Reichenbach’s Common Cause Principle and Indeterminism: A Review”; en González Recio, J. L. (ed.), *Philosophical Essays on Physics and Biology*, Hildesheim, Georg Olms Verlag.
- SUÁREZ, M. (2007), “Causal Inference in Quantum Mechanics: A Reassessment”; en Russo, F. y Williamson, J. (eds.), *Causality and Probability in the Sciences*, Londres, London College Publications, pp. 65-106.
- VAN FRAASSEN, B. C. (1982a), “The Charybdis of Realism: Epistemological Implications of Bell’s Inequality”, *Synthese*, 52, pp. 25-38.
- (1982b), “Rational Belief and the Common Cause Principle”; en McLaughlin, R. (ed.), *What? Where? When? Why? Essays in Honour of Wesley Salmon*, Dordrecht, Reidel, pp. 193-209.
- WILLIAMSON, J. (2005), *Bayesian Nets and Causality: Philosophical and Computational Foundations*, Oxford, Oxford University Press.