

¿La explicación explica la abducción? Encuentros y desencuentros entre Inteligencia Artificial y Filosofía de la Ciencia

David Gaytán Cabrera

UACM/ Programa *Principhia* / UNAM,
México

Abstract. Dov Gabbay y John Woods defendieron un enfoque no tradicional para entender formalmente la abducción. Una parte significativa de la defensa que proponen Gabbay y Woods está basada en una robusta crítica contra lo que podría llamarse “el enfoque formal tradicional de la abducción”. En este artículo realizaré una evaluación de un punto específico de esta crítica, el cual está relacionado con la noción de explicación. Haré esta evaluación a la luz del importante debate que sobre la caracterización de la explicación científica ha tomado ya más de 5 décadas de discusión fructífera en el ámbito de la Filosofía. Finalmente propondré que hay una falla importante en la crítica realizada por estos autores y que tal falla nos lleva a revalorar la importancia de la interacción entre dos disciplinas de investigación involucradas en estos enfoques: la Inteligencia Artificial y la Filosofía de la Ciencia.

Palabras clave: abducción, explicación, Inteligencia Artificial, Filosofía de la Ciencia.

1. Introducción

En un artículo reciente, Dov Gabbay y John Woods defendieron un enfoque no tradicional para entender formalmente la abducción. Este nuevo enfoque tiene virtudes considerables: puede decirse que se trata de una apuesta por una representación más pluralista y en cierto sentido debilitada de la noción de inferencia abductiva. Algunos rasgos importantes de este nuevo enfoque son, por ejemplo, la inclusión más franca de parámetros para la representación de relevancia, la integración de varios parámetros provenientes de diversas lógicas no clásicas y la distinción entre explicación y abducción. Una parte significativa de la defensa que proponen sus autores está basada en una robusta crítica contra lo que podría llamarse “el enfoque formal tradicional”. Uno de los ítems que constituyen la crítica del artículo se refiere a la perspectiva explicacionista en que según Gabbay y Woods, en el punto de vista tradicional es interpretado el carácter consecuencialista de la abducción. Realizaré una evaluación de este punto específico de crítica relacionado con la explicación, a la luz del importante debate que sobre la caracterización de la explicación científica ha tomado ya más de 5 décadas de discusión fructífera en el ámbito de la Filosofía de la Ciencia. El examen me llevará a mostrar una falla que considero importante en el punto de crítica mencionado. Lo interesante es

que parece ser una falla que penetra la forma en que es tratado el tema desde la Inteligencia Artificial. A partir de lo anterior, intentaré extraer algunas conclusiones finales acerca de la relación entre esta disciplina de investigación y la Filosofía de la Ciencia.

2. El enfoque formal tradicional para la elucidación formal de la abducción, según Gabbay y Woods

Las principales contribuciones originales del enfoque tradicional de la abducción descrito por Gabbay y Woods en [1], en su párrafo 12.1, se deben a varios autores, como Aliseda [2], Kowaliski [3], Kuipers [4], Magnani [5] y, conjuntamente, a Kakas, Kowalski y Mancarella [6]. Gabbay y Woods proponen llamar a la parte nuclear de este enfoque “el esquema *AKM*”, y sostienen que puede describirse simplícidamente mediante los siguientes 8 ítems (incluyo aquí el inciso adicional 1 que se agregó en [1] con el objetivo de facilitar en ese artículo la discusión que ellos desarrollaron alrededor de los requerimientos para toda elucidación de la noción abducción):

1. T!
Que puede leerse como: Se tiene el objetivo de encontrar una explicación para E.
2. E
Que puede leerse como: Se tiene el dato de que el suceso E ocurrió.
3. No ocurre que: $K \Rightarrow E$
Que puede leerse como: K, nuestra base de conocimientos, no tiene a E como una de sus consecuencias.
4. No ocurre que: $H \Rightarrow E$
Que puede leerse como: H, nuestra hipótesis abductiva, no tiene, por sí sola, a E como una de sus consecuencias.
5. $K(H)$ es consistente
Que puede leerse como: la revisión de K, resultante de agregar H a K, es consistente
6. $K(H)$ es mínima
Que puede leerse como: la revisión de $K(H)$ es mínima
7. $K(H) \Rightarrow E$
Que puede leerse como: la revisión $K(H)$ sí tiene como una de sus consecuencias a E
8. Por lo tanto: H

Antes de enumerar los 8 ítems del esquema *AKM*, Gabbay y Woods le atribuyen a éste dos ideas básicas: I1) que está estructurado para capturar la abducción consecuencialista. Esto significa que en *AKM*, la abducción está concebida como la inferencia de una hipótesis H, a partir de condiciones específicas para una noción de consecuencia, en este caso, la relación expresada por “ $\alpha \Rightarrow \beta$ ”, para cualesquier α, β . La segunda idea, I2), es que la manera estándar en que los filósofos de la ciencia interpretan la noción de consecuencia que caracterizaría a $\alpha \Rightarrow \beta$, es tomarla como una relación explicativa, expresaría que α explica β . Finalmente, en [1] se concluye que el esquema *AKM* subrepresenta la estructura lógica de la abducción:

“Proposition 12.1 (Under-representation) The *AKM*-schema under-represent the logical structure of abduction.”

Para esta conclusión los autores se apoyan en dos tesis contenidas en la siguiente aseveración del párrafo 12.1:

*We are suggesting, both that the *AKM*-parameters are too few so that their interpretation is too narrow to afford a suitably comprehensive representation of the structure of abduction.*

La distinción de las dos tesis a las que he aludido, puede quedar como sigue:

- a) El esquema *AKM* incluye pocos parámetros para permitir una representación adecuadamente comprensiva de la inferencia abductiva.*
- b) La interpretación que el esquema *AKM* supone para estos parámetros es demasiado estrecha para permitir una representación adecuadamente comprensiva de la inferencia abductiva.*

La primera de las tesis, a), se sostiene en [1] sobre la idea de que valdría la pena elucidar formalmente los vínculos explicativos que supone la relación de consecuencia en el esquema *AKM*. Si no se hacen explícitos tales vínculos o condiciones la idea de consecuencia representada por la flecha “ \Rightarrow ” no funcionaría como una elucidación de la abducción. Así, se demostraría la escasez de parámetros para la elucidación de la inferencia abductiva. La segunda tesis, b), se apoya de modo central en el supuesto de que entre los filósofos de la ciencia la interpretación estándar que elucidaría la relación de consecuencia es la interpretación explicativa de esta relación. No obstante, también se sostiene en otro supuesto: la asunción de que esta relación explicativa es caracterizable identificando las condiciones de la relación en cuestión, con las condiciones del modelo Nomológico Deductivo de Hempel y Oppenheim, de 1948, [7] (en adelante “modelo D-N”). Con estos antecedentes, el argumento para b) iría como sigue: Si ésta es la caracterización de la noción de consecuencia, es claro que *K* tendría que incluir un conjunto de leyes y *H* tendría que ser una condición inicial. No obstante, el carácter hipotético de *H* haría imposible que se tratara de una condición inicial. De lo inmediatamente dicho se sigue que la interpretación de los parámetros es demasiado estrecha para una representación adecuada.

3. Problemas en la crítica de Gabbay y Woods, en el caso de la interpretación explicativa

Defenderé en este apartado que al punto de crítica descrito en el apartado anterior le ocurre que: T1) no toma en cuenta seriamente el contexto de la discusión en el ámbito de la Filosofía de la Ciencia en sus términos más generales y, T2), no considera las diversas posiciones que sobre la teoría de la explicación se han defendido en esta disciplina filosófica.

Es casi inevitable comparar la perspectiva de Woods y Gabbay, respecto al problema de la representación adecuada de la inferencia abductiva, con el problema de la

representación formal adecuada de la explicación científica, tal como se le ha asumido a éste desde finales de la década de los 40's y hasta la fecha. Uno podría defender este punto de vista de un modo relativamente fácil como sigue. i) Las tesis a) y b) suponen, como explícitamente se enuncia en el artículo [1], que una evaluación de las capacidades del modelo pasa por dos criterios, resumidos en dos conceptos: parámetros de representación suficientes e interpretaciones adecuadas de éstos. ii) De un modo análogo, la discusión sobre la posibilidad de construir un modelo formal de explicación, a lo largo de la segunda mitad del siglo XX, corrió sobre estos mismos supuestos. Obsérvese por ejemplo, la crítica de Scriven en 1962 [8] contra la necesidad y contra la suficiencia del modelo Nomológico Deductivo propuesto por [7]. También puede consultarse la crítica de Achinstein en 1983 [9] que denunció la no inclusión formal de la noción de causalidad en los modelos de explicación; o la de van Fraassen en 1980 [10], sobre la no inclusión del contexto en los modelos formales de explicación conocidos hasta entonces. Así iii), *prima facie*, es un procedimiento adecuado el de mostrar aspectos del objeto modelado que no están incluidos en un modelo, con la intención de mostrar las debilidades de este último en su objetivo de representación.

No obstante, si se le examina con detenimiento, se puede mostrar que el procedimiento seguido en [Gabbay & Woods, 2006] puede llevar a errores. La falla básica puede describirse como la relativización del conjunto de aspectos a los que el evaluador en turno tendría permitido apelar para construir su crítica contra algún modelo. Formulado en una pregunta, el problema es ¿Cualquier aspecto no capturado por el modelo puede articular una crítica de este tipo que sea sensata? La respuesta es no. El argumento central para una respuesta negativa a esta pregunta es que el conjunto de aspectos de un objeto, sin importar si es conceptual o físico, es indeterminado. Esta es una conclusión que puede derivarse de un argumento que aparece, en otro contexto, en un texto de Putnam de 1981[11] y en uno de la misma autoría, de 1994[12]. El argumento versa sobre la indeterminación del número de objetos en que uno podría dividir cualquier ontología propuesta. El argumento es una aplicación que hace Putnam al mundo empírico, de un teorema en Teoría de Modelos: el resultado conocido como el teorema de Skolem-Lowenheim. La idea de Putnam es que, del mismo modo en que la totalidad de las verdades acerca de los objetos matemáticos que pueden expresarse en el lenguaje de las matemáticas no puede determinar a qué objetos nos estamos refiriendo, no podemos tampoco determinar la totalidad de los objetos empíricos. Se puede ver que tampoco podemos hacerlo en un dominio conceptual diferente al de las matemáticas. Así, hay una indeterminación esencial acerca de los objetos que uno puede considerar en cada realidad supuesta. Esto vale para los componentes que uno considere importantes en cada objeto modelado. Así, siguiendo esta línea de pensamiento, siempre tendremos la posibilidad de exigir nuevos aspectos que modelar. Consecuentemente, la tesis a), por sí sola, no puede tomarse como razón suficiente para que a partir de la incapacidad, que tiene uno cualquiera de los modelos considerados, de capturar alguno de los aspectos estipulados, uno pueda intentar restringir (o hacer racional) una elección entre los diversos modelos. Por lo tanto, dado el análisis presentado hasta ahora, tal como se ha expresado la tesis a) en [1], ésta es una tesis que no tiene el alcance que se le atribuye ahí mismo.

Analizaré ahora no la tesis b), sino cómo se le defiende en [1]. Puede concederse que la interpretación estándar de la abducción consecuencialista, entre los filósofos de

la ciencia, sea una relación explicativa. No obstante, lo anterior no hace que la afirmación de existencia de esa interpretación estándar deje de ser una aseveración que tendría que justificarse, lo cual no se realiza en el artículo. Sin soslayar lo anterior, lo que es claro es que esa aseveración, aun concediéndola, no implica que la caracterización de esa interpretación tenga que estar basada en el modelo D-N de explicación. Este modelo puede considerarse un modelo clásico, quizá paradigmático para los fines del debate de la segunda mitad del siglo XX, pero en ningún sentido es necesario que se le tome como *la* caracterización de la relación de explicación. Tampoco puede sostenerse que sea la caracterización estándar de la interpretación estándar entre los filósofos de la ciencia. Por el contrario, el debate en la Filosofía de la Ciencia, alrededor del problema de si es posible un modelo formal de explicación científica, ha producido una cantidad muy grande de muy diversas alternativas de modelación. Aún más, muchas de estas alternativas fueron justificadas parcialmente mediante una dura crítica contra el modelo D-N y, también, contra el modelo Inductivo-Estadístico (en adelante, “modelo I-S”) propuesto algunos años después que el D-N, en 1965 [13]. Algunos casos que pueden ilustrar lo anterior son los siguientes: la crítica de Railton en 1981 [14] contra el modelo I-S, sobre la cual se erige un modelo basado en una descripción deductiva y un *addendum* descriptivo; también la crítica de Salmon, de 1984, [15] dirigida contra la estructura argumentativa de las explicaciones, a partir de lo cual desarrolla un modelo que puede llamarse “de relevancia estadística”; otro caso muy significativo es la crítica contra la posibilidad de determinar un contexto de relevancia uniforme, que esgrime van Fraassen, 1980, [10], y el modelo de explicación que se propone ahí mismo, basado en lógica erotética. En consecuencia, no debería asumirse que la única caracterización de una interpretación explicacionista de la abducción consecuencialista sea el modelo D-N Hempeliano, ni siquiera sería conveniente asumir que es su caracterización estándar.

Todas estas consideraciones sobre la defensa de la tesis b) tienen por objetivo mostrar que, por lo tanto, no deberíamos asumir que la caracterización de la interpretación de la abducción consecuencialista deba ser el modelo D-N de Hempel y Oppenheim, ni siquiera debería considerarse a este modelo clásico la caracterización estándar de la interpretación estándar de la abducción consecuencialista. En este sentido, la defensa de la tesis b) tal como se presenta en [1], es inadecuada en el sentido de que intenta rechazar la interpretación explicacionista con base en una caracterización no representativa de la relación explicativa.

Ahora bien, Gabbay y Woods pueden seguir manteniendo la tesis b), si hallaran una forma de sostener que las interpretaciones de los parámetros en *AKM* son demasiado estrechas para permitir una representación adecuadamente comprensiva de la estructura de la inferencia abductiva. Si, efectivamente, *AKM* asumiera el modelo D-N Hempeliano como la caracterización de su interpretación explicacionista de la abducción consecuencialista, entonces *AKM* estaría en problemas. Notemos, de cualquier forma, que hemos mostrado que ello no necesariamente pondría en problemas a toda interpretación explicacionista de la versión consecuencialista de la abducción.

Un examen de este problema más particular contra *AKM*, requeriría un análisis más detallado de las propuestas que conformarían, según Gabbay y Woods, el esquema *AKM*; es decir, se necesitaría un examen de las propuestas de Aliseda, Kuipers, Kowalski, Kakas y Magnano. No realizaré este análisis en este artículo. Sin embargo, me interesa enfatizar que en términos generales, no sólo en *AKM*, sino en varias otras propuestas en Inteligencia Artificial que intentan caracterizar la explicación, existe una

tendencia a la sobresimplificación de la noción de explicación, identificando los ítems importantes de esta noción con las condiciones de los modelos D-N e I-S, los que podríamos llamar “modelos clásicos de la explicación”. Lo que es más importante, en algunas propuestas, no solo se asumen los modelos clásicos de Hempel y Oppenheim como la interpretación estándar de la explicación en Filosofía de la Ciencia, sino que se omite gran parte de los problemas planteados en el debate que ha dado esta disciplina filosófica. En el caso de Aliseda, [2], que puede entenderse quizá como la propuesta más importante y dominante dentro del enfoque formal tradicional de la abducción, se dedica un análisis centrado en estos modelos clásicos y una consideración de algunos problemas planteados por Salmon. Aunque la propuesta contenida en [2] debe considerarse una teoría lúcida y fundacional en el estudio de la abducción, en lo que respecta a su vinculación con la explicación se mantiene en los límites todavía estrechos de la caracterización de los modelos clásicos.

El caso del enfoque *GW*, el propuesto por Gabbay y Woods, no corre mejor suerte en relación con la noción de explicación. Una muestra de esto es, tal como lo he descrito aquí, la manera en que se organiza en [1] la defensa de la tesis b); pero también puede verse esta tendencia en algunas asunciones que pueden tomarse de un artículo de Woods de 2007 [16]. La propuesta del enfoque *GW* para la abducción tiende no solo a asumir la reducción de la caracterización de la explicación a los modelos clásicos, sino a minimizar el papel de la explicación en la abducción.

Los problemas para sostener a) y b), hacen que la conclusión de que el esquema *AKM* sub-representa la estructura lógica de la abducción, la proposición 12.1, aparezca como no bien fundada, por lo menos por esta vía.

El examen presentado hasta aquí de este punto de crítica de [1], debe entenderse como una defensa de T1), a saber, que la crítica en relación con la explicación no toma en cuenta seriamente el contexto de la discusión en el ámbito de la Filosofía de la Ciencia en sus términos más generales. También es una defensa de T2), es decir, que no considera las diversas posiciones que sobre la teoría de la explicación se han defendido en esta disciplina filosófica.

4. Las intuiciones conceptuales: Inteligencia Artificial y Filosofía de la Ciencia

En el caso de la tesis b), la asunción del modelo D-N es crucial para el fallo de la argumentación en ese apartado del artículo de Woods y Gabbay y se debe, en gran medida, a la sobresimplificación que suele hacerse en otras disciplinas, en este caso en la Inteligencia Artificial, de las discusiones epistemológicas acontecidas en el ámbito de la Filosofía de la Ciencia. Por el contrario, quizá deberíamos confiar en que una actitud más abierta y participativa entre las dos disciplinas, Inteligencia Artificial y Filosofía de Ciencia, podría venir bien al enriquecimiento de las producciones teóricas de ambas. En el caso que nos ocupa ahora, y en especial con relación al caso de la tesis b), bastaría con diversificar nuestra idea de los problemas involucrados en el debate que sobre la explicación se ha generado en la Filosofía de la Ciencia, y conocer los diversos modelos de explicación que se han propuesto en él. En cambio, para el caso de la tesis a), es necesario notar que, siguiendo a Putnam, el que uno pueda usar cualquier aspecto del objeto modelado para realizar una crítica contra un modelo dado, no es indicativo

de la imposibilidad de discutir sobre la adecuación de un modelo para ciertos fines de representación. Sabemos que la caracterización de un conjunto de aspectos sobre los que se pudiera establecer un criterio de elección o crítica entre modelos de la abducción es un asunto bien problemático, pero esto no quiere decir que no puedan establecerse, con algún criterio razonable, algunos aspectos a los que uno podría apelar. La clave está en que esos aspectos, como ocurre en el caso de la discusión sobre la explicación científica en Filosofía de la Ciencia, deberían tener algún grado de atrincheramiento en la cultura general que subyace a esa discusión. La discusión sobre la explicación, en Filosofía de la Ciencia, se dio en el marco de una serie de supuestos alrededor de la explicación, que estaban más o menos atrincherados en la cultura científica y filosófica pertinente al caso. Puede confrontarse para este asunto el texto iluminador Salmon [17]. Por ejemplo, la idea de la necesidad de una vinculación causal en la explicación o la idea de la necesidad de la apelación a leyes, no fueron una construcción ad-hoc de un ítem particular de ese debate en Filosofía de la Ciencia. Por el contrario, fueron aspectos recogidos y precisados a partir de las intuiciones de quienes, a lo largo de siglos, habían tocado el tema de la explicación. En la propuesta defendida en [1] la reconstrucción histórica de estos aspectos de la abducción que deberían modelarse, no parece bien desarrollada. Se hacen algunas pocas menciones a la tradición de Peirce en la abducción, sin un análisis de esa tradición. En todo caso, incluso, la propuesta peirciana parece más coincidente con la posición del esquema *AKM*, que con el modelo que ellos defienden, *GW*. Esto es cierto por lo menos en lo que respecta a la estructura con la que en ocasiones Peirce describió la abducción, pero también respecto a la tendencia de este pensador a vincularla con las relaciones de causalidad.

5. La relevancia explicativa: Inteligencia Artificial y Filosofía de la Ciencia

Finalmente, daré apenas un bosquejo de una aportación positiva para toda esta discusión. Un avance en la dirección correcta para la construcción de un contexto que sirviera de contraste para la elección de aspectos a tomar en cuenta en la elección racional entre modelos de la abducción podría partir de las conexiones de la abducción con el programa general de la disciplina de la Inteligencia Artificial. Un análisis importante en esta agenda de reconstrucción de la tradición, en particular sobre las nociones de relevancia en Inteligencia Artificial, podría iniciar prometedoramente con un análisis de los problemas del *frame* y de la cualificación, los cuales están en la base de la tradición en Lógica No Monotónica para lograr cierta clase de representaciones en Inteligencia Artificial. Particularmente, un examen de las discusiones de McCarthy y Hayes en 1969 [18] y del artículo posterior, 1977, de este primer autor [19] seguramente sería un estudio clave. Vale decir también que Minsky, en 1974 [20] también hizo algunas reflexiones que deberían reconsiderarse y que vinculan, con toda claridad, la Inteligencia Artificial con la Filosofía de la Ciencia. Exploremos el caso del problema del *frame* y el de la cualificación con mayor detalle, tomando en cuenta su caracterización en [18] y [19].

El problema del *frame*, puede pensarse como el problema de saber cuáles supuestos deben modificarse y cuáles no, con base en nueva información. En el mismo tenor, el problema de la cualificación puede pensarse como el problema de saber cuáles son los

supuestos posibles que deben tomarse en cuenta ante la resolución de un problema. Visto desde un punto de vista un poco más abstracto, ambos problemas tratan de la forma en que determinamos, a partir de una situación dada (sea una situación inicial o el resultado del cambio de esta situación inicial), lo que es asumible. Una forma de interpretar estos problemas, así unidos es mediante la siguiente pregunta: ¿Cómo determinar lo que es pertinente para sacar consecuencias a partir de una situación dada? Pueden así plantearse dos subcasos de este problema: Cuando la situación se toma como punto de partida; es decir, sin considerar aumentos de información (cualificación) y, el caso contrario, cuando debe tomarse en cuenta un aumento de información en el conjunto-base de datos (*frame*). Podemos llamar a este problema compuesto “el problema general de la relevancia (o de la pertinencia)”.

Un detalle, sin embargo, es imprescindible de tomarse en cuenta: el problema de la pertinencia general no puede caracterizarse cabalmente si no se supone que la determinación que exige tiene que hacerse con base en información incompleta. El problema de la relevancia general surge en el marco de una situación con información incompleta. Una situación así, si se aborda de modo consecuencialista, implica dos subproblemas más:

- 1) El problema de determinar estructuras lógicas que permitan saltar a la conclusión o, en otras palabras, que permitan conjeturar, en un razonamiento.
- 2) El problema de manipular razonamientos basados en supuestos falibles.

Lo que está de fondo en ambos subproblemas es cómo llegar a conclusiones de manera “sensata”, a partir de premisas que son insuficientes. Si entendemos “conjetura” como un razonamiento que supone un salto a la conclusión, la cuestión central se traduce en cómo hacer y manipular conjeturas sensatas. Podemos llamar a este problema general, el “problema de la conjetura”. El problema de la conjetura puede caracterizarse como una de las formas de abordar, desde un punto de vista lógico, el problema general de la relevancia. Así, aparece con mayor naturalidad el desiderata de que los problemas surgidos en Inteligencia Artificial sean considerados como problemas que involucran discusiones de Lógica pero también de Epistemología, pues el problema general de la relevancia tiene profundas raíces en cuestiones de índole epistemológica.

El problema de la conjetura, una forma de abordar el problema general de la relevancia, está estrechamente vinculado tanto con el problema de la modelación de la abducción, como con el problema de la modelación de la explicación. Las condiciones que se añadan a los núcleos esquemáticos de los modelos de una y otra noción deberían capturar adecuadamente los rasgos de relevancia atinentes a cada caso. En el primer caso, el de la abducción, para concluir la hipótesis, en el segundo, la explicación, para concluir el hecho que se pretende explicar. Ambos problemas de modelación requerirán de algún tipo de solución al problema general de la relevancia y, si se les aborda de modo consecuencialista, esto se traducirá en la búsqueda de una solución al problema de la conjetura. Dos insumos para abordar estos problemas son pertinentes. Por un lado, reflexiones sobre las implicaciones que tienen ciertos formalismos no monotónicos en la noción de racionalidad (las que considera reflexiones en el límite de la Filosofía de la Lógica y la Epistemología) y, por otro lado, reflexiones epistemológicas sobre las condiciones que puedan constituir una noción de relevancia vinculada a una conjetura particular, por ejemplo, la conjetura de un hecho particular a partir de una teoría. Un ejemplo del primer tipo de reflexiones es el conjunto de disquisiciones elaboradas por

Morado en [21]. Ejemplos del segundo tipo de reflexiones pueden encontrarse en las propuestas de Salmon en [15] y de van Fraassen en [10].

6. Conclusiones

El bosquejo de la sección anterior muestra, en una dirección, cómo el problema general de la relevancia y la forma en que se aborda formalmente desde la Inteligencia Artificial, puede tener importantes consecuencias en la modelación de la explicación, desde el ámbito de la Filosofía de la Ciencia. En otra dirección, ilustra la pertinencia de tomar en cuenta los amplios estudios que sobre la noción de relevancia explicativa se han hecho desde la Filosofía de la Ciencia, a la hora de la construcción de restricciones de relevancia en un modelo de abducción que se considere, desde el punto de vista de la Inteligencia Artificial, vinculado a la noción de explicación.

A través de las reflexiones contenidas en este artículo, se pueden establecer tentativamente las siguientes aseveraciones:

- 1) El punto de crítica de Gabbay y Woods vinculado a la caracterización explicacionista de la interpretación consecuencialista de la explicación, tiene fallas que deben considerarse.
- 2) Estas fallas muestran una deficiencia en la interacción entre investigaciones de dos diferentes disciplinas: Inteligencia Artificial y Filosofía de la Ciencia.
- 3) Un caso posible de cooperación, a nivel metodológico, entre ambas disciplinas, es el uso de conceptos intuitivos atrincherados en torno a la noción de explicación y de abducción.
- 4) Un caso plausible de revisión es el de las condiciones de relevancia para una interpretación consecuencialista de la abducción. Por un lado, es importante recuperar las intuiciones básicas del problema general de la relevancia en Inteligencia Artificial, por otro, reflexiones epistemológicas sobre los formalismos y, finalmente, reflexiones sobre las condiciones de relevancia de una explicación.
- 5) Puede resultar conveniente unir los esfuerzos de la Inteligencia Artificial y de la Filosofía de la Ciencia, por lo menos en el problema de la modelación de la abducción y en el de la modelación de la explicación.

Agradecimientos. Agradezco el apoyo para esta investigación, por parte de la Secretaría de Ciencia y Tecnología del Distrito Federal a través del proyecto PI2014-47.

Referencias

1. Gabbay, Dov, Woods, John: *Advices on Abductive Logic*. *Logic Journal of IGPL*. 14(2) pp. 189–219 (2006)
2. Aliseda Llera, Atocha: *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. Amsterdam: Institute for Logic, Language and Computation (1997)
3. Kowalski, R. A.: *Logic for Problem Solving*. New York, Elsevier (1979)

4. Kuipers, Theo, A.F.: Abduction aiming at empirical progress of even truth approximation leading to a challenge for computational modelling. *Foundations of Science* (1999)
5. Magnani, Lorenzo, *Abduction, Reason and Science: Processes of Discovery and Explanation*. New York: Kluwer, Plenum (2001)
6. Kakas, A., Kowalski, R.A., Mancarella, P.: Knowledge assimilation and abduction. In: *Proceedings of the European Conference on Artificial Intelligence, ECAI, Springer-Verlag* (1990)
7. Hempel, Carl G., Oppenheim, Paul: *Studies in the Logic of Explanation*. *Philosophy of Science*, vol. 15, pp. 567–579 (1948)
8. Scriven, Michael: *Explanations, Predictions, and Laws*. University of Minnesota press, Minneapolis (1962)
9. Achinstein, Peter: *The Nature of Explanation*. Oxford University Press, New York (1983)
10. Van Fraassen, Bas: *The Scientific Image*, Oxford University Press, New York (1980)
11. Putnam, Hilary: *Reason, Truth and History*. Cambridge University Press (1981)
12. Putnam, Hilary: *The Dewey Lectures 1994: Senses, Non senses and The Senses: An Inquiry into the Powers of the Human Mind*. *The Journal of Philosophy Inc.*, New York (1994)
13. Hempel, Carl: *Aspects of Scientific Explanation*. The Free Press, New York (1965)
14. Railton, Peter: A Deductive-Nomological Model of Probabilistic Explanation. *Philosophy of Science*, 45 (1978)
15. Salmon, Wesley: *Scientific Explanation and the Causal Structure of the World*. Princeton University Press (1984)
16. Woods, John: Ignorance and Semantic Tableaux: Aliseda on Abduction. *Theoria*. vol. 60, pp. 305–318 (2007)
17. Salmon, Wesley: *Four decades of Scientific Explanation*. University of Minnesota (1989)
18. McCarthy, John, Hayes, P.: Some Philosophical Problems from the Standpoint of Artificial Intelligence. B. Meltzer and D. Michie (eds.), *It Machine Intelligence*, Vol. 4, Edinburgh, N.Y., p. 468 (1969)
19. McCarthy, John: Epistemological Problems of Artificial Intelligence. In: *International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1038–1044 (1977)
20. Minsky, Marvin: A frame for Representing Knowledge. AI memo, 306 (1974)
21. Morado, R.: Problemas filosóficos de la lógica no monotónica. Raúl Orayen & Alberto Moretti (eds.) *Filosofía de la lógica*, Enciclopedia Iberoamericana de Filosofía, Vol. 27, Edit. Trotta, Madrid (2004)