DOI 10.36446/af.e1004

ALGUNAS REFLEXIONES SOBRE LAS METAINFERENCIAS: ACERCA DE "WHY A LOGIC IS NOT ONLY ITS SET OF VALID INFERENCES" DE EDUARDO A. BARRIO Y FEDERICO PAILOS

Some Thoughts on Metainferences: On "Why a Logic is not only its Set of Valid Inferences" by Eduardo A. Barrio and Federico Pailos

JAVIER CASTRO ALBANO [a, b]
https://orcid.org/0000-0001-8504-929X
javiercastroalbano@gmail.com

[a] Universidad de Buenos Aires, Buenos Aires, Argentina.
[b] Universidad Nacional de Rosario, Rosario, Argentina.

Resumen

Se presenta un argumento en contra de la afirmación de Eduardo Barrio y Federico Pailos según la cual "para caracterizar una lógica no solo el nivel inferencial, sino también cada uno de los niveles metainferenciales es necesario".

Palabras clave: Metainferencias; Caracterización de una lógica.

Abstract

An argument is presented against the claim by Eduardo Barrio and Federico Pailos according to which "to characterize a logic, not only the inferential level, but also every metainferential level is necessary".

Key words: Metainferences; Characterization of a Logic.

Una buena parte del trabajo que se lleva a cabo hoy en día en la lógica tiene que ver con la construcción y la comparación de sistemas lógicos. La construcción de sistemas lógicos suele seguir el patrón habitual en el álgebra abstracta: se parte de un sistema lógico conocido y se modifican algunos de sus rasgos característicos para crear un sistema derivado, parcialmente análogo al sistema inicial. La comparación de sistemas lógicos se hacía, hasta hace muy poco tiempo, tomando en cuenta diferencias en los conjuntos de leyes lógicas de los sistemas o en los conjuntos de sus inferencias válidas. En los últimos años, sin embargo, algunas personas se han interesado también por comparar los sistemas lógicos al nivel de ciertas construcciones llamadas "metainferencias”. En particular, Eduardo Barrio y Federico Pailos han sostenido que “para caracterizar una lógica no solo el nivel inferencial, sino también cada uno de los niveles metainferenciales es necesario” (Barrio & Pailos, 2021). Explicar por qué la tesis de Barrio y Pailos no me resulta convincente es el objetivo de este trabajo.

1. Metainferencias

Las metainferencias no son criaturas exóticas. Son generalizaciones sencillas de las familiares inferencias que desde siempre han sido objeto de la lógica. Siguiendo el simbolismo que utilizan Barrio y Pailos, si Γ es un conjunto de fórmulas de un lenguaje L y A es una fórmula de L, entonces $\Gamma \Rightarrow A$ es una inferencia sobre L. Los elementos de Γ son las premisas de $\Gamma \Rightarrow A$ y A es su conclusión. Una metainferencia no es más que una inferencia de un nivel superior, cuyas premisas y conclusión son inferencias típicas. Las metainferencias suelen representarse así:

$$\frac{\Gamma_1 \Rightarrow A_1, \ldots, \Gamma_n \Rightarrow A_n}{\Gamma \Rightarrow A}$$

Por supuesto, puede haber también metametainferencias (que tienen como premisas y como conclusión metainferencias), metametametainferencias, y así sucesivamente.[1]

Hay dos maneras de definir el concepto de *validez* de una metainferencia:[2]

(Validez global) La meta inferencia $\Gamma_1 \Rightarrow A_1, \ldots, \Gamma_n \Rightarrow A_n$ / $\Gamma \Rightarrow A$ es *globalmente válida* si y solo si siempre que sus premisas son válidas, también lo es su conclusión.

La segunda definición de validez metainferencial requiere una definición preliminar:

[1] Para expresar los diferentes niveles metainferenciales pueden usarse subíndices: $\Rightarrow_1$, $\Rightarrow_2$, etc. Afortunadamente, para los fines de este trabajo no será necesario aventurarse más allá del primer nivel metainferencial, por lo que podemos aprovechar la sencilla y clara representación gráfica que se ofrece en el texto.

[2] Una presentación detallada de estos dos conceptos de *validez metainferencial* y una discusión sobre la relevancia para la definición de la lógica de cada uno de ellos puede encontrarse en Teijeiro (2021).

(Confirmación) Una valuación confirma una inferencia Γ⇒A si y solo si no es un contraejemplo de esa inferencia.

(Validez local) La meta inferencia Γ1⇒A1, …, Γn⇒An / Γ⇒A es localmente válida si y solo si toda valuación que confirma todas sus premisas confirma también su conclusión.

Argumentaré en la sección 3 que ninguna de las dos definiciones de validez permite justificar la tesis de Barrio y Pailos. Pero antes, en la sección 2, intentaré explicar por qué Barrio y Pailos han llegado a pensar que las metainferencias pueden ser útiles para caracterizar una lógica.

2. La identidad de las lógicas

Consideremos el lenguaje proposicional L cuyas conectivas son ¬ (negación), ∧ (conjunción) y ∨ (disyunción inclusiva). Diferentes sistemas de lógica se han construido sobre la base de ese lenguaje. Una *valuación* V para L es una función del conjunto de fórmulas de L a un conjunto de valores de verdad. En el sistema de lógica clásica, **CL**, las valuaciones son funciones del conjunto de fórmulas de L al conjunto {1, 0} que respetan las conocidas tablas de verdad bivalentes:

	¬
1	0
0	1

∧	**1**	**0**
1	1	0
0	0	0

∨	**1**	**0**
1	1	1
0	1	0

Una inferencia Γ⇒A es válida en **CL** si y solo si toda valuación que asigna el valor 1 a todos los elementos de Γ, le asigna el valor 1 también a A.

Una valuación en el sistema **K3** es una función del conjunto de fórmulas de L al conjunto {1, ½, 0} que respeta las siguientes tablas de verdad trivalentes:

	¬
1	0
½	½
0	1

∧	**1**	**½**	**0**
1	1	½	0
½	½	½	0
0	0	0	0

∨	**1**	**½**	**0**
1	1	1	1
½	1	½	½
0	1	½	0

La definición de *inferencia válida* en **K3** es análoga a la de **CL**: una inferencia Γ⇒A es válida en **K3** si y solo si toda valuación que asigna el valor 1 a todos los elementos de Γ, le asigna el valor 1 también a A. El sistema **LP** tiene las mismas tablas de verdad trivalentes que **K3** y la misma definición de *valuación,* pero define de una manera diferente la validez de una inferencia: Γ⇒A es válida en **LP** si y solo si toda valuación que asigna a todos los elementos de Γ elementos del conjunto {1, ½} asigna también a A un elemento del conjunto {1, ½}. La diferencia entre **K3** y **LP** es el rol que juega el valor intermedio ½: en **K3** ½ no está entre los valores designados del sistema (los valores que la validez preserva en el sistema), mientras que en **LP** sí lo está. Suele decirse que la noción de *validez* de **K3** es *estricta* (preserva solo el valor 1) mientras que la noción de *validez* de **LP** es *tolerante* (preserva los valores 1 y ½). Hay otros dos sistemas lógicos que comparten las tablas de verdad trivalentes de **K3** y **LP** y la definición de *valuación*. Son los sistemas "híbridos" **ST** y **TS**, que combinan rasgos de las nociones de validez de **K3** y **LP**. La noción de validez del sistema **ST** es *estricta-tolerante*: Γ⇒A es válida en **ST** si y solo si toda valuación que asigna el valor 1 a todos los elementos de Γ, asigna a A un elemento del conjunto {1, ½}. La noción de validez del sistema **TS** es *tolerante-estricta*: Γ⇒A es válida en **TS** si y solo si toda valuación que asigna a todos los elementos de Γ elementos del conjunto {1, ½} asigna a A el valor 1.

En los orígenes de la lógica contemporánea era habitual identificar la lógica que codifica un sistema lógico con el conjunto de las fórmulas válidas determinado por el sistema. Esta identificación seguramente fue alentada por la estrecha conexión que existe entre las fórmulas válidas y las inferencias válidas en los sistemas lógicos más estudiados en la época, como la lógica clásica y la lógica intuicionista, y estaba muy a tono con la práctica común en la época de presentar los sistemas lógicos como sistemas axiomáticos. Pero si se considera que un sistema como **LP** es más que una mera variación algebraica del sistema de lógica clásica **CL** y se lo reconoce como la formulación de una lógica diferente, entonces la identificación de una lógica con el conjunto de sus fórmulas válidas no puede sostenerse, porque **CL** y **LP** tienen el mismo conjunto de fórmulas válidas, pero diferentes conjuntos de inferencias válidas. La existencia de sistemas como **CL** y **LP** sugiere que la lógica que codifica un sistema lógico no debe identificarse con el conjunto de sus fórmulas válidas, sino con el conjunto de sus inferencias válidas. Y esta ha sido la versión oficial en la lógica en las últimas décadas.

Sin embargo, la creación de sistemas lógicos como **ST** ha originado un debate sobre la identificación de una lógica, pues **ST** determina el mismo conjunto de inferencias válidas que **CL**, pero también hay razones para creer que ambos sistemas determinan lógicas diferentes. En particular, a diferencia de lo que ocurre con **CL**, **ST** tiene extensiones que no son transitivas.

De esta posibilidad, que admite **ST** pero no **CL**, de tener extensiones no transitivas, algunas personas han querido derivar importantes consecuencias filosóficas. Pablo Cobreros, Paul Égré, David Ripley y Robert van Rooij han demostrado que es posible extender **ST** sumando un predicado de verdad transparente que se aplique al propio lenguaje del sistema, pero sin que el sistema quede expuesto a las conocidas paradojas semánticas de la verdad, como la paradoja del mentiroso. Es precisamente la falla de la transitividad en el sistema extendido lo que bloquea la derivación de las paradojas. Apoyándose en la identificación de una lógica con un conjunto de inferencias válidas y en el hecho de que las inferencias válidas de **CL** y **ST** son exactamente las mismas, Cobreros, Égré, Ripley y van Rooij han sostenido que su proyecto constituye "una manera de sumar un predicado de verdad transparente a la lógica clásica", en contra del conocido resultado de Tarski que afirma la imposibilidad de hacer eso consistentemente.

Esta supuesta "solución" de las paradojas semánticas en la lógica clásica suena sospechosa y creo que, con razón, Barrio y Pailos la han cuestionado, argumentando que no es cierto que **CL** y **ST** determinen la misma lógica, aunque determinen el mismo conjunto de inferencias válidas. Es aquí donde entran a escena las metainferencias. Del mismo modo en que hay sistemas como **CL** y **LP** que tienen el mismo conjunto de fórmulas válidas pero diferentes conjuntos de inferencias válidas, Barrio y Pailos piensan que es posible que existan lógicas que no puedan distinguirse por el conjunto de sus inferencias válidas, pero sí por el conjunto de sus metainferencias válidas. Esto es lo que afirman en un artículo escrito junto a Damian Szmuc, en el que hablan de "la existencia de dos sistemas de lógica diferentes que tienen el mismo conjunto de inferencias válidas, pero diferentes metainferencias válidas" y en donde sostienen que "la lógica ST es un exponente de este fenómeno" (Barrio, Pailos & Szmuc, 2021), pues valida el mismo conjunto de inferencias que **CL**, pero sin validar la metainferencia de transitividad.

3. La validez metainferencial y la identidad de las lógicas

No es difícil probar que la validez global no puede usarse para justificar la creencia en "la existencia de dos sistemas de lógica diferentes que tienen el mismo conjunto de inferencias válidas, pero diferentes metainferencias válidas".

Proposición 1: si S1 y S2 son dos sistemas lógicos que determinan el mismo conjunto de inferencias válidas, entonces S1 y S2 determinan el mismo conjunto de meta inferencias globalmente válidas.

Prueba: Supongamos que S1 y S2 tienen las mismas inferencias válidas, pero existe una metainferencia $\Gamma_1 \Rightarrow A_1, \ldots, \Gamma_n \Rightarrow A_n / \Gamma \Rightarrow A$ que es globalmente válida en S1 y globalmente inválida en S2. Entonces (por la definición de validez global) en S2 las premisas de esa metainferencia, $\Gamma_1 \Rightarrow A_1, \ldots, \Gamma_n \Rightarrow A_n$, son válidas, pero la conclusión $\Gamma \Rightarrow A$ no lo es. Pero entonces (por el supuesto de partida) las premisas también son inferencias válidas en S1. Y si la metainferencia es globalmente válida en S1, entonces la conclusión $\Gamma \Rightarrow A$ debe serlo también (por la definición de validez global) contradiciendo el supuesto de que las inferencias válidas de S1 y S2 son las mismas.

La situación es diferente cuando se considera la validez local.

Proposición 2: Existen sistemas lógicos que tienen el mismo conjunto de inferencias válidas pero distintos conjuntos de metainferencias localmente válidas.

Prueba: los sistemas **CL** y **ST** tienen el mismo conjunto de inferencias válidas, pero la metainferencia (i) es localmente válida en **CL** y no en **ST** (en donde p, q y r son fórmulas atómicas):

(i) $$\frac{p \Rightarrow q,\ q \Rightarrow r}{p \Rightarrow r}$$

En efecto, no hay ninguna valuación en **CL** que verifique ambas premisas y no verifique la conclusión de (iv), pero en **ST** hay una valuación V tal que $V(p)=1$, $V(q)=½$ y $V(r)=0$ que satisface sus dos premisas sin satisfacer su conclusión, por lo que V invalida localmente a (i) en **ST**.

¿Es suficiente la proposición 2 para justificar la tesis de Barrio y Pailos ("para caracterizar una lógica no solo el nivel inferencial, sino también cada uno de los niveles metainferenciales es necesario")? No creo que lo sea.

La proposición 2 es un hecho algebraico, un resultado matemático acerca de la relación que existe entre los conjuntos de inferencias válidas de **CL** y **ST** y sus conjuntos de metainferencias localmente válidas. La tesis de Barrio y Pailos, en cambio, es una tesis de filosofía de la lógica, una tesis que sostiene que la lógica que determinan **CL** y **ST** contiene tanto a sus conjuntos de inferencias válidas como a sus conjuntos de metainferencias localmente válidas. No se puede derivar una tesis filosófica como esa directamente de un hecho algebraico como la proposición 2. Para justificar la tesis de Barrio y Pailos a partir de la proposición 2 es necesario mostrar que el concepto matemático de *metavalidez local* de un sistema no es una mera construcción algebraica, sino que juega un rol en la lógica de ese sistema. Lo que esperamos, desde el punto de vista filosófico, de la lógica de un sistema (al menos lo que yo espero de ella), es que contribuya a la comprensión de la práctica humana de dar, pedir y evaluar razones. Pero razonar, dar, pedir, evaluar razones, es algo que hacemos manipulado inferencias, no metainferencias. Si una noción de metavalidez local de **ST**, por ejemplo, aspira a jugar un rol en la comprensión de las prácticas inferenciales humanas, entonces lo mínimo que debe pedírsele es que esté conectada con la noción de validez inferencial de **ST**, que contribuya a iluminar lo que se puede inferir en el sistema. Sin esa conexión no creo que esté justificado decir que el conjunto de las metainferencias localmente válidas forman parte de la lógica de **ST.**[3]

Existe una manera muy clara de establecer esta conexión. Consiste en considerar a las metainferencias como "traducciones" de las conocidas propiedades estructurales de la noción de consecuencia. Las propiedades estructurales de una noción de consecuencia son aquellas

[3] Mi argumento se basa en una distinción que a menudo se pasa por alto, aunque Alberto Moretti (entre otras personas) la hayan expresado con claridad: "La lógica se distingue de lo que podríamos llamar sistemas de lógica. Sus tesis son del tipo de: las inferencias esquematizadas por el *modus ponens* son correctas, o del tipo de: es correcto aseverar el contenido A, y no se puede rechazar ese contenido, toda vez que se hayan afirmado los contenidos (si B, A) y B. La lógica no es una teoría matemática sino, digamos, filosófica. Los sistemas de lógica, en cambio, son estructuras matemáticamente definidas" (Moretti, 2006). El estudio de los sistemas de lógica es una ocupación (matemática) absolutamente legítima y no hay nada que impida que la investigación de los niveles metainferenciales pueda ser útil para iluminar aspectos relevantes de los sistemas de lógica (si legitiman la transitividad, por ejemplo). Pero eso no basta (creo yo) para sostener que el estudio del nivel metainferencial es útil para la lógica, entendida a la manera de Moretti como teoría filosófica de las inferencias o de las aseveraciones humanas. Para garantizar la relevancia del estudio de las metainferencias para la lógica habría que establecer la conexión entre el nivel metainferencial de un sistema de lógica y las actividades humanas de inferir o aseverar, que son el objeto de la lógica. Cuando me pregunto en el texto principal si la noción de *metavalidez* "juega un rol en la lógica del sistema" me refiero a si juega un rol en la aplicación del sistema a la aclaración de la lógica (las prácticas de la aserción y de la inferencia humanas), no a si juega un rol en la determinación de las propiedades de un sistema de lógica (lo que ciertamente hace). Un sistema de lógica puede exhibir muchas propiedades matemáticas interesantes que nada tienen que ver con su aplicación al estudio de la aserción y la inferencia humanas.

que, como la reflexividad, la monotonía o la transitividad, no dependen del vocabulario lógico. Decimos, por ejemplo, que una relación de consecuencia lógica $\vDash$ de un sistema lógico es transitiva si y solo si $\vDash$ satisface la siguiente condición (para formulas cualesquiera A, B y C del lenguaje del sistema):

(ii) Si $A \vDash B$ y $B \vDash C$, entonces $A \vDash C$

La transitividad así definida es, pues, una propiedad de clausura en el conjunto de las inferencias válidas de un sistema lógico. Si convenimos en usar la notación $\vDash A \Rightarrow B$ para indicar que la inferencia $A \Rightarrow B$ es válida de acuerdo a la relación de consecuencia lógica $\vDash$, entonces $A \vDash B$ y $\vDash A \Rightarrow B$ dicen lo mismo y la condición (ii) puede reformularse como:

(iii) Si $\vDash A \Rightarrow B$ y $\vDash B \Rightarrow C$, entonces $\vDash A \Rightarrow C$

Tenemos, pues, dos maneras equivalentes de formular la noción de transitividad. Bario, Pailos y Szmuc sugieren una tercera: "en lugar de entender las características estructurales como propiedades de clausura del conjunto de inferencias válidas de una lógica, nosotros las entendemos como metainferencias" (Barrio, Pailos & Szmuc, 2021). La idea es que la transitividad del sistema podría expresarse por medio de la validez de la siguiente metainferencia:

(iv) $$\frac{A \Rightarrow B,\ B \Rightarrow C}{A \Rightarrow C}$$

Si las metainferencias localmente válidas de **ST** fueran "traducciones" al lenguaje inferencial de las propiedades estructurales del conjunto de las inferencias válidas de **ST**, entonces la garantía de que el conjunto de metainferencias localmente válidas es parte de la lógica de **ST** sería absolutamente inobjetable. El problema es que la metavalidez local no puede entenderse de ese modo. Como hemos visto, la metainferencia (i) es localmente invalida en **ST**, pero (v) y (vi) son verdaderas en **ST**:

(v) Si $p \vDash q$ y $q \vDash r$, entonces $p \vDash r$
(vi) Si $\vDash p \Rightarrow q$ y $\vDash q \Rightarrow r$, entonces $\vDash p \Rightarrow r$

La validez local de **ST** no puede entenderse como una "traducción" de las propiedades estructurales de la validez inferencial de **ST**. Y sin esa conexión la noción de metavalidez local tampoco puede justificar la tesis de Barrio y Pailos. Los niveles metainferenciales de un sistema de lógica no son parte de la lógica de ese sistema, cuando la validez metainferencial se entiende localmente.[4]

La metavalidez global, por otra parte, sí puede ser entendida como una traducción al lenguaje inferencial de las propiedades estructurales. Las propiedades estructurales son propiedades de clausura del conjunto de inferencias válidas y la metavalidez global es preservación de la validez inferencial. Cuando la metavalidez se entiende globalmente, pues, los niveles metainferenciales sí pueden ser considerados niveles lógicos. Pero ya hemos visto que la metavalidez global no puede hacer distinciones que no estén presentes al nivel inferencial. De lo que se sigue que ni la metavalidez global, ni la metavalidez local, pueden justificar la tesis de Barrio y Pailos.[5]

[4] Un/a referí anónimo/a me ha llamado la atención sobre la conexión que suele establecerse entre la noción local de metavalidez y la lógica modal y se pregunta si mi ataque a la noción local no implicaría un ataque a la lógica modal. No creo que lo haga, pues mi punto principal es que las metainferencias no tienen relevancia para el estudio de la lógica (clásica, intuicionista, modal o cualquier otra) aunque puedan tenerla para el estudio matemático de los diferentes sistemas de lógica que se usan para sistematizar esas lógicas (para la diferencia véase la nota 3).

[5] Mi intención al escribir este trabajo era poner en cuestión la tesis de Barrio y Pailos, no ofrecer mi propia opinión sobre la identidad de las lógicas. Pero un/a referí anónimo/a mostró un amable interés en conocer mi opinión sobre este asunto y no sería de buen gusto negarse. Dicho brevemente, creo que la lógica, a diferencia de la teoría de los sistemas de lógica (para la diferencia véase la nota 3), es esencialmente una *teoría semántica*. Entiendo "teoría semántica" de una manera amplia, para incluir no solo la semántica de condiciones de verdad sino la semántica inferencialista, la semántica algebraica, la semántica de teoría de juegos y cualquier otro enfoque de la semántica que se quiera proponer. En mi opinión, las inferencias válidas de una lógica (y sus metainferencias válidas en todos los niveles), son consecuencia de la teoría semántica que define esa lógica. Es al nivel de la teoría semántica que se evalúa la calidad de una teoría lógica, es la semántica la que pone en contacto la lógica con la práctica lingüística de dar y pedir razones que es su objeto. La semántica bivalente de **CL** y la semántica trivalente de **ST** determinan lógicas diferentes, el hecho de que sus inferencias válidas sean las mismas no tiene mayor importancia para el tema de la identidad de la lógica. De hecho, cuando se identifica la lógica con una teoría semántica, la coincidencia de las inferencias válidas de diferentes lógicas es moneda corriente: piénsese, por ejemplo, en la coincidencia de las inferencias válidas de la lógica intuicionista determinada por condiciones de verdad por medio de la teoría de modelos y la lógica intuicionista determinada por una semántica inferencialista. La intención de definir una lógica por el conjunto de sus inferencias válidas o por el conjunto de sus metainferencias válidas me parece que no es más que un síntoma de la confusión (a la que hice referencia en la nota 3) entre la lógica y la teoría (matemática) de los sistemas de lógica.

Bibliografía

Barrio, E. A., & Pailos, F. (2021). Why a logic is not only its set of valid inferences. *Análisis Filosófico, 41*(2), 261-272. https://doi.org/10.36446/af.2021.461

Barrio, E. A., Pailos, F., & Szmuc, D. (2021). (Meta)inferential levels of entailment beyond the Tarskian paradigm. *Synthese, 198*(Suppl 22), 5265–5289. https://doi.org/10.1007/s11229-019-02411-6

Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind, 122*(488), 841-866. https://doi.org/10.1093/mind/fzt110

Moretti, A. (2006). Logica y semantica. *Revista de Filosofía, 31*(2), 31-43. https://revistas.ucm.es/index.php/RESF/article/view/RESF0606220031A

Teijeiro, P. (2021). Strength and stability. *Análisis Filosófico*, *41*(2), 337-349. https://doi.org/10.36446/af.2021.459

Recibido el 14 de julio de 2023; aceptado el 2 de noviembre de 2023.