Lógica Natural vs. Lógica Formal
Un debate inaplazable

Enrique Alonso

Introducción

En esta contribución pretendo repasar los fines y objetivos de dos formas muy distintas de entender el objetivo de la Lógica: una canónica, la defendida por la Lógica Formal o matemática del siglo XX, y otra que se abre paso gracias, ante todo, a la explosión de la IA-generativa[1], aunque existiera antes incluso de la popularización de esta última.

Presentaré las que, a día de hoy, son las principales escuelas que desarrollan su actividad en el ámbito de la Lógica Natural, para exponer algunas de mis críticas al respecto. Finalmente dedicaré algún espacio a analizar si la Lógica Natural podría aspirar a ser un reemplazo razonable de la Lógica Formal en los planes de estudio de los Grados de Humanidades.

Lógica Formal vs. Lógica Natural

Como parece obvio, toca fijar qué vamos a entender aquí por *Lógica Formal o matemática* y sobre todo por *Lógica Natural.* Definir la primera de ellas no requiere especial esfuerzo ya que existe una larga tradición al respecto.

> 1. La *Lógica Formal* es la disciplina que estudia la noción de *consecuencia* desde el punto de vista de la estructura de los enunciados que constituyen las premisas y la conclusión.

En cuanto a la Lógica Natural, la cuestión es más disputada, pero existe un cierto consenso que podría resumirse en las siguientes palabras de Lawrence S, Moss: "Por lógica natural entiendo el estudio de la inferencia en el lenguaje ordinario, realizado tan cerca como sea posible de sus 'formas explícitas' " (Moss, 2010, p. 84).

Por tanto podríamos decir que:

> 2. La *Lógica Natural* se centra en el estudio de las inferencias que tienen lugar en el discurso o razonamiento ordinario intentando respetar la forma de tales inferencias en el mismo lenguaje en el que tienen lugar.

Esta definición constituye una suerte de declaración de intenciones con respecto

1 La IA-generativa se apoya en la tecnología del *machine learning* y está a su vez en el modelo de *redes neurales.* Se trata este último de un paradigma que no gozó de la atención de la comunidad del mismo modo que lo hizo el orientado al diseño de software específico a través de lenguajes de alto nivel, como LISP, por ejemplo. Queda claro qué modelo parece finalmente el ganador de esta competición.

a la Lógica Formal. El análisis de la consecuencia lógica y, en consecuencia, de las razones por las que obtenemos conclusiones a partir de ciertas piezas de información no se *explicarían* ya recurriendo a la Lógica Formal, sino a las herramientas típicas del estudio de los lenguajes naturales: Gramática generativa, categorial, etc.

Como veremos más adelante, esta orientación reconecta de una manera nada disimulada con la Silogística aristotélica, aunque sin rechazar por ello mismo las herramientas de la Matemática, la Gramática y la Lógica formales (Van Benthem, 2008, pp. 13-14), pero ya habrá momento de hablar de ello.

Es fácil entender que la tensión principal entre estas dos maneras de concebir el objetivo de la disciplina de la Lógica reside en la conveniencia, o no, de considerar un lenguaje distinto del ordinario para estudiar la relación de consecuencia. En el caso de la Lógica Formal la decisión es, al menos desde la decisiva contribución de Frege, la de introducir un lenguaje formal capaz de representar adecuadamente la estructura de los enunciados que justifican o soportan una inferencia.

> Estas discrepancias con lo tradicional encuentran su justificación en que la lógica, hasta ahora, siempre se ha ajustado muy estrechamente al lenguaje y a la gramática. En especial, creo que la sustitución de los conceptos de sujeto y predicado por los de argumento y función, se acreditará con el tiempo (Frege, 1972, p. 4).

Según esta tradición, que es la que nos acompaña hasta el presente, la tarea del lógico difiere sustancialmente de la del gramático precisamente en este punto. El lógico puede y debe generar un lenguaje que cuente con las categorías apropiadas para representar la estructura de un enunciado, al menos aquellas que permiten dar cuenta de sus relaciones inferenciales. Dichas categorías estarán presididas por la dicotomía fundamental *argumento y función* y no por la tradicional de *sujeto y predicado*. En consecuencia, la Lógica Formal puede ingeniar lenguajes cuya estructura y aspecto disten de los del lenguaje ordinario tanto como sea preciso. Lo único que se exige es que cumpla con la tarea de explicar la inferencia a través de los elementos estructurales identificados.

Esta es la razón por la que en un curso de Lógica Formal elemental sostengamos cosas como la siguiente:

3. La estructura del enunciado:

Todos los hombres son mortales es

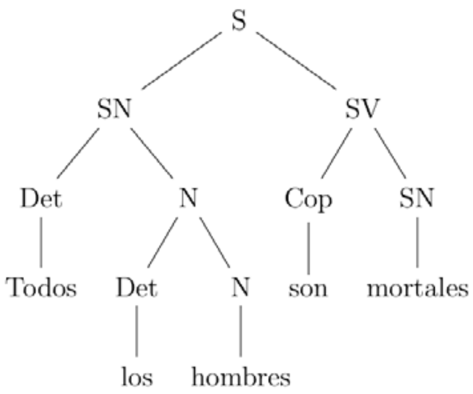
$\forall x\ (Hx \rightarrow Mx)$

Sin aclarar, en ocasiones, que esa pretendida estructura, casi *inmanente*, no es otra cosa que la que resulta de un proceso de traducción a un lenguaje previamente definido, el de la Lógica de Primer Orden –FOL– en este caso.

Desde la Lógica Natural se intenta evitar este tipo de maniobra reemplazando las herramientas de análisis lógico por las propias de la Gramática, en particular la Generativa y la Categorial.

Una forma común de expresar y analizar en Lógica Natural la estructura de un enunciado como el anterior podría ser la propia, por ejemplo, de la Gramática Generativa:

4.



Pero podríamos haber considerado igualmente la gramática categorial o el cálculo lambek, por ejemplo (Van Benthem, 1991).

Todas estas formas de analizar la estructura de un enunciado basadas en la Gramática de los lenguajes ordinarios han tenido y tienen uso, en los estudios actuales en Lógica Natural. ¿Pero es realmente este giro el que caracteriza a la Lógica Natural frente a la tradición de la Lógica Formal? Es posible que, en su origen, constituyera una idea interesante promovida por los muchos fracasos que la Lógica Formal acumulaba y acumula en el estudio del discurso ordinario. Sin embargo, el paso de los años y la experiencia adquirida han ido debilitando sustancialmente ese aspecto tan aparentemente definitorio de la Lógica Natural. Se ha generado una cierta tensión entre la estructura superficial de nuestras expresiones e inferencias y las herramientas que se hacen necesarias para explicar las relaciones inferenciales entre premisas y conclusión. Estas últimas pueden llegar a ser realmente complejas imitando o reproduciendo, en definitiva, la presunta artificialidad que el modelo de la Lógica Formal impone en su análisis de las inferencias que tienen lugar en el lenguaje ordinario. Pero de todo ello hablaremos con algo más de detalle en breve.

En definitiva, el balance entre una y otra iniciativa ya no es tan claro en este punto. El uso de lenguajes formales artificiales para el estudio de la inferencia ordinaria en la Lógica Formal, en el caso de la Lógica Natural quedaría compensado por el empleo de herramientas complejas basadas en la Gramática Categorial, el Cálculo Lambek o la Teoría de Tipos, por citar solo algunos ejemplos.

Sin embargo, este comentario no intenta ocultar o minimizar las dificultades de la estrategia seguida por la Lógica Formal que se basa, en lo que podríamos denominar, el principio de *traducción como análisis*. Es decir, el diseño de un lenguaje formal al que traducir los enunciados del lenguaje ordinario como forma de exponer su estructura.

Es evidente que esta estrategia no ha logrado satisfacer de forma simultánea estos dos requisitos:

5.
i. Obtener un lenguaje suficientemente potente desde el punto de vista expresivo y
ii. Diseñar herramientas que permitan decidir la aceptabilidad de los argumentos expresables en dicho lenguaje.
Y ello no porque estemos aún a la espera de alguna aportación trascendental, sino por la misma imposibilidad de la empresa.[2]

La estrategia habitual en Lógica Formal, y que hemos practicado durante un siglo largo, ha pasado siempre por estudiar pequeños fragmentos de las inferencias ordinarias generando lenguajes en los que quepa representar su estructura. Es la vieja conocida táctica de *divide et impera*. Es así como se llegaron a establecer las divisiones que aún están vigentes en el estudio de los sistemas formales: lenguaje sentencial, Primer Orden, Orden Superior, lenguaje modal, lenguaje temporal, deóntico, etc. La idea subyacente consistía en integrar finalmente todos estos estratos en un lenguaje universal obtenido por agregación. Pero todos sabemos que esto no es posible. Cada vez que se intenta generar un sistema complejo capaz de integrar varios de estos niveles se obtiene una suerte de modelo *frankstein* apenas manejable y con muy pocas ventajas técnicas que ofrecer. Esta es la maldición de la Lógica Formal contemporánea: cada logro en capacidad expresiva implica una consiguiente pérdida en control computacional, es decir, en capacidad para distinguir las inferencias correctas de aquellas que no lo son.[3]

Lo que procede en este momento es una discusión más detallada sobre el modo en que la Lógica Natural lidia con estas mismas dificultades.

La escuela holandesa

A finales de la década de 1980, un grupo informal[4] liderado por Johan van Benthem en la Universiteit van Amsterdam dio inicio a una serie de estudios cuyo rasgo fundamental era el reemplazo de los lenguajes formales, estándar en el estudio de la inferencia, por herramientas procedentes de la Gramática formal (Van Benthem, 1990).

2 Me refiero, claro está, a los teoremas de limitación sobradamente conocidos y estudiados desde su formulación durante el segundo tercio del siglo XX.

3 Esta maldición, expuesta en muchas ocasiones como una suerte de balanza, ha sido material común en los cursos impartidos por María Manzano y yo mismo en el Máster de Lógica y Filosofía de la Ciencia impartido en España a través de un consorcio de varias de las principales universidades del país.

4 Dicho grupo, conocido entonces como *Institute for Logic, Language and Information*, daría lugar en la década siguiente a la creación del *Institute for Language, Logic and Computation* que permanece hasta nuestros días.

El trabajo más destacado dentro de esa pequeña comunidad fue, durante algún tiempo, la tesis doctoral defendida por Victor Sánchez-Valencia en 1991 (Sánchez-Valencia, 1991), en la que se ofrece un elaborado tratamiento de las inferencias basadas en la dicotomía *monotonía/antitonía* existente entre los términos léxicos de nuestros enunciados. Lo explicaré con un ejemplo absolutamente elemental:

6.

Todos los griegos son virtuosos.
Los atenienses son griegos.
Por tanto,
Los atenienses son virtuosos.

Como se puede ver, se trata de una forma elemental de silogismo, en particular una instancia de la figura conocida como *Bárbara*.

La explicación contemporánea de la corrección de dicho argumento se obtiene a partir de su traducción al lenguaje de FOL y de la demostración mediante las herramientas apropiadas de que el argumento expuesto en [7] es formalmente correcto.:

7.

$\forall x(Gx \rightarrow V\, x)$
$\forall x(Ax \rightarrow Gx)$
Por tanto,
$\forall x(Ax \rightarrow V\, x)$

Para Sánchez-Valencia existe una forma mucho más aquilatada de justificar este tipo de inferencias. Sin ser totalmente fieles a su formalismo, se podría exponer del siguiente modo:[5]

8.

Every G is V
$[A] \subseteq [G]$
Every A is V

Lo que tiene lugar en este tipo de inferencias, es un silogismo en *Bárbara*. En este caso en particular, se trata de la substitución de un término léxico por otro en función de dos piezas de información:

i. la relación entre el término sustituido y el sustituyente en términos de extensión: $[A] \subseteq [G]$, y

5 En [Sánchez-Valencia, 1991, p.16 y ss] se ofrece una traducción de las figuras del silogismo a la notación prevaleciente en su trabajo.

ii. la posición del término léxico sustituido, *G* en este caso, en el enunciado en el que figura.[6]

A su vez, los puntos i. y ii. dependen de la partícula cuantificacional, cuantor, dominante en cada expresión. En el caso, por ejemplo, de que el término léxico a ser sustituido en [8] fuera *V*, la relación léxica con el término sustituyente, llamémosle [C] debería ser esta: [V] ⊆ [C] , dando lugar a:

9.

Every G is V
[V] ⊆ [C]
Every G is C

Es decir, los términos léxicos que intervienen en una expresión cuantificacional universal se comportan exigiendo relaciones inversas con el término reemplazante según la posición que ocupan. El término *G* puede ser reemplazado por uno de menor extensión, mientras que el término *V* lo tendrá que ser por uno de menor extensión. Esto es lo que nos lleva a afirmar que el comportamiento de *V* en "Every G is V" es *monótono*, mientras que el de "G" es *antítono*. Este rasgo parece, pues, estar asociado a la partícula cuantificacional en primera instancia, dando lugar, en el caso del cuantor existencial, a una conducta monótona en ambos términos:

10.

Some G is V
[G] ⊆ [A]
Some A is V

y

11.

Some G is V
[V] ⊆ [A]
Some G is A

A este comportamiento de un término léxico dentro de una expresión se le denomina *su polaridad* y se representa como (↑) en el caso de la monotonía, y como (↓) en el caso de la antitonía. En consecuencia se puede sostener ahora lo siguiente:

12.

i. $Q_{\forall}$(A, B) = (↓↑), estando A y B afirmados.
ii. $Q_{\exists}$(A, B) = (↑↑), estando A y B afirmados.

6 El término al que aquí nos hemos referido como el término sustituido es lo que en la tradición silogística se ha venido denominando término medio.

Es decir, (↓↑) representaría la polaridad de los términos léxicos sobre los que el cuantor universal actúa, mientras que (↑↑) supondría lo propio con respecto al cuantor existencial. Obsérvese que esta forma de tratar a los cuantores difiere, y de manera sustancial, de la que es típica en Lógica formal. Ahora se interpretan, básicamente, como operadores binarios que toman como argumentos términos léxicos, y no como constantes lógicas que operan sobre fórmulas. Creo que a nadie se le escapa que esta maniobra, motivada por el análisis subyacente en el lenguaje ordinario, complica, y mucho, el manejo de expresiones que en el lenguaje de FOL resultan elementales. [7]

Llegados a este punto resulta evidente que el objetivo principal de la propuesta holandesa tiene mucho que ver con la identificación de la polaridad de cada término léxico en contextos de complejidad creciente. Es notable el esfuerzo hecho, a este respecto, y también la complejidad de los algoritmos diseñados a tal efecto. La idea detrás de esta concepción de la Lógica Natural no puede ser más sencilla y al mismo tiempo más sugerente. Las inferencias que los seres humanos hacemos en contextos reales se basan en una operación muy simple: reemplazar términos léxicos por otros cuya relación mutua se conoce,[8] para dar lugar a un enunciado nuevo que actúa como conclusión del primero. Para ello nos basta con conocer la polaridad del término léxico en cuestión, en la expresión en la que va a ser reemplazado.

Obsérvese que, bajo esta perspectiva, una inferencia natural es tan solo el tránsito de un enunciado a otro que se obtiene del inicialmente dado al reemplazar un término léxico dentro de su estructura. Ese reemplazo viene legitimado por una pieza de conocimiento, tal vez implícita, que garantiza que la substitución puede ser hecha. Como puede verse, es algo que difiere, y de forma no menor, de la concepción tradicional de la consecuencia lógica. En Lógica Formal, las inferencias basadas en meras transformaciones léxicas nunca han tenido un tratamiento diferenciado.[9] El objetivo era, por el contrario, caracterizar la conducta de las constantes lógicas descritas en el lenguaje. Es así como podemos entender algunas de las polémicas más características del siglo pasado, como por ejemplo aquella que priorizaba los cálculos basados en *Deducción Natural*, *Tablas analíticas* o *Cálculo de Secuentes* sobre los *Sistemas axiomáticos*. En los primeros, las reglas describen de forma independiente el comportamiento de cada constante, mientras que en los *Sistemas axiomáticos* se hace, más bien, un estudio de sus interacciones.[10]

En definitiva, todo indica que estaríamos descubriendo una tensión nada disimulada entre una suerte de *Lógica de términos* y una *Lógica de constantes*,

7 Por ejemplo, el manejo de relaciones de cualquier ariedad con cuantores anidados.
8 Determinada en este caso por la relación de inclusión.
9 Ni siquiera la regla de substitución uniforme puede ser entendida en ese sentido.
10 De ahí, por ejemplo, la polémica en torno a la conectiva *plonk* (Belnap, 1962).

o alternativamente entre una *Lógica categoremática* y otra *Lógica sincategoremática*. La Lógica Natural se podría considerar, entonces, como un intento de recuperar la tradición del estudio de las relaciones entre los términos léxicos y, por lo tanto, constaría como un ejemplo de *Lógica categoremática*. Lo que quiero hacer explícito con estas consideraciones es, en primer lugar, que Lógica Natural y Lógica Formal difieren en aspectos sustanciales que nos llevan muy atrás en el orden de las decisiones adoptadas a lo largo del tiempo. En este sentido, la Lógica Natural sería pre-fregeana . Y en segundo lugar, que sus diferencias no se expresan por completo en el hecho de adoptar un lenguaje formal o las herramientas de la Gramática como instrumento de análisis. En ambos casos, se trata de aspectos que nos llevan más atrás en el tiempo del hito fundacional representado por Frege. Pero creo que el primero de ellos es, en cierto sentido, más radical y definitorio que el segundo.

Veamos si estas características se aprecian igualmente en la otra gran escuela dentro del ámbito de la Lógica Natural.

¿Dónde interviene la Gramática entonces? Como ya he dicho, en el estudio de la polaridad de los términos léxicos en expresiones complejas. El algoritmo descrito por Sánchez-Valencia (Sánchez-Valencia, 1991, cap. 5 y 6) reposa en un ingenioso empleo de la Teoría de Tipos sobre una estructura básica aportada por el Cálculo Lambek. Debo indicar, como ya insinué páginas atrás, que el uso de las reglas, de lo que Sánchez-Valencia denomina como un *System of Natural Logic*, no es sencillo ni tiene por qué serlo. Sí es importante indicar que con dicho sistema, Sánchez-Valencia consigue ir más allá, tanto de la Silogística como de la Lógica de Primer Orden. Se supera la Silogística al incorporar de forma natural verbos transitivos, típicamente caracterizados como relaciones binarias en LPO, algo que la Silogística no supo manejar adecuadamente. Y va más allá de la Lógica de Primer Orden al incorporar de manera bastante natural los cuantores generalizados, al menos los más característicos: *MOST* y *FEW* (Sánchez-Valencia, 1991, p. 136).

La Escuela americana

En este caso las referencias históricas me resultan menos familiares, por lo que no podría ofrecer un lugar concreto o un acontecimiento particular como actos fundacionales de esta escuela. Sin embargo, resulta relativamente fácil identificar a una serie de autores reconocibles por sus colaboraciones e intereses mutuos. Entre ellos se encuentran al menos los siguientes: B. MacCartney, T.F. Icard, y L.S. Moss.

El punto de partida de sus análisis lo constituye, en buena medida, los trabajos de la escuela holandesa en torno a la monotonía y en particular el estudio que Sánchez-Valencia desarrolla sobre ese tema (Sánchez-Valencia, 1991). En este sentido, la posición de la Escuela americana podría considerarse como un intento de ampliar el alcance que el estudio de la monotonía ofrece en el contexto de una Lógica de términos o categoremática, como aquí la hemos denominado.

Si tomamos la teoría de clases como marco de estudio para el análisis de los términos léxicos, no tenemos por qué limitarnos al estudio de relaciones $A \subseteq B$ y $A \supseteq B$ que son, precisamente las de monotonía y antitonía tomando el término léxico A como referencia.

En la escuela americana se señalan hasta siete relaciones primitivas que son las que se indican en el siguiente cuadro (MacCartney and Manning, 2009, p. 143):

Símbolo	Denominación	Ejemplo	Definición
$x \equiv y$	equivalencia	niño $\equiv$ chaval	x=y
$x \sqsubset y$	monotonía	cuervo $\sqsubset$ pájaro	$x \subset y$
$x \sqsupset y$	antitonía	europeo $\sqsupset$ francés	$x \supset y$
$x^\wedge y$	negación	humano $^\wedge$ no humano	$x \cap y = \emptyset \wedge x \cup y = U$
$x \mid y$	alternancia	gato $\mid$ perro	$x \cap y = \emptyset \wedge x \cup y \neq U$
$x \smile y$	covertura	animal $\smile$ no humano	$x \cap y \neq \emptyset \wedge x \cup y = U$
$x \# y$	independencia	hambriento $\#$ hipopótamo	el resto de los casos

Tabla 1: Conjunto $\mathfrak{B}$ de relaciones semánticas básicas

Este listado se obtiene de manera combinatoria y no tiene mucho sentido. Entretenernos en ello en este momento. Ahora tomaremos un ejemplo ofrecido por MacCartney y Manning (MacCartney and Manning, 2009, p. 145) para aclarar el modo en que se puede operar con estas relaciones básicas. Vamos a considerar tres términos léxicos: *fish*, *human* y *nonhuman*, para analizar algunas relaciones entre ellos tomados de dos en dos.

13.

i. *fish*|*human*

ii.human^nonhuman

iii.fish ⊏ nonhuman

Estos autores consideran que lo que está teniendo lugar, al relacionar términos léxicos de esta manera, es una suerte de combinatoria que analiza las relaciones entre términos léxicos como los descritos en la tabla anterior. No se trata de las propias relaciones entre términos léxicos, sino de las relaciones entre esas mismas relaciones. Esto da lugar a la siguiente definición:

14.

$$R \bowtie S =^{def} \{\langle x, z\rangle : \exists y(\langle x, y\rangle \in R \wedge \langle y, z\rangle \in S)\}$$

Es decir, las relaciones | y ^, que serían en este caso, *R* y *S*, producen otra relación ⊏ que se da entre los términos léxicos *x* y *z*, que en el ejemplo resultarían ser *fish* y *non-human*. Si se analiza con detalle las expresiones listadas en [13] se observa que estas dan lugar a una inferencia que puede expresarse del siguiente modo:

14.
Ningún pez es un ser humano.
Cualquier individuo, o bien es humano o bien no lo es,
Por tanto,
Todo pez es no humano
Es inevitable encontrar similitudes con la Silogística una vez más, solo que en esta ocasión disponemos de recursos expresivos que van más allá de ella.

El siguiente paso consiste en describir tres tipos de posibles actuaciones sobre los términos léxicos de un enunciado:

15.

i. Inserción $(ins(A_i))$

ii. Borrado $(del(A_i))$

iii. Substitución $((sub(A_i, A_j)))$

Veamos ahora cómo funciona este sistema con un ejemplo elemental que, no obstante, simplificaré un poco más, ofrecido MacCartney y Manning (MacCartney and Manning, 2009, p. 153):

15.

i	x_i	e_i	$\beta(x_0, x_i)$
0	Stimpy is a cat		
1	Stimpy is a dog	sub(cat, dog)	$\mid$
2	Stimpy is not a dog	ins(not)	$\sqsubset$
3	Stimpy is not a poodle	sub(dog, poodle)	$\sqsubset$

En la 4 columna de la primera línea se lee | lo que indica que la relación $\beta(x_0, x_1)$, entre “Stimpy is a cat” y “Stimpy is a dog” es de *alternancia* –véase la tabla de las relaciones elementales–. A su vez, “Stimpy is a dog” se ha obtenido a partir de “Stimpy is a cat” por medio de la operación *sub(cat, dog)*. En la línea 2, lo que vemos es que “Stimpy is not a dog” se obtiene a partir de "Stimpy is a dog" mediante la operación *ins(not)* sobre esta última. Este movimiento provoca que la relación entre "Stimpy is a cat" y "Stimpy is not a dog" es $\sqsubset$. Por último, llegamos en la línea 3 a afirmar que la relación entre x_0, es decir, "Stimpy is a cat" y x_3, "Stimpy is not a poodle" es, de nuevo, $\sqsubset$. Esto supone que entre x_0 y x_3 existe una relación de consecuencia.

El mecanismo procede mediante una serie de pasos –*edits* (MacCartney and Manning, 2009, p.152)– en los que la relación entre el enunciado original, x_0 y los subsiguientes va variando en función de las relaciones entre los términos léxicos que se ven implicados en cada caso. Lo realmente importante es que la relación entre el primer enunciado x_0 de la serie y el último, sea este x_n, venga dada por $\beta(x_0, x_n) = \sqsubset$, lo que, como acabo de indicar, introduce una relación de consecuencia entre ambos enunciados.

Creo que el resto de la historia es bastante fácil de imaginar. Incorpora parte del modelo holandés de una forma natural y conecta, además, con el ámbito de la IA de una forma mucho más directa. El tipo de cadenas como la que se ilustra en [15] reflejan muy bien el género de manipulaciones léxicas que intentan implementarse en los diversos modelos de IA-gen de los que tanto se habla ahora. Más en concreto, son el fundamento de los sistemas de *NLP –Natural Language Processing*– que sirven, a su vez, para el diseño de las redes neurales en que se basan las IA-gen.

Es fácil entender el modo en que cadenas como la que hemos analizado en [15] son capaces de producir contenido nuevo a partir de la información aportada por un usuario. No se trata de un análisis del modo en el que nuestra cognición se comporta con respecto a ciertas estructuras inferenciales, sino de la forma en que generamos discurso consistente a partir de una información previamente dada. Tendremos ocasión más adelante de desarrollar este punto. Como puede verse, el sistema desarrollado resulta ser igualmente una *Lógica de términos*, pero difiere sustancialmente del modelo holandés en aquello que se entiende por un argumento. Dedicaremos el siguiente apartado a discutir por extenso este punto.

Sobre la noción de argumento

Resulta extraño pensar que una noción tan elemental como la de *argumento* pueda verse tensionada luego de haber gozado de un largo periodo de consenso dentro de la comunidad científica. Pero tal y como dije líneas atrás, estamos ante movimientos que replantean algunas de las decisiones más características de la fundación de la Lógica contemporánea.

En Lógica formal entendemos por argumento lo siguiente:

16.

$A_1, \ldots A_i, \ldots A_n \vdash B$

donde $A_1, \ldots A_i, \ldots A_n$, B son fórmulas previamente definidas dentro de un lenguaje formal dado. $A_1, \ldots A_i, \ldots A_n$ son las premisas y B la conclusión. Si B se *sigue* de las premisas $A_1, \ldots A_i, \ldots A_n$ entonces decimos que el argumento es correcto, en caso contrario incorrecto.[11]

Para el modelo holandés la cuestión parece ser muy distinta. Tal y como he dicho al presentar esta escuela, una inferencia natural parece estar basada en el reemplazo de un término léxico por otro en una expresión, de tal modo que se establezca entre dichas expresiones una relación de consecuencia. Para ello se precisan dos cosas: información acerca de la relación entre los términos léxicos involucrados –monotonía o antitonía– y la polaridad que presenta el término reemplazado en la expresión dada inicialmente. Un argumento, visto desde esta perspectiva es, de hecho, un argumento correcto. Visualicémoslo con un esquema apropiado:

17.

$$\frac{\varphi(A_1,A_2 ... A_i ... An)}{\varphi(A_1,A_2 ... A_j ... An)}\ (A_i \subseteq A_j)$$

$$\frac{\varphi(A_1,A_2 ... A_i ... An)}{\varphi(A_1,A_2 ... A_j ... An)}\ (A_i \supseteq A_j)$$

Para empezar, aquí no hablamos de premisas, sino de una única expresión inicial. La información relativa a la relación entre los términos léxicos bajo escrutinio –($A_i \subseteq A_j$) o ($A_i \supseteq A_j$)[12] – no parece contar como una premisa más, sino como una especie de *garantía* o *justificación* de la conclusión. Esta es la razón por la que se puede sostener que esta idea de argumento presenta, en realidad, lo que sería un *argumento correcto*, y no un argumento en general, como ocurre en Lógica Formal.

Esto mismo ocurre con las figuras del Silogismo, en las que se listan argumentos correctos formados por dos premisas y una conclusión con una estructura fija, dada por la distribución de los términos léxicos en las premisas y en la conclusión, y la tipología de los juicios, proposiciones, presentes.

¿Qué ocurre con el modelo americano? A diferencia de lo que sucede en la Escuela holandesa, aquí podemos tomar en consideración una cantidad finita, pero ilimitada en principio, de enunciados. Obsérvese que no he hablado de premisas ya que, en realidad, solo la primera expresión de una cadena como la que figura en [15] puede considerarse como tal. Si damos forma a estas intuiciones, obtenemos lo siguiente:

11 Al emplear la expresión correcto, pretendo evitar tomar partido por una interpretación sintáctica o semántica de la corrección argumental.

12 Estos términos léxicos pueden corresponder en LPO a relaciones de cualquier ariedad. Téngase en cuenta que en términos puramente conjuntistas, la monotonía/antitonía se predica de cualquier tipo de clases, ya estén formadas por individuos o n-tuplas de ellos.

18.

$\langle x_0, x_1, ...x_i, x_n \rangle$, donde

i. $x_0, x_1, ...x_i, x_n$, son enunciados.

ii. $x_{i+1} = edit(x_i)$, donde un *edit* de un enunciado x_i es otro obtenido a partir de este mediante el uso de las operaciones $ins(A_i)$, $del(A_i)$, $sub(A_i, A_j)$ sobre términos léxicos y

iii. $\beta(x_0, x_n) = \sqsubset$

Es decir, un argumento sería una secuencia de *edits* tales que entre el primer enunciado x_0 y el último, x_n, existe una relación de consecuencia. Téngase en cuenta que una de estas secuencias no constituye en absoluto una prueba en sentido clásico. Las operaciones admitidas para producir nuevos elementos en la secuencia no son preservadoras de la verdad y por tanto, x_{i+1} no tiene por qué ser consecuencia de x_i. Este requisito solo se exige entre x_0 y x_n.

Si se compara la noción de argumento del modelo americano [18], con la del modelo holandés [17], es inmediato observar un claro sesgo computacional en el primero de los casos. La noción de *edit*, central en esta escuela, está claramente orientada a la producción de contenido bajo determinados criterios de control. Se trata, básicamente, de *generar* nuevos enunciados a partir de uno dado, de modo que, en todo momento, se conserve noticia de la relación entre el primero y el último de los producidos. Esa relación viene dada por la función $\beta(x_0, x_n)$ que es la que garantiza en última instancia si x_0 es o no consecuencia de x_n, pero no es menos relevante que x_n guarde cualquier otra posible relación con alguno de los x_i en la secuencia de *edits*. Lo sustancial es que podamos conocerla en cada paso. Llegados a este punto, lo que procede es dar paso a la ingeniería del conocimiento diseñando algún modelo de *Natural Language Inference –NLI–* que pueda finalmente ser implementado en una aplicación funcional.

En comparación con tales objetivos, el modelo holandés parece mucho más centrado en el análisis formal y filosófico de la inferencia natural, posición que personalmente comparto y que será la que desarrollaré a continuación.

Una concepción extensa de la *inferencia natural*

Reemplazar los términos léxicos en una expresión dada permite, si se conoce la relación precisa entre el término substituido y el substituyente, obtener consecuencias lógicas a partir de la expresión inicial. Ciertamente se trata de una idea poderosa que dice mucho de las estrategias que empleamos en la vida cotidiana para poner en uso ciertas habilidades lógicas preformales. De hecho, creo que es justo reconocer que estaríamos ante un modelo extraordinariamente frecuente de ejercicio de nuestras habilidades formales. Pero me cuesta reconocer que las agote.

Es cierto que el reemplazo de un término léxico por otro, conservando la estructura en una oración previamente dada, resulta ser una operación de muy escaso coste cognitivo y con un gran rendimiento. No obstante, me resulta difícil aceptar, al igual que le sucede a la escuela americana, que los reemplazos léxicos deban enfocarse con tanta intensidad en el par *monotoníaa/antitoníaa*. Y esto incumbe no solo a las relaciones de ariedad 1, sino a las de cualquier ariedad. Es cierto que la monotonía se puede proyectar sin esfuerzo a relaciones de cualquier ariedad, pero creo que al obrar así se pierden o se oscurecen muchos otros tipos de relaciones entre términos léxicos que sí se tratan, por ejemplo, en FOL. Y, por ir aún un paso más allá, tampoco creo que las inferencias naturales se limiten solo a transformaciones léxicas que preserven la estructura. Creo que también es posible actuar sobre la estructura, aunque ello suponga un coste computacional mucho más alto. Lo primero que voy a hacer es proponer un esquema de argumento que generalice apropiadamente el modelo holandés tal y como lo hemos representado en [17]. Pero antes introduciremos algunas definiciones útiles:

18.

Por $\varphi(A_1, \ldots A_n)$ se entenderá una fórmula cuyo léxico viene dado a lo sumo por los términos en $\{A_1, \ldots A_n\}$

Recuerdo que el *léxico* en este caso no hace referencia tan solo a las relaciones 1-arias típicas de FOL –*First Order Logic*–. Un elemento léxico es, en este contexto, una variable de cualquiera de los tipos[13] que el lenguaje en que se formula $\varphi(A_1, \ldots A_n)$ pueda admitir. Por ejemplo, la fórmula $p \rightarrow q$ sería una instancia, según [18], de $\varphi(p, q)$ mientras que $\forall x(Px \rightarrow \exists yRxy)$ lo sería de $\psi(Px, Rxy)$.

La expresión más general de un argumento correcto bajo la óptica de una inferencia natural extendida podría ser la siguiente:

13 Empleo la expresión tipo en un sentido técnico.

19.

$$\frac{\varphi(A_1, A_2 ... A_i ... A_n)}{\psi(B_1, B_2 ... B_i ... B_n)} \quad \theta(A_1, A_2 ... A_i ... A_n, B_1, B_2 ... B_i ... B_n)$$

donde,

20.

i. $\varphi(A_1, A_2 ... A_i ... A_n)$ es la premisa,

ii. $\psi(B_1, B_2 ... B_i ... B_n)$ es la conclusión,

iii. $\theta(A_1, A_2 ... A_i ... A_n, B_1, B_2 ... B_i ... B_n)$ es una fórmula con el léxico a lo sumo de la premisa y la conclusión y tal que:

iv. $$\frac{\begin{array}{l}\varphi(A_1, A_2 ... A_i ... A_n) \\ \theta(A_1, A_2 ... A_i ... A_n, B_1, B_2 ... B_i ... B_n)\end{array}}{\psi(B_1, B_2 ... B_i ... B_n)}$$

21.

La fórmula θ (A_1, A_2...A_i...A_n, B_1, B_2...B_i...B_n) de [19] recibe el nombre de *postulado de significado* y puede ser de dos tipos:

a) Postulado de significado léxico,
b) Postulado de significado lógico.

Intentaré dar forma a todo este material. Para empezar, he de confesar que he retomado la expresión *Postulado de significado,* recuperando la tradición carpiana (Carnap, 1952). Si vamos a admitir la expresión de relaciones de cualquier complejidad entre el léxico disponible en la premisa y la conclusión, me parece que lo más conveniente es huir de la combinatoria, en ocasiones algo limitante, que muestran tanto el modelo holandés como el americano. Estamos interesados en algo más que la monotonía o la relación obtenida al comparar las extensiones de dos clases dadas. De hecho, un postulado de significado podría ser cualquier expresión en un lenguaje formal de complejidad igual o mayor a aquella en que se expresan premisas y conclusión. Y no solo eso, sino que se contempla, igualmente, la existencia de postulados que afectan a la forma lógica y, en consecuencia, no guardan una relación directa con el léxico involucrado en la inferencia.

Antes de entrar en detalles, me gustaría decir algo más acerca de la forma en que se concibe aquí un *postulado de significado*. La cláusula [20, iv] deja bastante claro que su función es garantizar que la conclusión es consecuencia de la premisa aunque, en este momento, no indiquemos qué mecanismo inferencial es el que opera para establecer este hecho. En efecto, ni siquiera es necesario que exista un procedimiento formal, de los considerados estándar, para probar la inferencia indicada en [20, iv].

Por otra parte, pedirle a un hablante que dé cuenta en tiempo real de las razones por las que la premisa, junto con el conocimiento aportado por el postulado de significado, arrojan como resultado la conclusión, no es, ni puede ser, un requisito exigible en Lógica Natural. Las inferencias naturales son actos situados en el tiempo que vienen limitados por las capacidades de procesamiento de los sujetos que las practican. La persona que lleva a cabo una de estas inferencias solo tiene que ser capaz de elegir adecuadamente el postulado de significado pertinente al caso y efectuar las adaptaciones que la premisa requiera. Cuáles sean las razones por las que el sujeto en cuestión confía en dicho postulado es algo que puede ser capaz de aclarar eventualmente, y en caso de hacerlo, el nivel de detalle puede ser asimismo variable. Puede haberlo aprendido, sin necesidad de reflexión, a través de la experiencia propia y, con más frecuencia, a partir de la ajena. Pueden ser contenidos memorizados con el tiempo, de los cuales puede dar razón en mayor o menor medida, pero sobre cuya justificación quizá nunca se haya parado a pensar.

Podría decirse, entonces, que esta forma de entender la inferencia natural la acerca peligrosamente al terreno de los razonamientos irreflexivos e incluso falaces. Y podría ser el caso si no existiera un respaldo objetivo y comunitario a la pertinencia de los postulados de significado que los hablantes emplean en cada caso. Lo que pretendo dejar claro es que, en una inferencia natural, no se le puede exigir al hablante que aporte de manera inmediata evidencia de la forma en la que el postulado de significado actúa como sostén de la inferencia, pero esto no significa que dicho sostén no exista. Las inferencias naturales son buenas inferencias porque existe una justificación que la comunidad de los hablantes asume a partir de una evidencia científica bien asentada y transmitida a través de los usos y costumbres de una comunidad de hablantes. No pertenecen, por lo tanto, al dominio de las falacias, ni tampoco al de la Teoría de la Argumentación, aunque se puedan apreciar similitudes.

Las inferencias naturales son buenas inferencias desde el punto de vista lógico, no meras inferencias plausibles, como sucede en Teoría de la Argumentación. Son inferencias que preservan la verdad en un sentido clásico, es decir, en las que la verdad de la premisa y el postulado de significado garantizan *necesariamente*[14] la verdad de la conclusión. Es precisamente esta facultad de nuestra cognición la que este enfoque de la Lógica Natural pretende estudiar, y no más bien el vasto dominio de los argumentos aceptables en el discurso ordinario. Aprovecharé este momento para ofrecer la definición de Lógica Natural

14 Cuando hablo aquí de necesidad no pretendo conectar con ninguna versión modal de la consecuencia, sino reflejar la forma en que normalmente calificamos el hecho de que no exista una interpretación que haga verdaderos la premisa y el postulado y falsa la conclusión.

a la que me llevan todas estas consideraciones y que, como se podrá ver de inmediato, difiere sustancialmente de la ofrecida en el punto [2].

22.
> La *Lógica Natural* es la disciplina que se ocupa del estudio de los postulados de significado, sus formas, tipologías y modos de aplicación dentro de las inferencias situadas que tienen lugar en el discurso ordinario.

Como se puede ver, no es el apego a las herramientas de análisis del lenguaje ordinario, Gramática Categorial, etc., lo que marca la diferencia, sino el papel de los postulados de significado a partir de un modelo de argumento como el que se presenta en [19] y [20]. Lo que queda de esta sección lo dedicaré, precisamente, a presentar algunos de los objetivos de este tipo de estudio.

En [21] he hablado de *postulados de significado léxico* y de *postulado de significado lógico* como una forma de ampliar el alcance que los primeros desempeñan en las inferencias naturales. Empezaré por lo más conocido tratando los postulados léxicos en primer lugar.

Postulados de significado léxico

Los postulados de significado léxico, o simplemente, postulados léxicos expresan relaciones entre los términos del léxico en el que se exprese la premisa. Son, básicamente, todos los que se tratan en profundidad tanto en el modelo holand´es como en el americano, con la salvedad de que tenemos a nuestra disposición muchos más recursos para describir relaciones entre términos léxicos, al menos todas las de FOL.

El modelo de argumento se puede restringir de la siguiente manera:

23.

$$\frac{\varphi(A_1, A_2 ... A_i ... A_n)}{\varphi(B_1, B_2 ... B_i ... B_n)} \quad \theta(A_1, A_2 ... A_i ... A_n, B_1, B_2 ... B_i ... B_n$$

Obsérvese que en este caso la premisa y conclusión son ambas expresiones del tipo φ (Θ_1, Θ_2...Θ_i...Θ_n), dándose a entender de este modo que tienen la misma estructura y que lo único que varía es el léxico.

¿Cómo se trataría la monotonía/antitonía bajo esta concepción? Es decir, ¿cuáles son los postulados léxicos que expresan monotonía y antitonía? La respuesta es bastante evidente como se aprecia a continuación:

24.

Monotonía:

$$\frac{\forall x(Ax \to Bx)}{\forall x(Ax \to Cx)} \quad \forall x(Bx \to Cx)$$

El término léxico *Bx* en la premisa tiene polaridad positiva –(↑)–, el postulado de significado garantiza que ∀x(Bx → Cx), lo que en t´erminos de la extensio´n de *Bx* y *Cx* se puede expresar como [B] ⊆ [C], y por tanto, debemos concluir que ∀x(Ax → Cx).

Observarse que lo único que hemos hecho aquí es dar cuenta de que la inferencia siguiente es lógicamente válida

25.

$$\forall x(Ax \to Bx), \forall x(Bx \to Cx) \vdash \forall x(Ax \to Cx)$$

tal y como se exige en [20, iv].

El caso de la antitonía es igualmente trivial y no lo discutiré en detalle.:

26.

Antitonía:

$$\frac{\forall x(Ax \to Bx)}{\forall x(Cx \to Bx)} \quad \forall x(Cx \to Ax)$$

Nótese, en este caso, que el criterio [20, iv], junto con la identificación de las expresiones que, en FOL, se asocian con relaciones de inclusión entre términos léxicos, permiten emplear este procedimiento para identificar la polaridad de los términos léxicos en una expresión de cualquier complejidad. Supóngase dada una fórmula $\varphi(A_1, A_2 \ldots A_i \ldots A_n)$ en FOL. Deseamos ahora saber cuál es la polaridad del término A_i supuesto que tiene una única ocurrencia en $\varphi(A_1, A_2 \ldots A_i \ldots A_n)$. Tenemos entonces la siguiente solución:

26.

$$\frac{\varphi(A_1, A_2 \ldots A_i \ldots A_n)}{\varphi(A_1, A_2 \ldots A'_i \ldots A_n)} \qquad \forall x(A_i(x_0 \ldots x_n) \rightarrow A'_i(x_0 \ldots x_n))$$

Es decir, si reemplazamos en $\varphi(A_1, A_2 \ldots A_i \ldots A_n)$, A_i por A'_i y el argumento al que se añade como premisa el postulado léxico $\forall x(A_i(x_0 \ldots x_n)$ $A'(x_0 \ldots x_n))$ es clásicamente correcto, entonces A_i tiene polaridad positiva (↑) en $\varphi(A_1, A_2 \ldots A_i \ldots A_n)$. Para analizar la antitonía se hace lo propio, solo que cambiando el postulado por $\forall x(A'(x_0 \ldots x_n) \rightarrow A_i(x_0 \ldots xn))$. Evidentemente, este procedimiento tiene sus límites, ya que si aceptamos lenguajes de potencia expresiva mayor que FOL no podemos garantizar la existencia de mecanismos de prueba aceptables que actúen como los algoritmos que aquí se precisan. Algo que tampoco se puede afirmar si abandonamos el fragmento monádico de FOL. Pero, tal vez, no deberíamos detenernos en exceso en este punto.

¿Son todos los postulados de significado expresiones cuantificacionales del tipo $\forall x(A_i(x_0 \ldots x_n) \rightarrow A'(x_0 \ldots x_n))$? Es evidente que no. Si así fuera, no tendríamos manera de capturar las intuiciones, a mi juicio válidas, sobre la forma en que dos términos léxicos, en este caso relaciones 1-arias, pueden relacionarse. Veámoslo con un ejemplo muy básico que tiene que ver con la relación de *alternancia* entre términos léxicos introducida en la Tabla 1.

27.

I. Si lleva chip, entonces tiene que ser una mascota.

II. Si es una mascota no es un animal salvaje,

Por tanto,

III. Si lleva chip, entonces no es un animal salvaje.

Que se formalizaría en FOL como:

28.

$$\frac{\forall x(Chx \to Mx)}{\forall x(Chx \to \neg Sx)} \quad \forall x(Mx \to \neg Sx)$$

Siempre cabe aducir que estamos ante una instancia camuflada de monotonía en el que el postulado $\forall x(Mx \to \neg Sx)$ solo expresa que $[A] \subseteq [S^{-}]$. No creo, sin embargo, que esta crítica resulte efectiva. Pero, para ello, es cierto que tendríamos que encontrar postulados léxicos que no se expresaran en términos de una expresión cuantificada universalmente. Por fortuna existen, y podría ser que tuvieran que ver con la formulación de excepciones. Veámoslo con un ejemplo:

29.

Todos los comunistas son ateos.
Sin embargo, es bien sabido que hay Padres jesuitas que son comunistas,
Por tanto,
Hay Padres jesuitas que son ateos.

Su formalización sería entonces la siguiente:

30.

$$\frac{\forall x(Cx \to Ax)}{\exists x(Jx \to Ax)} \quad \exists x(Jx \& Cx)$$

Resulta curioso que si obviamos la triple distinción *premisa-postulado léxico-conclusión* y la reemplazamos por la clásica de *premisa mayor-premisa menor-conclusión*, lo que obtenemos es un silogismo del modo *DARII* perteneciente a la primera figura (Ferrater Mora, 1964, p. 688 y ss.). Como veremos en breve, esto no significa que el desarrollo que aquí estamos exponiendo se reduzca en definitiva a la silogística aristotélica. No lo hace ni en sus fines, ni tampoco en sus herramientas. Pero antes de ir a este punto, sí me gustaría reparar en un detalle interesante.

Según la cláusula [20, iv], lo que tiene que ocurrir para aceptar que, dado un determinado postulado de significado, una premisa arroje una conclusión, es lo siguiente:

31. *Premisa*, *Postulado de significado* ⊢ *Conclusión*,

admitiendo que existe un mecanismo capaz de evaluar la relación de consecuencia en dicho contexto que sea al menos consistente.[15]

Si la relación de consecuencia ⊢ es suficientemente clásica, las premisas, dadas en esta ocasión por la *premisa* y el *postulado de significado*, pueden alterar su posición sin que el resultado se vea afectado: la conclusión se seguirá igualmente. Sin embargo, desde el punto de vista de una inferencia natural, esto da lugar a un argumento completamente distinto en el que la premisa y el postulado de significado han alternado sus posiciones y en consecuencia su función. Si realizamos dicha operación sobre el argumento descrito en [30] se obtiene lo siguiente:

32.

$$\frac{\exists x(Jx \& Cx)}{\exists x(Jx \rightarrow Ax)} \qquad \forall x(Cx \rightarrow Ax)$$

que ahora representaría un silogismo de la 3 Figura en el modo DISAMIS.

¿Qué ha cambiado? Desde luego no la validez. Lo que ha cambiado es qué pieza de información actúa como *información fáctica* –la premisa– y cuál como *garantía*, o tal vez mejor, como *disparador* de la inferencia. Y no creo, desde luego, que pese a preservarse la validez, [30] y [32] representen la misma inferencia natural.[16]

Esta diferencia remite, más bien, al estatus epistemológico que tienen para el hablante ciertos enunciados. La información fáctica, típicamente expresada en la premisa, parece representar información acerca del mundo y, por lo tanto, contingente y potencialmente cambiante. Sin embargo, los postulados de significado, del tipo que sean, representan información distribuida e interiorizada por la comunidad de hablantes en mayor o menor medida, algo que podría caer dentro del *conocimiento común* tanto como de la *experiencia personal*. Es prematuro pronunciarse sobre la forma lógica privilegiada por los postulados de significado, y en particular, por los postulados léxicos. No obstante, también es justo reconocer que los enunciados de tipo universal pueden ser en gran medida más frecuentes que los

15 Apelo aquí a la consistencia y no a la corrección porque la primera puede expresarse al menos como una propiedad puramente sintáctica, mientras que la segunda depende de una interpretación semántica de los elementos de un lenguaje dado. Dudo que en estos contextos sea posible definir una clase de interpretaciones como las que son típicas en los desarrollos estándar de la Lógica Formal.

16 Del mismo modo que no representan la misma forma de Silogismo válido.

existenciales. Es decir, desde un punto de vista cognitivo, es más importante afianzar la norma en lugar de la excepción, si es que los postulados de tipo existencial representan realmente las excepciones.

Un tipo de postulados de significado que no parecen haber recibido demasiada atención en la literatura son aquellos que afectan a términos léxicos de ariedad mayor que uno, es decir, lo que típicamente venimos denominado como *relaciones*. En particular aquellos que no tienen que ver directamente con el par *monotonía/antitonía*. Veamos un ejemplo muy elemental:

33.

I. Ana está sentada a la derecha de Juan.
II. Si alguien está a la derecha de otra persona, esta se encontrar a su izquierda,
III. Por tanto,
IV. Juan se encuentra sentado a la izquierda de Ana.

Su forma en FOL podría expresarse del siguiente modo:

$$\frac{Daj}{Ija} \quad \forall x \forall y (Dxy \rightarrow Iyx)$$

En este caso da la impresión de que ni siquiera habría necesidad alguna de hacer explícito el postulado de significado, ya que forma parte del acervo más elemental de cualquier hablante. No parece posible ser un hablante competente del castellano sin conocer algo tan elemental como la relación existente entre este tipo de términos *posicionales*. ¿Quiere esto decir que estamos ante postulados de significado de una tipología distinta? Admito que la cuestión no parece tan fácil de dirimir como pudiera parecer. Es cierto que en el resto de los casos que hemos empleado como ejemplos, cabe ampliamente la posibilidad de que otros agentes no compartan la verdad del postulado de significado que actúa como disparador de la inferencia. Podemos ignorar si hay, por ejemplo, mascotas que sí se puedan considerar como animales salvajes en algún sentido del término, o podemos dudar de si realmente hay jesuitas que sean comunistas, pero no parece posible poner en cuestión las relaciones mutuas de términos como *izquierda* o *derecha* sin manifestar graves carencias de algún tipo. Creo que estas consideraciones solo apuntan, al menos de momento, a diferencias epistemológicas que afectan a la adquisición individual y colectiva del significado de ciertos términos. Conocido el significado de relaciones como *estar a la derecha –Dxy–* y *estar a la izquierda –Ixy–* parece necesario aceptar la validez de expresiones como $\forall x \forall y(Dxy \rightarrow Iyx)$ lo que quizá nos podría llevar a considerar que las relaciones Dxy e Ixy actúan como operadores lógicos de un tipo especial, como sucede, de hecho, con la *igualdad*. Sin descartar esta opción, creo que, por el momento, es preferible seguir considerando al postulado anterior como un postulado semántico y ello pese a que podamos ver en estos operadores de posición algo próximo a una constante lógica.

Otro terreno cuajado de complicaciones es el que afecta a las inferencias naturales procedentes del nivel sentencial. Es normal que tanto el modelo holandés como el americano presten nula atención a este ámbito de nuestra intuición formal. Si partimos de que el léxico de las expresiones propias de la Lógica sentencial son, precisamente, las variables de enunciado, una inferencia basada en un postulado léxico debería ser aquella en la que la conclusión preserva la estructura de la premisa procediendo tan solo al reemplazo de alguna variable sentencial. Lo analizaremos con un ejemplo relativamente trivial.

34.

$$\frac{p \to q}{p \to r} \quad q \to r$$

Cuesta trabajo no ver en esta inferencia un caso de monotonía aplicado a la segunda variable sentencial de la expresión condicional p → q, haciendo entonces de q → r un caso típico de postulado de significado léxico. Esta intuición se haría extensiva a la antitonía dando lugar a las siguientes inferencias naturales:

35.

$$\frac{p \to q}{p \to \vartheta} \quad q \to \vartheta$$

$$\frac{p \to q}{\vartheta \to q} \quad \vartheta \to p$$

Lo que se está retratando en [35] no es sino una versión extendida de la regla de sustitución de equivalentes. Se podría admitir que este tipo de transformaciones tienen que ver, ciertamente, con los postulados de significado léxico que hemos venido estudiando aquí. No obstante, en este momento no me atrevo a zanjar la cuestión de manera taxativa. Las inferencias naturales que afectan al nivel sentencial no parecen presentar el mismo manejo del léxico que aquel que se presenta en primer orden, pero no entraré en ello ahora.

Postulados de significado lógico

Como ya lo he mencionado, la idea de actuar sobre el léxico de una expresión conservando su estructura original, representa una forma extraordinariamente económica de obtener consecuencias necesarias a partir de una expresión inicial dada. Pero no es la única forma que tenemos de emplear nuestras intuiciones formales en el contexto de la inferencia natural. Disponemos igualmente de habilidades cognitivas que nos permiten actuar sobre la estructura de un enunciado para obtener las consecuencias necesarias del mismo. Consecuencias que transforman la estructura original conservando el léxico total o parcialmente.

Entre este tipo de operaciones destacan las que, a menudo, se han denominado como *reglas de interdefinición*. Veamos algunos ejemplos:

36.

$$\frac{\forall x Ax}{\neg\exists x \neg Ax}$$

$$\frac{p\&q}{\neg(\neg p \vee \neg q)}$$

$$\frac{\Box p}{\neg\Diamond\neg p}$$

Se trata solo de tres casos bien conocidos de inferencias cuya función es la de explicar el significado de una constante lógica apelando a otras. Para ello, resulta oportuno no actuar sobre el léxico al mismo tiempo.

Así visto, resulta difícil identificar cuál es el paralelismo con los esquemas inferenciales promovidos por postulados de significado léxico. En un esquema como el representado, por ejemplo en [31], el postulado de significado $\forall x(Cx \rightarrow Ax)$ que actúa como garante, o disparador, de la inferencia $\exists x(Jx\&Cx) \vdash \exists x(Jx \rightarrow Ax)$ parece recoger cierto conocimiento acerca de la extensión de los términos léxicos *Cx* y *Ax*. En particular: $[C] \subseteq [A]$. ¿Qué conocimiento actúa como garante o disparador, en el caso de las inferencias presentadas en [36]? Aparentemente solo puede ser el conocimiento de las reglas e interpretaciones admisibles que dan cuenta del significado de las constantes lógicas afectadas. Es decir, no parece preciso apelar a conocimiento factual de ningún tipo.

Este análisis es correcto, pero solo en cierta medida, ya que la mayoría de los hablantes del español son perfectamente capaces de aplicar tales reglas sin tener un conocimiento, siquiera aproximado, de los sistemas formales en que se tratan las constantes lógicas afectadas y que según lo anterior, son las que podrían dar cuenta de la corrección de estas inferencias.

La forma en la que los hablantes de una lengua aprenden este tipo de transacciones no es sustancialmente distinta del modo en que aprenden que la extensión de un término léxico está incluida en la de otro dado. Los postulados léxicos aportan conocimiento fáctico sobre el significado de ciertas expresiones, mientras que los postulados lógicos harían lo propio con las constantes lógicas. Y es en este punto donde las diferencias entre léxico y constantes se difuminan. Es cierto que es por completo contingente que los atenienses sean griegos, $\forall x(Ax \rightarrow Gx)$, mientras que, cualquiera de las inferencias listadas en [35] dependen de un conocimiento que afecta directamente a nuestra cognición. Pero no dependen en menor medida del conocimiento que los hablantes tienen del significado de ciertas partículas de su lengua y de sus relaciones mutuas.

¿Cómo expresaríamos las inferencias enumeradas en [36] en términos del formato básico de una inferencia natural? Una opción bastante razonable sería la siguiente:

37.

$$\frac{\forall x Ax}{\neg\exists x\neg Ax} \quad \forall\Theta\forall x(\forall x\Theta x \rightarrow \neg\exists x\neg\Theta x)$$

Al emplear una notación puramente formal, es completamente comprensible que, a continuación, esperemos una explicación en términos de los formalismos comprometidos, en este caso FOL y SOL (*Second Order Logic*). Sin embargo, aquí solo pretendemos resumir una explicación que sin ningún problema podríamos ofrecer en el lenguaje natural. La fórmula de 2° Orden $\forall x(\forall x\Theta x \rightarrow \neg\exists\neg x\Theta x)$ solo resume apropiadamente un principio como el siguiente: *Si una afirmación se predica de todos, entonces no puede haber algún individuo del que no se predique esa misma afirmación*. No hace falta un gran conocimiento de la Lógica de Primer Orden ni de la de Segundo Orden para encontrar un consenso suficiente acerca de la validez de este principio.

Y lo mismo sucede en el resto de los casos listados en [36]. Veámoslo:

38.

$$\frac{p\&q}{\neg(\neg p \lor \neg q)} \quad \forall\alpha\forall\beta(\alpha\&\beta \to \neg(\neg\alpha \lor \neg\beta))$$

$$\frac{\Box p}{\neg\Diamond\neg p} \quad \forall\alpha(\Box\alpha \to \neg\Diamond\neg\alpha)$$

El postulado de significado que apoya la primera de estas inferencias responde a un principio informal que se podría expresar del siguiente modo: *si es cierto que dos enunciados cualesquiera son verdaderos a la vez, entonces no puede ocurrir que alguno de ellos sea falso*. Y sucede, de igual modo, para el caso modal analizado, cuya traducción informal sería: *Si algo se da necesariamente, entonces no es posible que sea falso*.

Pero los postulados de significado lógico no se agotan en las interdefiniciones. Hay muchos más frentes en los que resultan reconocibles. Iré por partes:

39.

$$\frac{\forall x Ax}{\exists x Ax} \quad \forall\Theta\forall x(\forall x\Theta x \to \exists x\Theta x)$$

$$\frac{\Box p}{\Diamond p} \quad \forall\alpha(\Box\alpha \to \Diamond\alpha)$$

Los postulados de significado empleados en [39] expresan formalmente los siguientes principios intuitivos: *lo que se dice de todos, se dice de alguno* y *lo que se da necesariamente es igualmente posible*. Curiosamente ninguno de estos dos postulados son universalmente aceptados en el ámbito de la Lógica Formal. El primero de ellos porque está asociado a la condición de que no existan modelos cuyo dominio de cuantificación sea el conjunto vacío, y, el segundo, por razones semejantes: no puede haber mundos posibles incapaces de acceder a otro composible con ellos.[17]

¿Cuántos más de estos postulados de significado se pueden considerar? Esta pregunta esconde lo que, posiblemente, sea una crítica bastante consistente a los postulados de significado lógico. Porque en realidad, cualquier inferencia formal que pueda ser demostrada en algún sistema formal[18] aceptado por la comunidad científica, puede ser reinterpretada como un

17 De hecho, la f´ormula □α → ?α se tiene que añadir como axioma, el axioma D, en la interpretación deóntica de los operadores modales.

18 Y deberíamos añadir, que satisfaga el Teorema de Deducción.

postulado de significado formal. En ese sentido, se podría decir que los postulados lógicos son una forma de reintroducir la Lógica Formal, con todo su peso, en el territorio de la Lógica Natural. Y ello solo con una especie de ejercicio malabar consistente en expresar un inferencia válida como un principio en el que se cuantifica universalmente sobre las variables léxicas existentes.

Aquí hay varias consideraciones que hacer. La primera tiene que ver con la complejidad de los postulados a incorporar. Por definición, no parece aceptable que podamos manejar postulados de una complejidad excesiva, simplemente porque no los podríamos memorizar ni aplicar de manera efectiva. Por otra parte, debe tenerse en cuenta que una inferencia natural no incorpora en su aplicación las razones por las que un determinado postulado es o no aceptable. Estas habrán sido aportadas y discutidas en su momento e incorporadas en el acervo de una cierta comunidad de hablantes. No se precisa de una especie de Lógica universal que sea capaz de dar cuenta, a la vez, de todas los postulados que podemos tratar de este modo. De hecho sabemos que tal cosa no existe. Los postulados lógicos pueden encontrar respaldo en sistemas formales, cierto, pero también pueden basarse en razonamientos informales, a falta de algo mejor. De hecho, algunos de estos postulados pueden resultar contradictorios entre sí o simplemente inaceptables, tanto para algunos hablantes como para ciertos contextos. Por ejemplo, si empleo como postulado de significado $\forall\alpha\forall\beta(\alpha\&\neg\alpha \rightarrow \beta)$ para concluir *q* a partir de una premisa $p\&\neg p$, estaré corriendo el riesgo de que algún partidario de las Lógicas paraconsistentes o las relevantistas ponga en cuestión mi razonamiento. De hecho, solo en contextos muy específicos se acepta un postulado tan radical como el anterior, por mucho que represente el principio sacrosanto de *ex contradictione quodlibet*.

En definitiva, no hace falta una Lógica universal,[19] ni tan siquiera un único sistema formal. Podemos recurrir a una diversidad de ellos, incluso contradictorios entre sí, o en ausencia de alguno apropiado, a razonamientos informales que aporten razones a favor del postulado empleado en cada caso.

De hecho, el único rasgo de logicidad compartido por cualquier inferencia natural basada en un postulado lógico es la aplicación del *modus ponens* –MP– en los siguientes términos:

40.

Premisa, *Premisa*   *Conclusión* ⊢ *Conclusión*

¿Ofrece alguna ventaja que nuestros postulados de significado se justifiquen por medio de una sistema formal al uso? Sin duda. Es algo que puede ayudar, y mucho, a analizar inferencias naturales que no resulten tan cla-

19 No hago referencia aquí a los contenidos propios de los Congresos *Universal Logic*, sino a un sistema capaz de dar cuenta de todas las inferencias en las que intervienen partículas portadoras de logicidad.

ras a primera vista. Todo dependerá entonces de la potencia expresiva del lenguaje en el que surja la disputa. Si está limitada a un rango de recursos para los que existe algún cálculo, o aún mejor, algún algoritmo decidible de prueba, la disputa siempre será soluble mediante el cálculo. En caso contrario, deberemos conformarnos con lo que podamos conseguir en cada caso.

Creo, y con esto concluyo, que la identificación de los postulados de significado lógico y su función en la inferencia natural tiene un componente empírico (Geurts, 2005, secc. 6.2) que no puede ser desestimado. Me gustaría pensar que existe un hilo conductor capaz de jerarquizar su uso en función de variables como el tiempo de ejecución, universalidad, tipología de los operadores lógicos involucrados, etc. Pero dudo de que estemos siquiera cerca de ese logro.

La Lógica Natural en las Humanidades

La cuestión que sobrevuela en este trabajo no es otra que esa. ¿Debemos considerar a la Lógica Natural como un saludable reemplazo de la Lógica Formal en los estudios de Filosofía? La cuestión del lugar de la Lógica en las Humanidades, no es nueva, ni tiene una fácil respuesta –véase, por ejemplo (Alonso, 2019b)– pero es, en cierta medida, la auténtica cuestión de la Lógica.

Años atrás, la Lógica Formal inició un paulatino retroceso frente a la Teoría de la Argumentación en los planes de estudio de Humanidades (Alonso, 2019a). Unas veces tuvo lugar como un genuino reemplazo de una asignatura por otra en los planes de estudio y otras, más frecuentemente, como un cambio subrepticio de los contenidos propios de la Lógica Formal ¿Se va a repetir este proceso con la Lógica Natural? Creo que estamos muy lejos de ello.

Es cierto que la Lógica Natural presenta contenidos que, en buena ley, resultan más próximos a las preocupaciones e intereses del humanista que aquellos que en la actualidad se estudian en algunos cursos de Lógica Formal. Esto, que parece una virtud, podría ser también un riesgo si se aprovecha para alimentar una reintroducción de la vieja Silogística escolástica en los planes de estudio. Muchos estarían conformes con este anacronismo, pero este no es mi caso.

La Lógica Natural requiere, para ser justamente comprendida, conocimientos no elementales de Gramática Formal. A eso se debe sumar cierto dominio de ramas muy técnicas de la Lógica Formal: Teoría de Tipos, Cálculo de Tablas Analíticas y Cálculo lambda, entre otros. Es posible, no lo descarto en absoluto, que en un tiempo razonable seamos capaces de reducir la carga técnica al mínimo, al tiempo que vayan componiéndose manuales útiles para exponer sus contenidos de forma sistemática. Al fin y al cabo, esa es la intrahistoria de la Lógica Formal en los estudios de Humanidades. Por el momento, habrá que esperar a que dichos manuales estén disponibles y a que la Academia asimile la necesidad o la conveniencia de revisar los contenidos de sus objetivos formativos.

Con esto no quiero, ni por asomo, promover la desaparición de la Lógica Formal de los planes de estudio de las Humanidades. Sería fácil sucumbir a la retahíla de discursos que ven en esta disciplina una suerte de Matemáticas encubiertas bien alejadas del genuino espíritu de la Filosofía. Pero es inmediato observar que esos mismos argumentos servirían para cerrar el paso a la Lógica Natural y, en general, a cualesquiera de los estudios que traigan al presente los métodos y enseñanzas de la Filosofía. Solo una visión muy específica y pobre de esta disciplina podría reivindicar tal cosa. Yo solo pido una reflexión para analizar cuánta Lógica y de qué tipo es aquella que debe saber un filósofo, no si debe conocer alguna.

Por otra parte, aún está pendiente de aclarar qué derechos arropan la petición de una presencia para la Lógica Natural en la Filosofía. Un argumento, y no menor, podría ser el de su peso en las infraestructuras básicas de la IA-gen. Ya hemos hablado de ello al comentar algunos de los enfoques de la escuela americana y su orientación eminentemente computacional. Que la Filosofía ignore los principios básicos de una disciplina con un potencial transformador como lo es la IA-gen, me parece de una irresponsabilidad que muy bien haríamos en evitar, aunque también pueda reconocer mi escepticismo al respecto. Pero no es esa la razón por la que defiendo que se estudie el modo en que la Lógica Natural podría formar parte de la formación del filósofo. La pregunta que no puedo, no podemos, dejar de hacernos es ¿por qué somos capaces de razonar lógicamente? Para entender la dimensión de esta pregunta fundamental tengo que pedirle al lector que imagine la posibilidad de que no lo hiciéramos, o que lo hiciéramos de forma sustancialmente distinta a como, de hecho, ocurre. No es fácil, pero lo poco que la paleoantroplogía nos dice de la condición humana es que ésta se ha expresado en un número, además creciente, de especies de las cuales solo la nuestra ha sobrevivido en el tiempo ¿Tendrían una Lógica, siquiera rudimentaria? ¿sería la nuestra? La etología también nos regala, con el tiempo, estudios en los que distintas especies, a veces muy distintas en realidad, presentan formas de logicidad nada desdeñables ¿Cuál es, comparativamente, nuestra *firma* característica? ¿Cuál es la estructura de esta habilidad cognitiva que nos permite inferir nuevas[20] piezas de información a partir de otras dadas?

Para un materialista militante, como es mi caso, la existencia de la habilidad formal en nuestra cognición solo puede explicarse como parte del éxito evolutivo de nuestra especie. En consecuencia, y en la medida en que la presión selectiva siga presente entre nosotros, estaría inmersa en un constante proceso de cambio. Entender la estructura de esa parte de nuestra cognición, y hacerlo tanto desde el plano conceptual como desde el empírico, es una aventura a la que no me gustaría renunciar como humanista. Es cierto que estaríamos abandonando una multitud de supuestos que afectan a los

20 No quiero decir nuevas en el sentido de que amplíen la información de la premisa: eso nos llevaría fuera de la Lógica.

estudios de Lógica situándola en un territorio intermedio entre las matemáticas, la antropología, la psicología y la teoría de la mente, pero, ¿qué perdemos dando una oportunidad a esta opción? En nuestra tradición reciente la hibridación de disciplinas ha dado, a menudo, los mejores resultados, no creo que debamos negarnos a intentarlo, una vez más, en este caso.

Referencias bibliográficas

Alonso, E., (2019a). "Lógica y teoría de la argumentación. Anatomía de una reforma". *Quadripartita ratio*, (7), pp. 12-31.

------- (2019b). "¿Qué lógica necesita saber un filósofo?" *Andamios, Revista de Investigación Social*, 16(41), pp. 143-163.

Belnap, N. D. (1962). "Tonk, plonk and plink". *Analysis*, 22(6), pp. 130-134.

Carnap, R. (1952). "Meaning postulates". *Philosophical Studies*, 3(5), pp. 65-73.

Ferrater Mora, J. (1964). *Diccionario de filosofía de bolsillo*. Alianza.

Frege, G. (1972). *Conceptografía*. Instituto de Investigaciones Filosóficas. Colección: Filosofía Contemporánea.

Geurts, B. (2005). "Monotonicity and processing load". *Journal of Semantics*, 22(1), pp. 97-117.

MacCartney, B. and Manning, C. (2009). "An extended model of natural logic". *Proceedings of the 8th International Conference on Computational Semantics*, pp. 140-156.

Moss, L. S. (2010). "Natural Logic and Semantics". En Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Ter- zopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Aloni, M., Bastiaanse, H., De Jager, T., and Schulz, K., editors, *Logic, Language and Meaning*, volume 6042, pp. 84-93. Springer Berlin Heidelberg.

Sánchez-Valencia, V. (1991). *Studies on Natural Logic and Categorial Grammar*. Universiteit van Amsterdam.

Van Benthem, J. (1990). "Categorial Grammar and Type Theory". *Journal of Philosophical Logic*, 19(2), pp. 115-168. http://www.jstor.org/stable/30226424.

Van Benthem, J. (1991). "Logic in action". *Journal of Philosophical Logic*, 20(3), pp. 225-263.

Van Benthem, J. (2008). *A brief history of natural logic. Technical report, ILLC.* Amsterdam and Stanford.