

Capítulo 6

Razonamiento por defecto mediante δ -resolución

Fernando Soler Toscano¹

6.1. Razonamiento por defecto y no monotonía

Una de las características propias de los sistemas formales (deductivos) clásicos es su *monotonía*. Así, dado $\mathcal{L}$, un lenguaje de primer orden con las convenciones habituales, si $\Gamma \cup \{\alpha\} \subset \mathcal{L}$ entonces, si se verifica

$$\Gamma \models \alpha \tag{6.1}$$

es decir, que α es *consecuencia lógica* de Γ , entonces, dado cualquier conjunto de fórmulas Λ , también se verificará

$$\Gamma \cup \Lambda \models \alpha \tag{6.2}$$

Sin embargo, muchos razonamientos que realizamos en la vida ordinaria no son monótonos, pues no poseen ninguna cualidad similar a la propiedad formal de la relación de consecuencia lógica clásica que permite pasar de (6.1) a (6.2) en todos los casos, sean cuales sean Γ , α y Λ . Por ejemplo, fácilmente

¹Este trabajo se ha realizado en el marco del Proyecto de Investigación HUM2004-01255 del Ministerio de Educación y Ciencia. Asimismo, se ha contado con la ayuda de la Junta de Andalucía, a través de la financiación al Grupo HUM609 del Plan Andaluz de Investigación, y de la concesión de una beca de Formación de Doctores.

reconoceríamos como válido el razonamiento² que aparece en el cuadro 6.1 si sólo tomamos en cuenta las premisas a y b .

- | | |
|--|---------------------------|
| $a.$ Si se acciona el interruptor, la bombilla se enciende.
$b.$ Voy a accionar el interruptor.
$c.$ La bombilla está fundida. | La bombilla se encenderá. |
|--|---------------------------|

Cuadro 6.1: Ejemplo de un razonamiento no monótono

Si tomamos Γ como las dos primeras premisas y α como la conclusión, entonces podemos comparar el anterior razonamiento (sin considerar la premisa c) con la relación de consecuencia lógica expresada en (6.1). Sin embargo, si hacemos que Λ contenga la premisa c , entonces no es válido lo expresado en (6.2), puesto que desde a , b y c no se puede inferir, mediante sentido común, la conclusión “La bombilla se encenderá”. Es más, podemos inferir precisamente lo contrario, “La bombilla no se encenderá”.

El ejemplo anterior es un caso de *razonamiento por defecto*, muy habitual en las inferencias que solemos llamar *de sentido común*. Uno de los recursos propios del razonamiento humano es la generalización inductiva, basada con frecuencia en lo que en términos formales se conoce como *asunciones de normalidad*. Así, la mente comprende la premisa a del cuadro 6.1 realmente más o menos como:

Normalmente, si se acciona el interruptor, la bombilla se enciende.

o bien:

Si se acciona el interruptor *y todo va bien*, la bombilla se enciende.

con lo que en condiciones normales es verdad lo expresado en a , pero se reserva la posibilidad de que haya situaciones *anormales* o *atípicas*, como que la bombilla se funda, que impiden usar la premisa. Entonces, la regla a se empleará siempre que no se pueda inferir que la situación dada es uno de los casos de anormalidad. Por *defecto*, se asumirá que toda situación es normal (respecto de la aplicación de cada regla) mientras no se demuestre lo contrario.

¿Por qué funciona así la mente humana? Los psicólogos cognitivos no han llegado a un acuerdo sobre por qué la mente se comporta de forma distinta a la lógica formal. Lo que resulta claro es que este modo de razonar resulta mucho más eficaz que si intentáramos corregir nuestra mente para que sólo realizara inferencias lógicas. Entonces, por ejemplo, la premisa a debería considerarse

²Como es habitual, las oraciones que se encuentran sobre la línea horizontal son premisas; la que aparece debajo, es la conclusión. Intuitivamente, un razonamiento es *válido* si no es posible que las premisas sean verdaderas y simultáneamente la conclusión sea falsa.

estrictamente falsa. Tal vez sería verdadera la premisa a' , que reza: “Si se acciona el interruptor y la bombilla no está fundida, entonces la bombilla se enciende”. Pero en realidad, a' tampoco es válida. Quizás lo sea a'' : “Si se acciona el interruptor y la bombilla no está fundida y el interruptor no está roto, entonces la bombilla se enciende”. O mejor, a''' : “Si se acciona el interruptor y la bombilla no está fundida y el interruptor no está roto y el suministro eléctrico no se ha cortado, entonces la bombilla se enciende”. Pero tampoco es estrictamente verdadera a''' ; haría falta siempre un nuevo refinamiento, y al final toda regla sería infinita, y seguramente intratable para la mente humana. Si la evolución nos ha permitido sobrevivir razonando de este modo *ilógico* seguramente sea porque es mucho más económico considerar reglas del tipo de a , y tener la mente siempre abierta a que toda regla puede tener excepciones, algunas quizás no descubiertas.

A pesar de que la relación de consecuencia lógica clásica no logre capturar las inferencias que hemos llamado *de sentido común*, existen numerosos acercamientos formales a modos de razonamiento no monótonos. Concretamente, el razonamiento por defecto puede formalizarse a través de los llamados *sistemas Reiter*, que permiten añadir, a los cálculos propios de la lógica clásica, un conjunto de reglas de la forma siguiente:

$$\frac{\alpha_1, \dots, \alpha_n, \diamond \beta_1, \dots, \diamond \beta_m}{\gamma} \quad (6.3)$$

Dada una teoría Γ (un conjunto de fórmulas de $\mathcal{L}$), la regla (6.3) permite inferir γ siempre que puedan demostrarse $\alpha_1, \dots, \alpha_n$ y además $\Gamma \cup \{\beta_1, \dots, \beta_m\}$ sea un conjunto satisfactible. El sentido del operador “ $\diamond$ ” es restringir la aplicabilidad de la regla (6.3) sólo a los casos en que sea consistente con la teoría Γ asumir simultáneamente como verdaderas todas las β_i , $1 \leq i \leq m$. Entonces, cada una de las β_i formaliza una de las *asunciones de normalidad* que deben hacerse para aplicar la regla (6.3). Además, cuando la regla no puede aplicarse es que nos encontramos ante una situación *anormal*, puesto que las asunciones de normalidad que exige (6.3) no pueden ser satisfechas dentro de la teoría Γ .

Para ilustrar lo anterior, veamos un ejemplo típico de regla por defecto. Es la siguiente:

$$\frac{Ave(x), \diamond Ave_normal_vuela(x)}{Vuela(x)} \quad (6.4)$$

Tengamos en cuenta que la variable libre x que aparece en (6.4) se debe entender implícitamente como cuantificada universalmente, puesto que la regla se trata, realmente, de un esquema que aceptamos para cada instancia de la variable x . Por tanto, (6.4) es simplemente la formalización de la idea común de que *generalmente* todas las aves vuelan. El predicado $Ave_normal_vuela(x)$,

que aparece marcado con $\diamond$, es una asunción de normalidad, en este caso de la normalidad de las aves respecto de la capacidad de volar.

Las siguientes fórmulas nos ayudarán a comprender cuándo puede aplicarse la regla (6.4) y cuándo no está permitido:

$$\forall x(Pingüino(x) \rightarrow \neg Ave_normal_vuela(x)) \quad (6.5)$$

$$Ave(a) \quad (6.6)$$

$$Pingüino(a) \quad (6.7)$$

Consideremos una teoría Γ_1 que contenga exclusivamente las fórmulas (6.5) y (6.6). Además, la única regla por defecto es (6.4). Entonces, dado que por (6.6) tenemos $Ave(a)$ y que $Ave_normal_vuela(a)$ es consistente con Γ_1 (para probar esto necesitaríamos recurrir a algún procedimiento semántico de búsqueda de modelos, aunque en este caso sería bastante sencillo), podemos aplicar la siguiente instancia de (6.4),

$$\frac{Ave(a), \diamond Ave_normal_vuela(a)}{Vuela(a)} \quad (6.8)$$

y concluir, por defecto, $Vuela(a)$, es decir, que el ave a vuela.

Pero, ¿qué ocurre si ampliamos nuestro conocimiento sobre a ? Sea Γ_2 la teoría que amplía Γ_1 con (6.7). Entonces ya no podemos usar la regla (6.8), ya que la asunción de normalidad que su aplicación requiere resulta inconsistente con la teoría, puesto que desde (6.5) y (6.7) podemos deducir $\neg Ave_normal_vuela(a)$. Por tanto, al saber que a es un pingüino, concluimos que pese a ser un ave no posee la capacidad común de volar. Se ha activado, pues, una excepción a nuestra regla por defecto (6.4).

Las conclusiones que se obtienen empleando el razonamiento por defecto no sólo pueden rebatirse porque se añada conocimiento específico sobre un cierto individuo, como ha ocurrido en el ejemplo anterior al añadir a la base de conocimientos la información de que a es un pingüino. También es posible añadir reglas generales, como las siguientes:

$$\begin{aligned} \forall x(Tiene_alas_rotas(x) &\rightarrow \neg Ave_normal_vuela(x)) \\ \forall x(Gallina(x) &\rightarrow \neg Ave_normal_vuela(x)) \\ &\vdots \end{aligned}$$

Con estas reglas, seguramente se impediría aplicar la regla (6.4) para aves que, de otra forma, asumiríamos, *por defecto*, que sí vuelan. Son nuevas excepciones a la regla general de que todas las aves vuelan.

Antes de continuar, realicemos una observación. Al formular la regla (6.4), la asunción de normalidad podría haber sido el mismo predicado $Vuela(x)$, con

lo que tendríamos la siguiente regla:

$$\frac{Ave(x), \diamond Vuela(x)}{Vuela(x)} \quad (6.9)$$

Entonces, si cambiamos (6.5) por

$$\forall x (Pingüino(x) \rightarrow \neg Vuela(x))$$

en la teoría Γ_1 seguiríamos concluyendo por defecto $Vuela(a)$, pero ahora, en Γ_2 no sólo no podríamos obtener esta conclusión, sino que se demostraría $\neg Vuela(a)$. De hecho, pueden emplearse como asunciones de normalidad los mismos predicados que aparecen en las conclusiones de las reglas por defecto. Sin embargo, en este trabajo vamos a usar predicados distintos, para que siempre resulte claro cuáles son los que se reservan como asunciones de normalidad. En lo fundamental, el tratamiento que presentaremos no varía si se usan los mismos predicados, tan sólo haría falta tener en cuenta ciertos detalles.

Un problema que surge al usar sistemas Reiter es que para aplicar reglas como (6.3) debe comprobarse que todas las α_i , $1 \leq i \leq n$ son demostrables en la teoría que se está usando, pero además que todas las β_j , $1 \leq j \leq m$ son consistentes con la teoría. En primer orden, la forma de hacer lo segundo es mediante procedimientos semánticos de búsqueda de modelos. Pero esto amplía enormemente la complejidad del sistema, pues hay que hacer en paralelo dos procesos, la comprobación de que se demuestran las α_i y la de la consistencia de todas las β_j con la teoría. Por ello, resultaría útil encontrar un mismo proceso que pudiese integrar ambas comprobaciones. Es lo que vamos a hacer en este trabajo, tratar las inferencias por defecto como un modo especial de abducción, y aplicar entonces el cálculo de δ -resolución, que originalmente desarrollamos para su aplicación en la resolución de problemas abductivos, pero que demuestra ser útil para abordar la formalización de otros modos de razonamiento no monótonos.

6.2. Abducción y razonamiento por defecto

C. S. Peirce, padre del pragmatismo norteamericano, define un tipo de razonamiento, la *abducción*, que guarda muchas semejanzas con los razonamientos que hemos presentado en la sección anterior. Para Peirce, sólo en lógica y matemáticas pueden existir juicios seguros, irrefutables. Para el resto de disciplinas, y por supuesto también para los razonamientos de sentido común, no es posible esta seguridad; nos movemos siempre en el terreno de la *hipótesis*. Por ello, las conclusiones que ahora alcanzamos podrán ser mañana rebatidas si la nueva información que obtenemos así lo exige. En este tipo de razonamientos resulta

fundamental el papel de la *abducción*, usada para explicar observaciones nuevas con el conocimiento que tenemos. Peirce proporciona el siguiente esquema para ilustrar cómo procede la abducción:

El hecho sorprendente, C , es observado.

Pero si A fuera verdad, C sería aceptado como algo evidente.

Por lo tanto, hay razón para sospechar que A es verdad (*CP 5.189, 1903*).

La abducción parte de una sorpresa, la observación del hecho inesperado C , para el que desea buscarse una explicación. Analizando nuestro conocimiento concluimos que si A fuera verdad entonces C sería una consecuencia obvia. Por tanto, la conclusión abductiva nos presenta A como una explicación posible de la observación C .

Este tipo de razonamiento ha sido estudiado en numerosas disciplinas, por el protagonismo que encuentra prácticamente en todo ámbito que se caracterice por el empleo de inferencias explicativas. Por ejemplo, en la Inteligencia Artificial se emplean modelos formales de inferencia abductiva para la construcción de programas de diagnosis que pueden detectar cuáles son los fallos (la explicación, A) que hay detrás del comportamiento defectuoso (la observación, C) de cierto sistema (una red de ordenadores, un microprocesador, etc.). También en Filosofía de la Ciencia se ha discutido cuál es el papel que representa el razonamiento abductivo dentro de la *explicación científica*, es decir, dentro de las operaciones que se llevan a cabo (en la mente de cierto individuo o dentro de la comunidad científica) cuando se formula una explicación (que puede ser hasta una nueva teoría) para un fenómeno nuevo o inexplicado.

Pasando a un plano más formal, veamos cómo podemos tratar lógicamente las ideas de *problema abductivo* y *solución abductiva*. En las definiciones siguientes seguimos, en parte, a M. Cialdea y F. Pirri [CP93] y A. Aliseda [Ali97].

Definición 1 (Problema abductivo) Sean³ $\Theta \subset \mathcal{L}$ y $\phi \in \mathcal{L}$. Entonces, decimos que $\langle \Theta, \phi \rangle$ es un problema abductivo si y sólo si se verifican

$$\Theta \not\models \phi$$

y

$$\Theta \not\models \neg\phi$$

La definición 1 establece que $\langle \Theta, \phi \rangle$ es un problema abductivo cuando desde Θ no se derivan ni ϕ ni su negación. Por tanto, Θ ocupa el lugar de una teoría

³La semántica de $\mathcal{L}$ se construye como es habitual en un lenguaje de primer orden, teniendo simplemente en cuenta que sólo se evalúan las sentencias, fórmulas sin variables libres.

y ϕ el de una observación que no se deriva de tal teoría (tampoco se deriva la negación de la observación, pues en tal caso ϕ sería una refutación de Θ , y no se plantearía un problema abductivo) pero que desea explicarse dentro de Θ . Las explicaciones de ϕ en Θ serán las *soluciones abductivas* al problema abductivo $\langle \Theta, \phi \rangle$, como explicamos en la siguiente definición.

Definición 2 (Solución abductiva) Sea $\langle \Theta, \phi \rangle$ un problema abductivo y $\alpha \in \mathcal{L}$. Entonces, decimos que α es una solución abductiva al problema abductivo $\langle \Theta, \phi \rangle$ si y sólo si se cumplen:

$$\Theta, \alpha \models \phi \quad (6.10)$$

$$\Theta, \alpha \not\models \perp \quad (6.11)$$

$$\alpha \not\models \phi \quad (6.12)$$

Cada uno de los requisitos que la definición 2 impone a ϕ establece una condición que deben cumplir las soluciones abductivas. Así, (6.10) formaliza la exigencia fundamental para que α sea una explicación de ϕ en la teoría Θ . Se trata de que la observación se derive de la teoría junto con la explicación. También, (6.11) impone un requisito importante, a saber, que la explicación sea consistente con la teoría. En otro caso, α explicaría con Θ cualquier cosa, según la propiedad de la relación de consecuencia lógica clásica conocida como *explosión*, que formaliza el principio de *ex contradictione quodlibet*. El requisito expresado en (6.12) no es incluido necesariamente por todos los autores dentro de la noción de solución abductiva. Lo que exige es que la observación ϕ no se derive de la sola explicación α . Así, α será una explicación válida para la teoría Θ , pero probablemente no lo será para otra teoría Θ' .

Es común exigir requisitos adicionales a las soluciones abductivas para de entre todas las explicaciones posibles (que en lenguajes formales como el de primer orden que estamos usando pueden ser infinitas) seleccionar *la mejor explicación*. También nosotros vamos a incorporar *criterios preferenciales* para elegir las explicaciones. Son los siguientes:

Forma sintáctica de las explicaciones. Las explicaciones seleccionadas serán sentencias (fórmulas sin variables libres) de $\mathcal{L}$ con la siguiente forma

$$Q_1 x_1 \dots Q_n x_n (\gamma_1 \wedge \dots \wedge \gamma_m)$$

tal que cada Q_i , $1 \leq i \leq n$ es un cuantificador $\forall$ ó $\exists$ (también puede no haber cuantificadores) y cada γ_j , $1 \leq j \leq m$, es un literal de $\mathcal{L}$.

Minimalidad semántica. Dadas dos explicaciones α_1 y α_2 , si

$$\alpha_1 \models \alpha_2$$

entonces preferimos α_2 a α_1 , ya que es una explicación más débil, y por tanto menos restrictiva. Decimos que α_2 es *minimal respecto de* α_1 . Determinar la minimalidad en lógica de primer orden es un problema indecidible, pero siempre que pueda determinarse que una explicación es minimal respecto de otra, aplicaremos el criterio anterior.

Uso de predicados abducibles. En ocasiones, interesará que los literales que forman parte de las explicaciones usen ciertos predicados específicos. Para ello, se definirá cierto conjunto de predicados de $\mathcal{L}$ como el conjunto de *predicados abducibles*, y sólo se aceptarán explicaciones tales que todos sus predicados sean abducibles.

Un rasgo fundamental del razonamiento abductivo es su carácter no monótono. Entre otras cosas, esto nos permitirá tratar el razonamiento por defecto como una forma de abducción. Veamos un ejemplo, tomado de [KKT98], donde se manifiesta la no monotonía de la abducción. Usaremos un lenguaje proposicional con las variables l (que tomamos por “llovió anoche”), a (por “los aspersores han regado”), c (“el césped está mojado”), z (“los zapatos están mojados”). Definimos el problema abductivo siguiente:

$$\langle \{l \rightarrow c, a \rightarrow c, c \rightarrow z\}, z \rangle$$

De modo que la teoría contiene la información siguiente:

- Si llovió ayer, el césped está mojado.
- Si los aspersores han regado, el césped está mojado.
- Si el césped está mojado, los zapatos están mojados.

y la observación que desea explicarse es que los zapatos están mojados. Según lo explicado, entre las soluciones abductivas se encuentran l , a y c . Pero ahora bien, sea otro problema abductivo:

$$\langle \{l \rightarrow c, a \rightarrow c, c \rightarrow z, \neg l\}, z \rangle$$

que sólo añade al problema abductivo anterior la fórmula $\neg l$, lo cual puede suceder, por ejemplo, si alguien nos dice que no llovió anoche. Entonces, l ya no cumple los requisitos establecidos por la definición 2 para ser una solución abductiva a este segundo problema abductivo. Sí lo seguirán siendo a y c , entre otras.

Este ejemplo muestra cómo la ampliación de las teorías puede alterar el conjunto de soluciones abductivas posibles. No sólo pueden perderse soluciones; según sean las ampliaciones de las teorías también pueden añadirse nuevas explicaciones.

En la próxima sección mostraremos el cálculo de δ -resolución, un cálculo abductivo que nos permitirá sistematizar la búsqueda de soluciones abductivas. Posteriormente, en la sección 6.4 veremos cómo pueden resolverse los problemas propios del razonamiento por defecto mediante δ -resolución. Pero primero, mostremos cómo es posible transformar un problema de razonamiento por defecto en un problema abductivo. En parte, seguimos ideas de P. Flach [Fla94], aunque adaptadas a nuestro tratamiento de la abducción.

Consideremos el ejemplo que presentamos en la página 87. Se trata de saber si determinada ave vuela. Para ello contamos con la regla por defecto (6.4). Para transformar este problema en un problema abductivo, comenzamos por formular la teoría. Las fórmulas (6.5), (6.6) y (6.7) permanecen tal cual. La clave está en transformar las reglas por defecto en fórmulas de la teoría. Concretamente, podemos transformar (6.4) en:

$$\forall x(Ave(x) \wedge Ave_normal_vuela(x) \rightarrow Vuela(x)) \quad (6.13)$$

En general, una regla por defecto como (6.3) se transforma en la fórmula

$$\forall x_1, \dots, x_j(\alpha_1 \wedge \dots \wedge \alpha_n \wedge \beta_1 \wedge \dots \wedge \beta_m \rightarrow \gamma) \quad (6.14)$$

tal que $x_1, \dots, x_j$ son las variables libres que ocurren en (6.3).

Volviendo a nuestro problema, consideremos que la teoría, Θ , se compone de (6.5), (6.6) y (6.13). Entonces, $\langle \Theta, Vuela(a) \rangle$ es un problema abductivo, en el sentido de la definición 1, tal como puede comprobarse mediante cualquier sistema de decisión para lógica de primer orden. Esto significa que, en sentido clásico, no es posible determinar si a vuela o no. Igualmente, según la definición 2, se verifica que $Ave_normal_vuela(a)$ es una solución abductiva a nuestro problema abductivo, lo cual implica:

$$\Theta, Ave_normal_vuela(a) \models Vuela(a) \quad (6.15)$$

$$\Theta, Ave_normal_vuela(a) \not\models \perp \quad (6.16)$$

$$Ave_normal_vuela(a) \not\models Vuela(a) \quad (6.17)$$

Entonces, ¿cuál es la respuesta a la cuestión sobre si a vuela o no? Veamos, por (6.15) tenemos que si asumimos $Ave_normal_vuela(a)$ podemos concluir, junto a la teoría Θ , que a vuela. Pero esto es una asunción de normalidad, por lo que puede asumirse siempre que no resulte contradictoria con la teoría (tal era el sentido de $\diamond$); precisamente (6.16) asegura que no existe tal contradicción. Por último, (6.17) indica que para concluir que a vuela no sólo hace falta aceptar las asunciones de normalidad, sino que la teoría resulta imprescindible.

Ahora bien, si añadimos (6.7) a Θ , perdemos la solución abductiva $Ave_normal_vuela(a)$, ya que (6.16) deja de verificarse, porque la asunción de normalidad resulta inconsistente con la nueva teoría, que ahora nos dice que por ser a un pinguino no es posible asumir $Ave_normal_vuela(a)$.

Generalizando, para tratar un problema de razonamiento por defecto como un problema abductivo, seguimos los siguientes pasos:

1. Se transforman las reglas por defecto tipo (6.3) en fórmulas tipo (6.14). Estas fórmulas, junto con la base de conocimientos, formarán la teoría Θ . Los predicados que formaban las asunciones de normalidad se toman ahora como predicados abducibles.
2. Sea ϕ lo que deseamos derivar por defecto desde Θ (en el ejemplo anterior ϕ era $Vuela(a)$). Entonces, si $\langle \Theta, \phi \rangle$ es un problema abductivo, existe un problema de razonamiento por defecto, porque lo que deseamos saber, ϕ no se infiere en sentido clásico de nuestra base de conocimientos.
3. Si existe al menos una solución abductiva tipo $\delta_1 \wedge \dots \wedge \delta_l$ al problema abductivo $\langle \Theta, \phi \rangle$ tal que cada δ_k , $1 \leq k \leq l$, es un literal formado por un predicado abducible (para que pueda asumirse por defecto) y sin variables (para que las asunciones de normalidad sean siempre sobre individuos concretos, determinados), entonces es posible inferir por defecto ϕ desde Θ . Ello es así porque, por la definición 2:
 - Por (6.10), sólo hace falta aceptar el conjunto $\{\delta_1, \dots, \delta_l\}$ de asunciones de normalidad para derivar ϕ desde Θ .
 - Por (6.11), pueden aceptarse tales asunciones de normalidad, pues son consistentes con Θ (requisito de $\diamond$).
 - Finalmente, por (6.12), las asunciones de normalidad tienen un papel complementario de la teoría, no bastan por sí mismas para derivar ϕ .

6.3. El cálculo de δ -resolución

En esta sección presentamos el cálculo de δ -resolución, dual a la resolución clásica propuesta por J. A. Robinson [Rob65], que en principio desarrollamos para abordar problemas abductivos pero que, como posteriormente veremos, muestra su utilidad para otras formas de razonamiento no monótono; concretamente, para el razonamiento por defecto.

Aunque para resolver problemas interesantes de razonamiento por defecto necesitamos al menos una expresividad de primer orden, comenzaremos mostrando brevemente una versión del cálculo de δ -resolución para lógica proposicional. Esta versión es decidible, tal como permite la lógica proposicional, por lo que podremos construir un proceso abductivo decidible mediante δ -resolución proposicional. Posteriormente, indicaremos la extensión a primer orden, en que la decidibilidad se pierde, pero, a pesar de todo, el sistema resulta adecuado

para un gran número de aplicaciones, entre las que se encuentra la resolución de los problemas de razonamiento por defecto que mostraremos en la sección 6.4.

Dado el carácter esquemático de este trabajo, tan sólo proporcionaremos indicaciones de las pruebas de los resultados más interesantes. En cualquier caso, como el cálculo de δ -resolución es dual a la resolución de Robinson, también los son las pruebas de los teoremas principales.

6.3.1. Versión proposicional

A continuación, presentamos el cálculo abductivo para el lenguaje proposicional $\mathcal{L}_p$, definido como es habitual. Comenzamos con algunas definiciones.

Definición 3 (δ -cláusula) Una δ -cláusula Σ es un conjunto de literales

$$\Sigma = \{\lambda_1, \dots, \lambda_n\}$$

y equivale a la conjunción de los literales

$$\lambda_1 \wedge \dots \wedge \lambda_n,$$

de modo que dada una valoración v , v satisface Σ si y sólo si v satisface todos los λ_i , $1 \leq i \leq n$. En adelante, nos referiremos a las δ -cláusulas mediante letras griegas mayúsculas. Con $\square$ nos referimos a la δ -cláusula vacía, universalmente válida.

Definición 4 (Forma δ -clausal) Una forma δ -clausal A es un conjunto de δ -cláusulas

$$A = \{\Sigma_1, \dots, \Sigma_n\}$$

y equivale a la disyunción de sus δ -cláusulas, de forma que dada una valoración v , v satisface A si y sólo si v satisface al menos una de las δ -cláusulas Σ_i , $1 \leq i \leq n$. En adelante, nos referiremos a las formas δ -clausales mediante letras mayúsculas latinas. La forma δ -clausal vacía es no satisfactible⁴.

Las definiciones de δ -cláusula y forma δ -clausal nos permiten extender la notación usada para expresar relaciones de consecuencia lógica en $\mathcal{L}_p$. Así, por

⁴Mientras que la δ -cláusula vacía $\square$ es universalmente válida (puesto que una δ -cláusula se evalúa, por la definición 3, igual que un conjunto de fórmulas), la forma δ -clausal vacía no es satisfactible. Se trata de dos conjuntos vacíos de diferente nivel, pues los elementos que integran las δ -cláusulas y las formas δ -clausales son diferentes (en el primer caso son literales de $\mathcal{L}_p$; en el segundo, δ -cláusulas). Obsérvese que una forma δ -clausal Σ no se evalúa igual que un conjunto de fórmulas, ya que para que una valoración v satisfaga Σ debe cumplirse que haya al menos una δ -cláusula $\gamma \in \Sigma$ que sea satisfecha por v . Esto nunca puede verificarse para la forma δ -clausal vacía, como resulta obvio por no contener ninguna δ -cláusula.

ejemplo, mediante $\Sigma \models A$ expresamos que toda valoración v que satisfaga la δ -cláusula Σ (en el sentido de la definición 3) satisface igualmente la forma δ -clausal A (en el sentido de la definición 4).

Teorema 5 (Conversión a forma δ -clausal) *Para toda fórmula proposicional $\phi \in \mathcal{L}_p$ existe una forma δ -clausal A tal que*

$$\models \phi \leftrightarrow A$$

Prueba informal: Para demostrar este teorema sólo hay que considerar que, por las definiciones 3 y 4, dada cualquier fórmula en forma normal disyuntiva

$$(\gamma_1^1 \wedge \dots \wedge \gamma_1^{j_1}) \vee \dots \vee (\gamma_n^1 \wedge \dots \wedge \gamma_n^{j_n})$$

la forma δ -clausal

$$\{\{\gamma_1^1, \dots, \gamma_1^{j_1}\}, \dots, \{\gamma_n^1, \dots, \gamma_n^{j_n}\}\}$$

es equivalente. Entonces, dada cualquier fórmula ϕ podemos obtener $FND(\phi)$ en forma normal disyuntiva, y desde ahí una forma δ -clausal A_ϕ , según se ha indicado. Entonces, como ϕ y $FND(\phi)$ son equivalentes, y también lo son $FND(\phi)$ y A_ϕ , resulta que ϕ y A_ϕ son equivalentes. ■

Definición 6 (Regla de δ -resolución) *Dadas las δ -cláusulas $\Sigma_1 \cup \{\lambda\}$ y $\Sigma_2 \cup \{\neg\lambda\}$, la regla de δ -resolución se expresa de la siguiente forma:*

$$\frac{\Sigma_1 \cup \{\lambda\} \quad \Sigma_2 \cup \{\neg\lambda\}}{\Sigma_1 \cup \Sigma_2}$$

Decimos que la cláusula $\Sigma_1 \cup \Sigma_2$ es un δ -resolvente de las dos primeras.

Definición 7 (Demostración por δ -resolución) *Una secuencia de δ -cláusulas es una demostración mediante δ -resolución de la δ -cláusula Λ a partir de las δ -cláusulas $\Sigma_1, \dots, \Sigma_n$, lo que expresamos con*

$$\Sigma_1, \dots, \Sigma_n \vdash_\delta \Lambda$$

si y sólo si se cumplen las dos siguientes condiciones:

- Cada una de las δ -cláusulas de la secuencia es o bien una de las Σ_i , $1 \leq i \leq n$ o un δ -resolvente de δ -cláusulas anteriores.
- La secuencia termina con la δ -cláusula Λ .

Los dos siguientes teoremas establecen la corrección y completud débil del cálculo de δ -resolución. Sus pruebas son duales a las correspondientes para el cálculo de resolución proposicional.

Teorema 8 (Corrección) *Para toda forma δ -clausal A y toda δ -cláusula Σ , si $A \vdash_\delta \Sigma$, entonces*

$$\Sigma \models A$$

Teorema 9 (Completud débil) *Si un conjunto de δ -cláusulas A es universalmente válido, entonces*

$$A \vdash_\delta \square$$

Teorema 10 (Completud abductiva) *Sea A una forma δ -clausal no universalmente válida. Entonces, $A \vdash_\delta \Sigma$, para cada δ -cláusula Σ que verifique lo siguiente:*

- Σ es satisfactible.
- $\Sigma \models A$
- No existe ninguna δ -cláusula $\Sigma' \subset \Sigma$ tal que $\Sigma' \models A$

Prueba informal: Sea A una forma δ -clausal no universalmente válida. Consideremos que $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ es una δ -cláusula con las características que indica el enunciado del teorema. Entonces, como $\Sigma \models A$, tenemos que $\models A \cup \{\{\overline{\sigma_1}\}, \dots, \{\overline{\sigma_n}\}\}$, siendo $\overline{\sigma_i}$ el literal complementario de cada σ_i , $1 \leq i \leq n$. Entonces, $A \cup \{\{\overline{\sigma_1}\}, \dots, \{\overline{\sigma_n}\}\} \vdash_\delta \square$, por el teorema 9. Sea $\mathcal{Dem}$ la demostración por δ -resolución de $\square$ a partir de $A \cup \{\{\overline{\sigma_1}\}, \dots, \{\overline{\sigma_n}\}\}$ y construyamos una prueba paralela $\mathcal{Dem}'$, que sólo emplea δ -cláusulas de A . Por tanto, como las δ -cláusulas $\{\overline{\sigma_i}\}$ sólo intervienen eliminando literales σ_i , la δ -cláusula Σ' que demuestra $\mathcal{Dem}'$ sólo contiene literales de A . Por hipótesis, sabemos que no puede darse $\Sigma' \subset \Sigma$, así que $\Sigma' = \Sigma$, y como se ha obtenido exclusivamente a partir de δ -cláusulas de A , tenemos que $A \vdash_\delta \Sigma$. ■

Los dos siguientes resultados nos permiten eliminar, de las pruebas mediante δ -resolución, las δ -cláusulas *subsumidas* por otras y las no satisfactibles. La noción de subsunción es igual que en el cálculo de resolución. Decimos que una δ -cláusula Σ está subsumida por otra Σ' si y sólo si⁵ $\Sigma' \subset \Sigma$. Las pruebas de ambos corolarios son duales a las de los resultados equivalentes en el cálculo de resolución.

⁵En ocasiones se considera que Σ está subsumida por Σ' si y sólo si $\Sigma' \subseteq \Sigma$, y entonces se eliminan las δ -cláusulas subsumidas por otras *diferentes*, ya que entonces toda δ -cláusula está subsumida por sí misma.

Corolario 11 (Regla de subsunción) *Dada una δ -cláusula Γ , si $A \cup \{\Sigma\} \cup \{\Sigma \cup \Lambda\} \vdash_{\delta} \Gamma$, entonces existe una δ -cláusula Γ' tal que $A \cup \{\Sigma\} \vdash_{\delta} \Gamma'$ y $\Gamma' \subseteq \Gamma$.*

Corolario 12 (Eliminación de δ -cláusulas contradictorias) *Para toda δ -cláusula Σ satisfactible y todo conjunto de δ -cláusulas B , si $B \cup \{\{\lambda, \neg\lambda, \gamma_1, \dots, \gamma_n\}\} \vdash_{\delta} \Sigma$, entonces $B \vdash_{\delta} \Sigma$.*

Definición 13 (Saturación) *Dada la forma δ -clausal A , el conjunto saturación por δ -resolución desde A , que llamamos A^{δ} , es el más pequeño que contiene todas las δ -cláusulas satisfactibles demostrables mediante δ -resolución a partir de A que no están subsumidas por otras δ -cláusulas también demostrables por δ -resolución desde A .*

Intuitivamente, podemos construir A^{δ} aplicando la regla de δ -resolución tantas veces como sea posible a las δ -cláusulas de A , así como a las que vayan apareciendo. Finalmente, se eliminan del conjunto resultante todas las δ -cláusulas no satisfactibles o que resulten subsumidas por otras que hayan aparecido. De esta forma, para cualquier conjunto finito de δ -cláusulas A , el proceso de construcción de A^{δ} termina siempre en un número finito de pasos.

Corolario 14 (Propiedad fundamental de la saturación) *Sea A una forma δ -clausal. Entonces, A^{δ} contiene solamente cada δ -cláusula Σ tal que*

- Σ es satisfactible.
- $\Sigma \models A$
- No existe ninguna δ -cláusula $\Sigma' \subset \Sigma$ tal que $\Sigma' \models A$

Prueba informal: Este corolario es un resultado casi inmediato a partir de los teoremas 8 y 10 y de la definición 13. ■

Definición 15 (Solución abductiva minimal) *Dado el problema abductivo $\langle \Theta, \phi \rangle$, decimos que una δ -cláusula $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ es una solución abductiva minimal si y sólo si se verifican:*

- $\sigma_1 \wedge \dots \wedge \sigma_n$ es una solución abductiva al problema abductivo $\langle \Theta, \phi \rangle$, en el sentido de la definición 2.
- No existe ninguna δ -cláusula $\Sigma' \subset \Sigma$ que sea también una solución abductiva para $\langle \Theta, \phi \rangle$.

Al conjunto de δ -cláusulas que son soluciones abductivas minimales del problema abductivo $\langle \Theta, \phi \rangle$ lo llamamos $\mathcal{Ab}\delta(\Theta, \phi)$.

Teorema 16 Dado el problema abductivo $\langle \Theta, \phi \rangle$, $\Theta = \{\theta_1, \dots, \theta_n\}$, si N_Θ es la forma δ -clausal de $\neg(\theta_1 \wedge \dots \wedge \theta_n)$ y O la de ϕ , se verifica

$$\mathcal{Ab}\delta(\Theta, \phi) = (N_\Theta^\delta \cup O^\delta)^\delta - (N_\Theta^\delta \cup O^\delta)$$

Prueba informal: En primer lugar, podemos demostrar que toda δ -cláusula de $\mathcal{Ab}\delta(\Theta, \phi)$ pertenece a $(N_\Theta^\delta \cup O^\delta)^\delta - (N_\Theta^\delta \cup O^\delta)$. Sea $\Sigma = \{\sigma_1, \dots, \sigma_n\}$ una δ -cláusula de $\mathcal{Ab}\delta(\Theta, \phi)$. Entonces, por la definición 15, $\sigma_1 \wedge \dots \wedge \sigma_n$ es una solución abductiva al problema abductivo $\langle \Theta, \phi \rangle$, en el sentido de la definición 2, lo cual implica:

- $\theta_1 \wedge \dots \wedge \theta_n, \sigma_1 \wedge \dots \wedge \sigma_n \models \phi$. Pero entonces, $\sigma_1 \wedge \dots \wedge \sigma_n \models \theta_1 \wedge \dots \wedge \theta_n \rightarrow \phi$. Entonces, si C es la forma δ -clausal de $\theta_1 \wedge \dots \wedge \theta_n \rightarrow \phi$ puede probarse que $\Sigma \in C^\delta$, así como que C es equivalente a $N_\Theta^\delta \cup O^\delta$, y por tanto $\Sigma \in (N_\Theta^\delta \cup O^\delta)^\delta$.
- $\theta_1 \wedge \dots \wedge \theta_n, \sigma_1 \wedge \dots \wedge \sigma_n \not\models \perp$. Por tanto, $\sigma_1 \wedge \dots \wedge \sigma_n \not\models \neg(\theta_1 \wedge \dots \wedge \theta_n)$, y entonces $\Sigma \notin N_\Theta^\delta$.
- $\sigma_1 \wedge \dots \wedge \sigma_n \not\models \phi$. Entonces, $\Sigma \notin O^\delta$.

Los tres resultados a los que hemos llegado aseguran que $\Sigma \in (N_\Theta^\delta \cup O^\delta)^\delta - (N_\Theta^\delta \cup O^\delta)$.

Ahora, supongamos $\Sigma \in (N_\Theta^\delta \cup O^\delta)^\delta - (N_\Theta^\delta \cup O^\delta)$. La prueba de $\Sigma \in \mathcal{Ab}\delta(\Theta, \phi)$ puede realizarse siguiendo los tres pasos anteriores en sentido inverso. ■

Los resultados anteriores nos permiten definir un proceso abductivo para lógica proposicional que sigue los siguientes pasos, para cualquier conjunto de fórmulas $\Theta = \{\theta_1, \dots, \theta_n\}$ y cualquier fórmula ϕ :

Análisis de la teoría. Se obtiene N_Θ , forma δ -clausal de $\neg(\theta_1 \wedge \dots \wedge \theta_n)$.

Entonces, si en N_Θ no hay ninguna δ -cláusula satisfactible es que Θ es universalmente válida, y el proceso acaba, puesto que en tal caso $\langle \Theta, \phi \rangle$ no puede ser un problema abductivo. En otro caso, se obtiene N_Θ^δ , y si $\square \in N_\Theta^\delta$, entonces Θ es no satisfactible, y el proceso acaba. En otro caso,

Análisis de la observación. Se obtiene O , forma δ -clausal de ϕ . Entonces, si en O no hay ninguna δ -cláusula satisfactible es que ϕ no es satisfactible, y el proceso acaba. En otro caso, se obtiene O^δ , y si $\square \in O^\delta$, entonces ϕ es universalmente válida, y el proceso acaba. En otro caso,

Búsqueda de refutaciones. Si toda δ -cláusula de O^δ pertenece a N_Θ^δ o está subsumida por alguna de sus δ -cláusulas entonces es que $\Theta \models \neg\phi$ (esto puede probarse mediante un sencillo razonamiento semántico), y en tal caso la observación refuta la teoría. En otro caso,

Búsqueda de explicaciones. A partir de N_Θ^δ y O^δ se forma $(N_\Theta^\delta \cup O^\delta)$, y desde ahí $(N_\Theta^\delta \cup O^\delta)^\delta$. Si $\square$ pertenece a este último conjunto, entonces $\Theta \models \phi$, con lo que la teoría explica la observación. En otro caso, se construye $Ab\delta(\Theta, \phi) = (N_\Theta^\delta \cup O^\delta)^\delta - (N_\Theta^\delta \cup O^\delta)$, con lo que tenemos todas las soluciones abductivas minimales al problema abductivo $\langle \Theta, \phi \rangle$.

En el proceso abductivo que acabamos de presentar encontramos algunos aspectos epistemológicos bastante interesantes, que se asemejan a ciertas ideas provenientes de la filosofía de la ciencia. En primer lugar, hay dos rasgos refutacionistas que merece la pena resaltar. Uno de ellos es que el proceso comienza buscando los puntos débiles de la teoría, representados por las δ -cláusulas que se obtienen tras el primer caso; cada una de ellas contiene una posibilidad de refutar la teoría. Además, tras el segundo paso, de análisis de la observación (análisis cartesiano, si se quiere; se trata de reducir la observación a los elementos mínimos necesarios para explicarla) volvemos a encontrar otro aspecto refutacionista, pues antes de buscar las explicaciones se comprueba si la observación va o no en contra de la teoría.

Otro rasgo interesante es el carácter directo del proceso explicativo. Otros sistemas abductivos (como la abducción basada en tablas semánticas, o incluso en resolución al modo tradicional) trabajan obteniendo conclusiones que deducen de la teoría junto con la negación de la observación y después negándolas, de donde se obtienen las explicaciones. Esta forma de proceder, indirecta, tiene mucho que ver con la reducción al absurdo, pero no parece cercana al modo natural en que la mente humana trabaja cuando busca una explicación razonable a una observación nueva. Sin embargo, la δ -resolución obtiene directamente las explicaciones, pues cada δ -cláusula que resulta del proceso es ya una solución abductiva.

Por último, la versión proposicional del proceso abductivo mediante δ -resolución es decidible, como se deriva de los resultados presentados. Por tanto, dadas cualesquiera teoría Θ y observación ϕ , podemos saber:

- Si Θ es no satisfactible, universalmente válida o contingente (meramente satisfactible).
- Si ϕ es no satisfactible, universalmente válida o contingente (meramente satisfactible).

- Si $\Theta \models \neg\phi$ (refutación) o $\Theta \models \phi$ (observación ya explicada) o bien $\langle \Theta, \phi \rangle$ es un problema abductivo. En este último caso podemos obtener las soluciones abductivas minimales.

Veamos un pequeño ejemplo para ilustrar la resolución de un problema abductivo mediante δ -resolución. Volvamos al problema abductivo presentado al final de la página 92 (segunda versión, con teoría ampliada). Entonces,

$$\langle \Theta, \phi \rangle = \langle \{l \rightarrow c, a \rightarrow c, c \rightarrow z, \neg l\}, z \rangle$$

Veamos cada paso del proceso abductivo:

Análisis de la teoría. En este caso N_Θ (forma δ -clausal de $\neg((l \rightarrow c) \wedge (a \rightarrow c) \wedge (c \rightarrow z) \wedge \neg l)$) es

$$\{\{l, \neg c\}, \{a, \neg c\}, \{c, \neg z\}, \{l\}\}$$

Como en N_Θ hay δ -cláusulas satisfactibles, tenemos que Θ no es universalmente válida. Además,

$$N_\Theta^\delta = \{\{a, \neg z\}, \{a, \neg c\}, \{c, \neg z\}, \{l\}\}$$

y como $\Box \notin N_\Theta^\delta$, entonces Θ es meramente satisfactible.

Análisis de la observación. El conjunto O , forma δ -clausal de ϕ es $\{\{z\}\}$.

Como en O hay una δ -cláusula satisfactible, ϕ es satisfactible. Además, $O^\delta = O$, y como $\Box \notin O^\delta$, entonces ϕ es meramente satisfactible.

Búsqueda de refutaciones. Como hay al menos una δ -cláusula de O^δ , $\{z\}$, que no pertenece a N_Θ^δ ni está subsumida por alguna de sus δ -cláusulas entonces es que $\Theta \not\models \phi$, con lo que no hay refutación.

Búsqueda de explicaciones. Tenemos que:

- $(N_\Theta^\delta \cup O^\delta) = \{\{a, \neg z\}, \{a, \neg c\}, \{c, \neg z\}, \{l\}, \{z\}\}$
- $(N_\Theta^\delta \cup O^\delta)^\delta = \{\{a\}, \{c\}, \{z\}, \{l\}\}$, y como $\Box \notin (N_\Theta^\delta \cup O^\delta)^\delta$, entonces $\Theta \not\models \phi$.
- $\mathcal{Ab}\delta(\Theta, \phi) = (N_\Theta^\delta \cup O^\delta)^\delta - (N_\Theta^\delta \cup O^\delta) = \{\{a\}, \{c\}\}$. Por tanto, las dos soluciones abductivas minimales son a y c .

Fácilmente pueden incorporarse al proceso abductivo criterios preferenciales de selección de la mejor explicación. Así, en el ejemplo anterior se puede emplear un criterio basado en la historia de las δ -cláusulas para determinar que a es mejor explicación que c , pues tiene mayor historia, es decir, han sido necesarias para obtenerla más δ -cláusulas que para obtener c , por lo que se ha usado una mayor porción de la teoría.

6.3.2. Extensión a primer orden

En este apartado presentamos algunos aspectos de la extensión del cálculo de δ -resolución, así como del proceso abductivo mostrado, a lógica de primer orden. Asumiremos que $\mathcal{L}$ es un lenguaje de primer orden con las convenciones habituales. Para la semántica sólo se asignarán valores de verdad a las sentencias, o fórmulas de $\mathcal{L}$ sin variables libres. Una $\mathcal{L}$ -estructura $M = \langle \mathcal{D}, \mathfrak{I} \rangle$ está compuesta por el dominio no vacío $\mathcal{D}$ y la función interpretación $\mathfrak{I}$, definida como es habitual.

Definición 17 *Para cualquier $\alpha \in \mathcal{L}$, representamos con $\text{sko}(\alpha)$ la forma normal de Skolem de α .*

Dual a la forma normal de Skolem, definimos la forma normal de Herbrand, tomada de [Bus95]:

Definición 18 *Decimos que una fórmula $\alpha \in \mathcal{L}$ está en forma normal de Herbrand si y sólo si:*

1. α está en forma prenexa.
2. En α no aparecen cuantificadores universales.

Dada cualquier sentencia $\alpha \in \mathcal{L}$, definimos el procedimiento de transformación de α a su forma normal de Herbrand, que llamaremos $\text{her}(\alpha)$, de la siguiente manera:

1. A partir de α , obtenemos su forma prenexa, que llamaremos α_p .
2. Se suprimen los cuantificadores universales de α_p comenzando por el que está más a la izquierda y continuando sucesivamente hacia la derecha, observando estas reglas:
 - a) Los cuantificadores universales que no tengan a la izquierda ningún cuantificador existencial se eliminan y se sustituye cada ocurrencia de la variable que cuantifican por un parámetro que no ocurra en la fórmula.
 - b) Los cuantificadores universales que tengan a la izquierda algún cuantificador existencial se eliminan igualmente, pero en este caso se sustituye cada ocurrencia de la variable que cuantifican por un término que consiste en un functor (que no debe haberse usado antes) que tiene como argumentos todas las variables que aparecen cuantificadas existencialmente a la izquierda del cuantificador universal que se elimina.

El resultado de seguir los pasos anteriores es $\text{her}(\alpha)$, la forma normal de Herbrand de α . A los términos que se introducen en los pasos 2a y 2b los llamaremos términos de Skolem⁶.

Las propiedades de la forma normal de Herbrand son duales a las de la forma normal de Skolem. También el siguiente resultado puede demostrarse de forma dual a como se probaría el resultado correspondiente (también dual) para formas de Skolem.

Teorema 19 *Dada cualquier sentencia $\alpha \in \mathcal{L}$, se verifican:*

$$\begin{aligned} \alpha &\models \text{her}(\alpha) \\ \models \alpha &\text{ si y sólo si } \models \text{her}(\alpha) \end{aligned}$$

Definición 20 *Una δ -cláusula Σ es un conjunto de literales de $\mathcal{L}$,*

$$\Sigma = \{\lambda_1, \dots, \lambda_n\}$$

Si $\{x_1, \dots, x_m\}$, $m \geq 0$, es el conjunto de variables libres que aparecen en Σ , entonces, dada una $\mathcal{L}$ -estructura M , consideramos que M satisface Σ si y sólo si $M \models \exists x_1, \dots, x_m (\lambda_1 \wedge \dots \wedge \lambda_n)$. Entonces escribimos $M \models \Sigma$, como una abreviatura de lo anterior, ya que la semántica que usamos no contempla la interpretación de fórmulas con variables libres. Si $M \not\models \exists x_1, \dots, x_m (\lambda_1 \wedge \dots \wedge \lambda_n)$ escribimos (también como abreviatura) $M \not\models \Sigma$, y decimos que M no satisface Σ .

En adelante, nos referiremos a las δ -cláusulas mediante letras griegas mayúsculas. Con $\square$ nos referimos a la δ -cláusula vacía, universalmente válida.

Definición 21 *Una forma δ -clausal A es un conjunto de δ -cláusulas*

$$A = \{\Sigma_1, \dots, \Sigma_n\}$$

Dada una $\mathcal{L}$ -estructura M , consideramos que M satisface A , lo que escribimos $M \models A$ (en caso contrario, $M \not\models A$) si y sólo si $M \models \Sigma_i$, para alguna $\Sigma_i \in A$.

En adelante, nos referiremos a las formas δ -clausales mediante letras mayúsculas latinas. La forma δ -clausal vacía es no satisfactible.

Definición 22 *Dada una sentencia $\alpha \in \mathcal{L}$, definimos el procedimiento de transformación de α a A , su forma δ -clausal, que consta de los siguientes pasos:*

⁶Usamos la misma denominación que para los términos introducidos durante la transformación a forma normal de Skolem.

1. A partir de α obtenemos su forma normal de Herbrand, $her(\alpha)$.
2. Empleamos equivalencias proposicionales para transformar la matriz de $her(\alpha)$ a forma normal disyuntiva⁷. El resultado, que llamaremos $her(\alpha)_d$ tendrá la forma

$$\exists x_1, \dots, x_m ((\lambda_{1_1} \wedge \dots \wedge \lambda_{1_{j_1}}) \vee \dots \vee (\lambda_{n_1} \wedge \dots \wedge \lambda_{n_{j_n}}))$$

siendo n el número de conjunciones elementales de su matriz, y cada λ_{i_k} el k -ésimo literal, $1 \leq k \leq j_i$, de la i -ésima conjunción elemental, $1 \leq i \leq n$.

3. A partir de $her(\alpha)_d$, construimos directamente A , de la forma:

$$A = \{ \{ \lambda_{1_1}, \dots, \lambda_{1_{j_1}} \}, \dots, \{ \lambda_{n_1}, \dots, \lambda_{n_{j_n}} \} \}$$

es decir, lo que hacemos es quitar los cuantificadores iniciales y construir una δ -cláusula por cada una de las conjunciones elementales que aparecen en la matriz de $her(\alpha)_d$.

El siguiente teorema es una consecuencia de resultados anteriores y de considerar que la transformación a forma δ -clausal consiste únicamente en la obtención de la forma normal de Herbrand y la aplicación de ciertas equivalencias. Desde ahí, el corolario que viene a continuación es una consecuencia inmediata.

Teorema 23 *Dada cualquier sentencia $\alpha \in \mathcal{L}$, su forma δ -clausal, A , es equivalente a $her(\alpha)$.*

Corolario 24 *Dada cualquier sentencia $\alpha \in \mathcal{L}$ cuya forma δ -clausal es A , se verifica que $\models A$ si y sólo si $\models \alpha$.*

A continuación introducimos el cálculo de δ -resolución en lógica de primer orden. Pero antes, como la regla de δ -resolución se aplicará a δ -cláusulas con conjuntos disjuntos de variables, definimos la operación de renombramiento de variables.

Definición 25 *Dada la δ -cláusula Σ cuyas variables son $\{x_1, \dots, x_n\}$, definimos el procedimiento de renombramiento de variables de Σ como la obtención de $\Sigma(x_1, \dots, x_n/y_1, \dots, y_n)$, siendo cada y_i , $1 \leq i \leq n$, una variable que no aparece en Σ .*

⁷Aunque estemos ante una fórmula de primer orden, siguen siendo válidas las equivalencias proposicionales, ya que la semántica de $\mathcal{L}$ es una extensión de la de $\mathcal{L}_p$.

Definición 26 Dadas $\Sigma_1 \cup \{\lambda\}$ y $\Sigma_2 \cup \{\neg\gamma\}$, dos δ -cláusulas cuyos conjuntos de variables son disjuntos, y además σ es un unificador de máxima generalidad de $\{\lambda, \gamma\}$, se define la regla de δ -resolución de la siguiente forma:

$$\frac{\Sigma_1 \cup \{\lambda\} \quad \Sigma_2 \cup \{\neg\gamma\}}{\Sigma_1\sigma \cup \Sigma_2\sigma}$$

De la δ -cláusula $\Sigma_1\sigma \cup \Sigma_2\sigma$ decimos que es un δ -resolvente de las δ -cláusulas $\Sigma_1 \cup \{\lambda\}$ y $\Sigma_2 \cup \{\neg\gamma\}$.

Definición 27 Decimos que la δ -cláusula Λ puede demostrarse mediante δ -resolución a partir de las δ -cláusulas $\Sigma_1, \dots, \Sigma_n$, lo que representamos como

$$\Sigma_1, \dots, \Sigma_n \vdash_{\delta} \Lambda$$

si y sólo si existe una secuencia de δ -cláusulas, a la que llamamos demostración mediante δ -resolución de Λ a partir de $\Sigma_1, \dots, \Sigma_n$, que cumple los siguientes requisitos:

- Cada una de las δ -cláusulas de la secuencia es:
 1. Una de las Σ_i , $1 \leq i \leq n$, o bien
 2. Una cláusula $\Sigma\sigma$ obtenida por renombramiento de variables de otra δ -cláusula Σ anterior, o bien
 3. Un δ -resolvente de δ -cláusulas anteriores.
- La secuencia termina con la δ -cláusula Λ .

Los teoremas que siguen, de corrección y completud, son una extensión de los que presentamos para lógica proposicional. La corrección se demuestra de forma dual a la corrección del cálculo de resolución clásico. Para probar la completud, necesitaríamos demostrar una versión dual del Teorema de Herbrand, o bien recurrir a la dualidad de la δ -resolución con la resolución de Robinson y obtener la completud del cálculo de δ -resolución como corolario de la completud de la resolución.

Teorema 28 (Corrección de la δ -resolución) Sea la fórmula $\alpha \in \mathcal{L}$ tal que su forma δ -clausal A verifica $A \vdash_{\delta} \square$. Entonces, $\models \alpha$.

Teorema 29 (Completud de la δ -resolución) Sea $\alpha \in \mathcal{L}$ una sentencia tal que $\models \alpha$. Entonces, si A es la forma δ -clausal de α , se verifica $A \vdash_{\delta} \square$.

Definición 30 (δ -explicación) Sea $\alpha \in \mathcal{L}$ una sentencia, cuya forma δ -clausal es A , tal que $\not\models \alpha$. Entonces, decimos que la sentencia $\beta \in \mathcal{L}$ es una δ -explicación de α si se cumplen las siguientes condiciones:

1. β tiene la forma $Q_1x_1 \dots Q_{n_1}x_{n_1}(\lambda_1 \wedge \dots \wedge \lambda_{m_1})$, $n_1 \geq 0$, siendo cada λ_i , $1 \leq i \leq m_1$, un literal, y cada Q_j , $1 \leq j \leq n_1$, un cuantificador con la variable x_j como índice.
2. Es posible construir $\text{sko}(\beta) = \forall y_1, \dots, y_{n_2}(\lambda_1 \wedge \dots \wedge \lambda_{m_1})\sigma$, $n_2 \leq n$, siendo σ la sustitución que asigna a cada una de las variables cuantificadas existencialmente en β un término nuevo, de modo que ninguno⁸ de los funtores ni constantes que aparecen en A pero no en α esté en $\text{sko}(\beta)$.
3. Existe una δ -cláusula $\Sigma = \{\gamma_1, \dots, \gamma_{m_2}\}$ tal que:
 - a) $A \vdash_\delta \Sigma$.
 - b) Dada la forma normal de Skolem $\text{sko}(\beta) = \forall y_1, \dots, y_{n_2}(\lambda_1 \wedge \dots \wedge \lambda_{m_1})\sigma$ indicada en el punto 2 anterior, para cada $\gamma_k \in \Sigma$, $1 \leq k \leq m_2$ existe un $\lambda_i\sigma$, $1 \leq i \leq m_1$, tal que para alguna sustitución posible π_k se cumple $\lambda_i\sigma\pi_k = \gamma_k$.

Para obtener δ -explicaciones resulta interesante el algoritmo que P.T. Cox y T. Pietrzykowski [CP84] proporcionan para realizar *skolemización inversa*. Así, dado un literal λ que contiene términos de Skolem y las variables libres $x_1, \dots, x_n$, devuelve un conjunto C de literales cuantificados y sin variables libres con las siguientes propiedades:

Corrección: si $\sigma \in C$, entonces $\models \sigma \rightarrow \exists x_1, \dots, x_n(\lambda)$.

Compleitud: Para cada sentencia $\alpha \in \mathcal{L}$, si $\models \alpha \rightarrow \lambda$, entonces existe un literal cuantificado $\sigma \in C$ tal que $\models \alpha \rightarrow \sigma$.

No redundancia: Si $\sigma, \alpha \in C$, entonces $\not\models \sigma \rightarrow \alpha$ y $\not\models \alpha \rightarrow \sigma$.

Mediante una adaptación de este algoritmo a conjunciones de literales es posible construir las δ -explicaciones. Además, la propiedad de *no redundancia* es deseable para evitar δ -explicaciones no minimales, como ocurría en el ejemplo anterior. M. Cialdea y F. Pirri [CP93] presentan una extensión del algoritmo de Cox y Pietrzykowski que puede aplicarse a conjunciones de literales, y por tanto a la construcción de δ -explicaciones. A continuación proporcionamos una adaptación de su definición.

Definición 31 (Skolemización inversa) Sea $\alpha \in \mathcal{L}$ una fórmula en forma normal de Skolem y sea $ts(\alpha) = \{t_1, \dots, t_k\}$ el conjunto de términos de Skolem y variables libres que aparecen en α . Sea $\langle t_{p_1}, \dots, t_{p_k} \rangle$ un orden total de los elementos de $ts(\alpha)$ de forma que para cada i y j , si t_{p_i} ocurre en t_{p_j} entonces $i <$

⁸Este requisito hace que ninguno de los términos de Skolem que se introdujeron en A al hacer $her(\alpha)$ ni sus funtores correspondientes aparezcan ni en β ni en $\text{sko}(\beta)$.

j. Entonces $Q_1x_1 \dots Q_kx_k\alpha'$, $Q_l \in \{\forall, \exists\}$, se obtiene desde α por skolemización inversa a partir de $\langle t_{p_1}, \dots, t_{p_k} \rangle$ si y sólo si

- *α' se obtiene desde α reemplazando cada término t_{p_i} por la variable nueva x_i , y además*
- *Para cada i , si t_{p_i} es una variable, entonces Q_i es un cuantificador existencial; en otro caso, si t_{p_i} es un término de Skolem, Q_i es un cuantificador universal.*

Al conjunto de todas las fórmulas que pueden obtenerse por skolemización inversa desde α lo llamaremos $\text{desk}(\alpha)$.

Empleando la noción de δ -explicación podemos probar los dos siguientes resultados, que serán fundamentales para la definición del proceso abductivo mediante δ -resolución en primer orden.

Teorema 32 (Corrección abductiva) *Sea $\alpha, \beta \in \mathcal{L}$ dos sentencias tales que $\not\models \alpha$ y β es una δ -explicación de α . Entonces, $\beta \models \alpha$.*

Teorema 33 (Compleitud abductiva) *Sea $\alpha, \beta \in \mathcal{L}$ dos sentencias tales que:*

- *$\not\models \alpha$*
- *$\beta \models \alpha$*
- *β es una fórmula en forma prenexa cuya matriz es una conjunción de literales.*

Entonces, β es una δ -explicación de α .

La construcción de un proceso abductivo mediante δ -resolución en primer orden presenta algunos problemas que no aparecían en lógica proposicional. El proceso que entonces presentamos se basaba en la noción de saturación por δ -resolución. Sin embargo en lógica de primer orden, dada su indecidibilidad, la saturación de los conjuntos de δ -cláusulas no queda por lo general asegurada. Es más, la introducción de términos de Skolem durante la conversión a forma δ -clausal (al realizar la transformación a forma normal de Herbrand) impide ciertas unificaciones, con lo que no es posible, por lo general, reconocer durante el proceso abductivo situaciones en que la observación refuta la teoría, o bien en que las abducciones generadas (que serán δ -cláusulas a las que habrá que someter a *skolemización inversa* para obtener las δ -explicaciones correspondientes) no son explicativas.

Para ilustrar lo que acabamos de comentar, veamos un ejemplo muy simple. Sea la teoría $\Theta = \{\exists xTx\}$ y la observación $\phi = \forall y\neg Ty$. En este caso, la observación refuta la teoría, puesto que $\Theta \models \neg\phi$. Recordemos que en el proceso abductivo proposicional esta situación se reconocía porque, tras obtener N_Θ^δ y O^δ mediante saturación por δ -resolución de las formas δ -clausales de $\neg\Theta$ y ϕ , respectivamente, cada δ -cláusula de O^δ estaba subsumida por una de N_Θ^δ , o bien pertenecía a este conjunto. Si trasladamos esto a nuestro ejemplo en primer orden, tenemos que $N_\Theta^\delta = \{\{\neg Th_1\}\}$ y $O^\delta = \{\{\neg Th_2\}\}$, siendo h_1 y h_2 los términos de Skolem introducidos durante la transformación a forma normal de Herbrand. Como se observa, estos términos impiden que $\{\neg Th_2\}$ pueda ser subsumida por $\{\neg Th_1\}$, con lo que no es posible constatar durante el proceso abductivo que $\Theta \models \neg\phi$.

Sin embargo, sí es posible verificarlo en un proceso independiente. De hecho,

$$\Theta \models \neg\phi$$

si y sólo si

$$\models \Theta \rightarrow \phi$$

si y sólo si

$$\models \{\{\neg Th_1\}, \{Ty\}\}$$

al ser $\{\{\neg Th_1\}, \{Ty\}\}$ la forma δ -clausal de $\Theta \rightarrow \phi$. Pero esto último se verifica, ya que con una sola aplicación de la regla de δ -resolución se alcanza $\square$.

Del mismo modo también se puede comprobar si las abducciones que se obtienen son explicativas, aunque siempre en procesos independientes, y con las limitaciones propias de la lógica de primer orden. Es decir, a cada δ -cláusula que se obtenga puede aplicarse skolemización inversa, y comprobar luego si las δ -explicaciones obtenidas cumplen los requisitos para ser consideradas abducciones explicativas.

La opción que nos parece más razonable es adaptar el proceso proposicional de la siguiente manera:

Análisis de la teoría. Dada la teoría Θ (conjunción de sentencias de $\mathcal{L}$) se obtiene la forma δ -clausal de $\neg\Theta$, que llamaremos N_Θ . Entonces se aplica δ -resolución a las δ -cláusulas de N_Θ . Si el proceso llega a saturarse (no es posible una aplicación más de la regla de δ -resolución en que aparezca una δ -cláusula que no está ya), sea N_Θ^δ el conjunto obtenido. Entonces,

- Si $\square \in N_\Theta^\delta$, entonces la teoría es no satisfactible.
- En otro caso, continúa el proceso.

Análisis de la observación. Dada la observación ϕ (sentencia de $\mathcal{L}$), se obtiene su forma δ -clausal, O , a la que se aplica δ -resolución. Si al proceso se satura, sea O^δ el conjunto obtenido. Entonces,

- Si $\square \in O^\delta$ entonces la observación es universalmente válida.
- En otro caso, continúa el proceso.

Búsqueda de refutaciones. Se eliminan de O^δ todas las δ -cláusulas subsumidas⁹ por alguna δ -cláusula de N_Θ^δ . Entonces,

- Si no queda ninguna δ -cláusula en O^δ entonces la observación refuta la teoría.
- En otro caso, continúa al proceso, siendo B el conjunto unión de las δ -cláusulas de N_Θ^δ y aquellas de O^δ no subsumidas.

Búsqueda de explicaciones. Se aplica δ -resolución a las δ -cláusulas de B . Si el proceso se satura, sea B^δ el conjunto obtenido. Entonces,

- Si $\square \in B^\delta$, entonces es que la observación es consecuencia lógica de la teoría.
- En otro caso, se devuelven aquellas δ -cláusulas de B^δ que no están en B como posibles abducciones explicativas. No sirven las de B puesto que todas ellas satisfacen a $\neg\Theta$ o a ϕ , con lo que no son explicativas.

A cada una de las δ -cláusulas obtenidas habrá que aplicar skolemización inversa para obtener las δ -explicaciones, y a su vez (para asegurar la corrección del proceso) verificar para cada δ -explicación si es realmente una abducción explicativa. Ello se realiza comprobando mediante δ -resolución (o por otro método de decisión) que para cada δ -explicación α no se verifican ni $\Theta \models \neg\alpha$ ni $\alpha \models \phi$. Si se elige la δ -resolución como método de decisión, esto supone verificar que no se obtiene $\square$ en ninguna de las dos búsquedas que hay que hacer. Pero de nuevo, la indecidibilidad evita que ambas condiciones puedan comprobarse por lo general.

El proceso anterior se ha definido contando con que los procesos de δ -resolución se saturaban. Sin embargo, en ciertas ocasiones, como hemos explicado, puede que esto no ocurra. Entonces debe elegirse algún otro criterio para detener las búsquedas y pasar a la siguiente etapa del proceso abductivo. Si estamos ante una implementación, tal criterio puede ser el número máximo de δ -cláusulas generadas, o bien un límite en la memoria de trabajo, etc.

A pesar de las limitaciones comentadas, existen muchos problemas abductivos en primer orden que pueden resolverse con éxito mediante δ -resolución, como veremos en la siguiente sección.

⁹En lógica de primer orden, una δ -cláusula Σ_1 está subsumida por otra δ -cláusula Σ_2 si y sólo si existe una sustitución σ tal que $\Sigma_2\sigma \subseteq \Sigma_1$.

6.4. Razonamiento por defecto mediante δ -resolución

Como mostramos en la sección 6.2, los problemas de razonamiento por defecto pueden comprenderse como problemas abductivos. Terminamos este trabajo presentando la solución mediante δ -resolución de algunos problemas sencillos. Comenzamos por el caso más simple; nuestra teoría será:

$$\Theta = \{\forall x(Ave(x) \wedge Ave_normal_vuela(x) \rightarrow Vuela(x)), Ave(a), Pingüino(a)\}$$

y lo que nos preguntamos es si resulta posible asumir por defecto que a vuela, es decir, $\phi = Vuela(a)$. El predicado $Ave_normal_vuela(x)$ es una asunción de normalidad, como vimos más arriba. Como hemos comentado, podemos buscar soluciones abductivas al problema abductivo $\langle \Theta, \phi \rangle$ y, si hay alguna compuesta sólo de asunciones de normalidad y sin variables libres, entonces la respuesta será positiva, y solamente entonces. El proceso abductivo transcurre de la siguiente manera (por brevedad sólo mostramos el resultado de cada paso):

Análisis de la teoría. La teoría es contingente (meramente satisfactible). Tenemos que

$$N_{\Theta}^{\delta} = \{\{Ave(x), Ave_normal_vuela(x), \neg Vuela(x)\}, \{\neg Ave(a)\}, \\ \{Ave_normal_vuela(a), \neg Vuela(a)\}, \{\neg Pingüino(a)\}\}$$

Análisis de la observación. Observación contingente: $O^{\delta} = \{Vuela(a)\}$.

Búsqueda de refutaciones. No se encuentran.

Búsqueda de explicaciones. Como fácilmente puede observarse, llegamos en un solo paso a la δ -cláusula $\{Ave_normal_vuela(a)\}$, que no pertenece a $N_{\Theta}^{\delta} \cup O^{\delta}$ pero sí se obtiene mediante δ -resolución desde dicho conjunto. Como esta δ -cláusula está compuesta sólo por una asunción de normalidad respecto de a , la solución al problema original es que por defecto podemos asumir que a vuela.

Compliquemos un poco el problema, tal como hicimos anteriormente. Ahora,

$$\Theta' = \Theta \cup \{\forall x(Pingüino(x) \rightarrow \neg Ave_normal_vuela(x))\}$$

y entonces:

Análisis de la teoría. La teoría, sigue siendo contingente. En este caso:

$$N_{\Theta'}^{\delta} = \{\{Ave(x), Ave_normal_vuela(x), \neg Vuela(x)\}, \\ \{\neg Ave(a)\}, \{\neg Pingüino(a)\}, \{Pingüino(x), Ave_normal_vuela(x)\}, \\ \{Ave_normal_vuela(a)\}\}$$

Análisis de la observación. Contingente, y también $O^\delta = \{\{Vuela(a)\}\}$.

Búsqueda de refutaciones. No se encuentran.

Búsqueda de explicaciones. La única δ -cláusula compuesta sólo de asunciones de normalidad que pertenece a $(N_\Theta^\delta \cup O^\delta)^\delta$ es $\{Ave_normal_vuela(a)\}$, que ya está en $N_\Theta^\delta \cup O^\delta$, por lo que no es una solución abductiva. Por tanto, la solución al problema original es que ahora, con la teoría Θ' no podemos inferir, por defecto, que a vuela.

Podemos construir una tercera teoría:

$$\Theta'' = \Theta' \cup \{\forall x(\neg Ave_normal_vuela(x) \rightarrow \neg Vuela(x))\}$$

Esta teoría aporta a Θ' la información de que las aves que no son normales en cuanto a la capacidad de volar no vuelan. Veamos qué ocurre ahora:

Análisis de la teoría. En este caso, $Vuela(a) \in N_{\Theta''}^\delta$.

Análisis de la observación. Contingente: $O^\delta = \{\{Vuela(a)\}\}$.

Búsqueda de refutaciones. Como la única δ -cláusula de O^δ pertenece a $N_{\Theta''}^\delta$, la observación refuta la teoría. Por tanto, $\Theta'' \models \neg\phi$, es decir, desde la teoría podemos inferir (deductivamente, en sentido clásico) que a no vuela. Por tanto, con la nueva teoría no sólo no podemos inferir por defecto que a vuela, sino que además podemos deducir que no es el caso. El proceso abductivo que hemos proporcionado, como vemos, nos sirve para detectar esta situación.

Bibliografía

- [Ali97] Atocha Aliseda. *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. Dissertation Series. Institute for Logic, Language, and Computation, Holland, 1997.
- [Bus95] Samuel R. Buss. "On Herbrand's theorem". *Lecture Notes in Computer Science*, 960:195–209, 1995.
- [CP84] Philip T. Cox y Tomasz Pietrzykowski. "A complete, nonredundant algorithm for reversed skolemization". *Theoretical Computer Science*, 28:239–261, 1984.
- [CP93] Marta Cialdea Mayer y Fiora Pirri. "First order abduction via tableau and sequent calculi". *Bulletin of the IGPL*, 1:99–117, 1993.

- [Fla94] Peter Flach. *Simply Logical. Intelligent Reasoning by Example*. John Wiley, 1994.
- [KKT98] Antonis Kakas, Robert Kowalski y Francesca Toni. “The role of abduction in logic programming”. En *Handbook of logic in Artificial Intelligence and Logic Programming*, páginas 235–324. Oxford University Press, 1998.
- [NS04] Ángel Nepomuceno y Fernando Soler. “Abducción y razonamiento por defecto”. En Vicente et al. [VFC⁺04], páginas 385–387.
- [Rob65] John Alan Robinson. “A machine-oriented logic based on the resolution principle”. *Journal of the ACM*, 12:23–41, 1965.
- [Sol04a] Fernando Soler. “Cálculo de δ -resolución proposicional”. En Vicente et al. [VFC⁺04], páginas 392–395.
- [VFC⁺04] Agustín Vicente, Patricia de la Fuente, Cristina Corredor, Juan Barba y Alfredo Marcos, editores. *Actas del IV Congreso de la Sociedad de Lógica, Metodología, y Filosofía de la Ciencia en España*. Sociedad de Lógica, Metodología, y Filosofía de la Ciencia en España, 2004.