Plausibilidad y Razonamiento No Monotónico *

Guillermo Simari † Claudio Delrieux ‡

Universidad Nacional del Sur
Alem 1253, (8000) Bahía Blanca
ARGENTINA

e-mail:grs@arcriba.edu.ar

PALABRAS CLAVE: INTELIGENCIA ARTIFICIAL, REPRESENTACIÓN DE CONOCIMIENTO, RAZONAMIENTO NO MONOTÓNICO, PLAUSIBILIDAD, CREENCIAS.

*Este trabajo fue parcialmente financiado con fondos de la Secretaría de Ciencia y Técnica de la Universidad Nacional del Sur, en el ámbito del Grupo de Investigación en Inteligencia Artificial (GIIA).

†Departamento de Matemática

‡Departamento de Ingeniería Eléctrica

Abstract

La investigación en Representación de Conocimiento ha dado lugar a sistemas de razonamiento no monotónico, es decir, sistemas con la capacidad de utilizar premisas tentativas para razonar y eventualmente revisar sus conclusiones a la luz de nueva información. Por *premisas* se puede entender tanto reglas tentativas (que no tienen el *status* de implicaciones materiales) como hechos asumidos *por suposición*. Distinción que nunca ha estado lo suficientemente clara.

Nuestro propósito es presentar un sistema de razonamiento que integre de una manera precisa y clara ambos tipos de conocimiento tentativo. El manejo de reglas tentativas será efectuado bajo el sistema de Reiter [4]. La incorporación de información asumida se realiza por medio de un sistema que, esencialmente, es una generalización del sistema de *Razonamiento Plausible* propuesto por Rescher [7].

El resultado es un sistema capaz de aceptar o rechazar premisas plausibles provenientes de otros agentes en función de un conocimiento previamente adquirido y del conjunto de premisas ya aceptadas. Más aún, el sistema no acepta todas las consecuencias de dichas premisas, sino solo aquellas consecuencias justificadas por su conocimiento previo.

1 Introducción

El razonamiento no monotónico intenta construir un sistema con la capacidad de incluír reglas tentativas, y que tenga la característica de *modificar* sus conclusiones cuando nueva evidencia así lo determine. El punto principal en este trabajo es explorar la necesidad de una distinción en el tratamiento que se da en dichos sistemas a las premisas tentativas *por suposición*, y a las premisas *revisables* como por ejemplo las reglas tentativas.

En la literatura inglesa las primeras se denominan *assumptions*, y se reconocen sintácticamente por ser literales de base. Las segundas son las *default rules*, y sintácticamente toman la forma de una implicación material abierta (con variables libres sin cuantificar), o de una regla en el metalenguaje.

Por ejemplo, R. Reiter en un trabajo de 1988 describe el razonamiento no monotónico como aquél que incorpora premisas del tipo "Normalmente A es del caso" ([5], pp. 439). Sin embargo, existe una diferencia esencial entre lo que puede inferir un razonador si A es un literal de base y lo que puede inferir si A es una implicación abierta.

En particular, Reiter se refiere al segundo caso, tanto en el resto de su trabajo como en sus publicaciones de los últimos diez años. Esta es la rama de la representación de conocimiento que se conoce como *default reasoning*. Pero queda el interrogante ¿Qué sucede cuando A es un literal?

La importancia de aclarar esta cuestión para el razonamiento no monotónico es que en general los sistemas propuestos deben basar su razonamiento en premisas consistentes. Esto claramente es una suposición muy fuerte, dado que un agente razonador con sentido común puede tener premisas inconsistentes -tal vez por estar mal informado- pero no por ello sus capacidades de razonamiento se suspenden. Por el contrario, el comportamiento racional es ubicar y eliminar la causa de la inconsistencia.

La solución propuesta en este trabajo es considerar que dichas premisas provienen de *informantes*, es decir, de otros agentes -falibles- que informan a nuestro razonador de cosas que son del caso en el mundo. Nuestro razonador incorpora dicha información, calificándola de acuerdo a la *plausibilidad* [7] que asigna a cada informante. Un informante distinguido dentro de dicha estructura es el razonador mismo, el cuál tiene máxima plausibilidad.

Cuando ocurre una inconsistencia, el sistema aceptará solamente la información proveniente del informador más plausible (en caso de que ésto pueda determinarse). Si no tiene forma de comparar la plausibilidad de los agentes que informan las premisas que dieron lugar a la inconsistencia, entonces se descarta toda la información conflictiva. Es decir, se suspende el juicio con respecto a dichas premisas.

El funcionamiento en detalle de este sistema, junto con algunos ejemplos y discusiones, será presentado en este trabajo de acuerdo a la siguiente secuencia: en la sección 2 se presentará en forma resumida el razonamiento con suposiciones propuesto por Reiter. En la sección 3 se presentará la idea esencial del razonamiento plausible de Rescher [7] y la generalización de la misma que adoptaremos en este trabajo. El sistema completo, la unión de las capacidades de razonamiento no monotónico y razonamiento plausible, se presenta en la sección 4. La sección 5 incluye ejemplos y discusiones, y en la sección 6 se señalan las conclusiones generales y las propuestas para futuros sistemas.

2 La Lógica Default de Reiter

El objetivo de esta sección es proveer el contexto necesario para el resto del trabajo, por lo que su lectura puede ser omitida por quienes conozcan la teoría. Podemos pensar que se desea construir un modelo del razonamiento de un agente idealmente racional **a**, el cuál tiene conocimiento lógico (*i.e.*, deductivamente válido) $\mathcal{K}$, y además tiene conocimiento *revisable*, es decir, conocimiento que expresa tendencias generales pero que pueden tener excepciones. Estas excepciones no son necesariamente conocidas, por lo que se espera que el sistema retracte o revise las conclusiones obtenidas con dicho conocimiento cada vez que aparece una excepción. Esto significa que se espera un comportamiento no monotónico en el sistema.

El conocimiento de **a** está representado del siguiente modo. $\mathcal{K}$ está expresado en un lenguaje de primer orden $\mathcal{L}$. En $\mathcal{K}$ destacamos el conocimiento general o necesario $\mathcal{K}_{\mathcal{N}}$ y el conocimiento particular o contingente $\mathcal{K}_{\mathcal{C}}$ de modo que $\mathcal{K} = \mathcal{K}_{\mathcal{N}} \cup \mathcal{K}_{\mathcal{C}}$. Sintácticamente la diferencia entre $\mathcal{K}_{\mathcal{N}}$ y $\mathcal{K}_{\mathcal{C}}$ es que el primer conjunto de conocimiento se expresa en sentencias universalmente clausuradas, mientras que el segundo conjunto de conocimiento se expresa exclusivamente con sentencias de base (*i.e.*, sin variables).

La lógica default de Reiter representa el conocimiento tentativo por medio de esquemas de **reglas por suposición** o por defecto (*default rules*) [4]. La no monotonicidad se manifiesta por medio de dichos esquemas de regla, que pueden tener excepciones, sin que dichas excepciones estén explíctamente representadas. Esto significa que sus conclusiones pueden revisarse en caso de que exista nueva evidencia. De ese modo tendremos que la regla por defecto

$$\frac{A(\overline{X}) : \mathsf{M}B(\overline{X})}{C(\overline{X})}$$

es un esquema que representa el proceso de derivar una conclusión plausible (o consecuente) $C(\overline{X})$ dados el conocimiento cierto (o prerrequisito) de que $A(\overline{X})$ y la consistencia del conocimiento plausible (o justificación) de que $B(\overline{X})$ [1]. La representación sintáctica sufrió desde entonces algunas simplificaciones. M puede ser omitida en la notación sin ambiguedad, y prerrequisito, consecuente y justificación toman la forma de literales $a(\overline{X})$, etc. Por último, el vector de variables $\overline{X}$ tambien se omite.

DEFINICIÓN 2.1 (Reiter [4])
Una Teoría con Suposiciones T *es un par* $T = \langle \mathcal{K}, \Delta \rangle$ *donde* $\mathcal{K}$ *es un conjunto de sentencias cerradas en un lenguaje de primer orden y* Δ *es un conjunto de reglas por suposición.*

Si Δ es cerrado, es decir, cada regla en Δ es cerrada, entonces la teoría con suposiciones es cerrada. Intuitivamente, $\mathcal{K}$ es el conocimiento deductivo de la

[1] $A(\overline{X})$, $B(\overline{X})$ y $C(\overline{X})$ son fbf arbitrarias donde un vector de variables $\overline{X}$ ocurre libre.

teoría, brindando un contexto a la misma. Δ intenta producir un conjunto de *extensiones* consistentes a partir de $\mathcal{K}$. La idea de extensión intenta capturar la idea del conjunto (consistente) de creencias que se puede extraer a partir de una teoría dada. Como veremos, las extensiones pueden ser múltiples o pueden no existir. Cada una de dichas extensiones es una posible visión del mundo compatible con el conjunto original de creencias que a originalmente poseía.

Una extensión E de una teoría $T = \langle \mathcal{K}, \Delta \rangle$ cerrada es un conjunto de sentencias que satisface las siguientes propiedades:

1. $\mathcal{K} \subseteq E$,
2. $Th(E) = E$,
3. Si existe una suposición $\frac{a:b}{c}$ en Δ tal que $a \in E$, $\neg b \notin E$, entonces $c \in E$.

Esto permite caracterizar las extensiones con la siguiente definición:

DEFINICIÓN 2.2 (Reiter [4])
Dada una teoría $T = \langle \mathcal{K}, \Delta \rangle$ cerrada, para cualquier conjunto $\mathcal{A}$ de sentencias sea $\Gamma(\mathcal{A})$ el menor conjunto que satisface

1. $\mathcal{K} \in \Gamma(\mathcal{A})$,
2. $Th(\Gamma(\mathcal{A})) = \Gamma(\mathcal{A})$,
3. *Si existe una suposición $\frac{a:b}{c}$ en Δ tal que $a \in \Gamma(\mathcal{A})$, $\neg b \notin \Gamma(\mathcal{A})$, entonces $c \in \Gamma(\mathcal{A})$.*

Entonces un conjunto E de sentencias $E \subseteq \mathcal{L}$ es una extensión *para Δ si y sólo si $\Gamma(E) = E$, es decir, E es un punto fijo del operador Γ.*

EJEMPLO 2.1 *Sea una teoría $T = \langle \{\}, \{\frac{\{\}:a}{\neg b}, \frac{\{\}:b}{\neg a}\} \rangle$. Tenemos dos extensiones: $E^1 = \{a, \neg b\}$ y $E^2 = \{b, \neg a\}$*

EJEMPLO 2.2 *Una teoría sin extensiones es $T = \langle \{\}, \{\frac{\{\}:a}{\neg a}\} \rangle$*

Un resultado importante es que una teoría T tiene una extensión E inconsistente si y sólo si $\mathcal{K}$ es inconsistente, y E es la única extensión de la teoría. La lógica con reglas por suposición consigue extender la lógica clásica de modo de poder representar algunos de los patrones de inferencia no monotónica que ocurren en sistemas que incorporan negación por falla, suposición de mundo cerrado, etc. Es decir, se extiende la lógica de aquellos sistemas donde existe esencialmente una sola regla por suposición. Cuando existen dos o más reglas cuyos consecuentes son inconsistentes, entonces la lógica no es lo suficientemente fuerte como para elegir una extensión entre las posibles.

EJEMPLO **2.3** *Sea la teoría*
$T = \langle\{ave(opus), herido(opus)\}, \{\frac{ave(X) : vuela(X)}{vuela(X)}, \frac{herido(X) : \neg vuela(X)}{\neg vuela(X)}\}\rangle$.
Tenemos dos extensiones: $E^1 = \{ave(opus), herido(opus), vuela(opus)\}$ *y* $E^2 = \{ave(opus), herido(opus), \neg vuela(opus)\}$.

El objetivo de este trabajo es dotar al sistema de un mecanismo que le permita elegir entre distintas extensiones introduciendo una relación de plausibilidad que afecta a $\mathcal{K}_C$, es decir, al conocimiento contingente que a tiene del mundo. De esa manera a puede seleccionar información de modo tal que su base de conocimiento $\mathcal{K}$ sea siempre consistente (en caso de que $\mathcal{K}_\mathcal{N}$ lo sea).

3 Razonamiento Plausible

El nombre "Razonamiento Plausible" y algunas de sus motivaciones fueron inspiradas por los trabajos de N. Rescher [6, 7], los cuales son, en esencia, un intento de volver a dotar a las lógicas modales de una pragmática, de la cuál fueron esterilizadas por los enfoques excesivamente algebraicos. En Rescher, la plausibilidad de una fórmula es una medida de su grado de soporte asertivo, *i.e.*, de su valoración epistemológica [1].

En ésto es necesario ser preciso. La plausibilidad no se refiere ni a la forma del conocimiento, ni a su contenido. La plausibilidad es *externa* al conocimiento.

Esto permite distinguir claramente a la noción de plausibilidad de otra valoración epistemológica más conocida: la probabilidad. Claramente la probabilidad es intrínseca al conocimiento. Conocer la probabilidad de un evento es de alguna manera conocer cómo se relaciona con su clase de referencia. Por ejemplo, una amiga arroja un dado y nos dice que salió el 4. La probabilidad de dicho evento es $\frac{1}{6}$. Pero nosotros no tenemos razones para desconfiar de nuestra amiga, por lo que la plausibilidad del mismo evento es mucho mayor que $\frac{1}{6}$.

En este ejemplo se evidencia que muchas de las "reglas" del razonamiento probabilístico no se aplican a plausibilidades. Por ejemplo, para cualquier sentencia, si ésta tiene una probabilidad p, su negación debe tener probabilidad $(1-p)$. En el razonamiento plausible ésto no es así: podemos ser totalmente agnósticos con respecto a una sentencia (la conjetura de Fermat, por caso), es decir, juzgarla poco plausible y al mismo tiempo juzgar poco plausible a su negación. Esto muestra que la plausibilidad no es una *medida*, lo cuál es una propiedad ventajosa.

En este sentido, el objetivo es tener conocimiento calificado por su plausibilidad, y razonar con el mismo utilizando la lógica clásica junto con algún procedimiento ampliativo. Es decir, el conocimiento plausible se utiliza junto con el conocimiento lógico en forma deductiva, hasta tanto suceda algo que nos obligue a revisar lo hecho.

La propuesta de Rescher consiste en asignar a cada fórmula de un lenguaje proposicional un índice de plausibilidad en una familia $\{\frac{1}{n}, \frac{2}{n}, \cdots, \frac{n-1}{n}, 1\}$ para un n dado. La plausibilidad de cualquier fórmula del lenguaje puede encontrarse con un conjunto de desigualdades. Ante inconsistencias, el criterio de racionalidad es rechazar la premisa menos plausible que las origina. Los detalles pueden consultarse en las obras citadas, pero son innecesarios para este trabajo, ya que en este punto nuestro enfoque se aparta del enfoque de Rescher por varias razones.

En principio parece más razonable referirse a la plausibilidad de *la fuente de información* que a la plausibilidad de la información misma. Por dicha razón adoptamos la visión de asignar plausibilidades a las sentencias en función del *informante* que las provee.

Al mismo tiempo la idea de un orden lineal parece desafortunada. No siempre es posible comparar la plausibilidad de los distintos informantes y decidir en quién confiamos más. Podemos leer un diario sensacionalista informando un determinado hecho y escuchar por televisión la desmentida del funcionario involucrado. En este caso, no estamos en condiciones de creer más en un reporte que en el otro. Por dicha razón adoptamos un *orden parcial* para representar la plausibilidad de los informantes.

Por último, dado que la plausibilidad es externa al conocimiento, es imposible dar una pragmática adecuada a la plausibilidad de un conocimiento intensional o general. Rescher no discute este punto, permitiendo que se asigne plausibilidad, por ejemplo, a implicaciones. Como nuestro interés es trabajar con un lenguaje de primer orden, no está claro de ningún modo qué significa que una sentencia $\forall X.p(X) \Rightarrow q(X)$ tiene una determinada plausibilidad.

Los fundamentos filosóficos de esta distinción son un tanto sutiles y se remontan a los orígenes mismos de la lógica modal. Podemos, sin embargo, resumir la idea en la distinción entre la modalidad *de dicto* (del dicho) y la modalidad *de re* (de la cosa en sí). La plausibilidad es claramente una forma *de dicto* de relativizar la valoración epistemológica de una fórmula. Es decir, una sentencia plausible es una "verdad" *prima facie*. Por dicha razón sólo tiene sentido hablar de plausibilidades en casos particulares y concretos. La relativización *de re* de una sentencia abierta solamente puede pensarse como un conocimiento tentativo y sujeto a excepciones (reglas *prima facie*). Este tipo de conocimiento es exactamente para el cuál se desarrolló el sistema de razonamiento no monotónico expuesto en la sección anterior.

Por dicha razón, solamente vamos a permitir que nuestro sistema asigne plausibilidades a conocimiento particular, i.e., literales de base. Este conocimiento plausible, por otra parte, es exactamente al que Reiter se refiere cuando dice "Normalmente A es del caso" y A es un literal de base.

4 El Sistema

Volvamos ahora al proceso de razonamiento de nuestro agente a. El conocimiento particular $\mathcal{K}_{\mathcal{C}}$ de a, *i.e.*, lo que el agente ve por sí mismo en el mundo, debe ser necesariamente consistente con $\mathcal{K}_{\mathcal{N}}$. Pero además a tiene conocimiento adquirido a través de un conjunto de informantes $\mathcal{I} = \{i_1, i_2, \cdots, i_n\}$, donde cada i_k es una fuente de información. Cada i_k, entonces, está capacitado para informar a a de un conjunto de literales. Ahora bien, como vimos en la sección 2, la base para el razonamiento revisable de a debe ser consistente. El propósito de esta sección es aplicar los conceptos de plausibilidad para restaurar la consistencia de la misma cuando los conjuntos de información dan origen a una inconsistencia.

a tiene establecido en $\mathcal{I}$ un orden parcial al que nos referiremos como *relación de plausibilidad* $\mathcal{P}$. Esto significa que, dados dos informantes i_p y i_q, a puede dar preferencia a la información de i_p, o bien a la de i_q, o bien no dar preferencia a ninguno. Podemos, entonces, completar el conjunto $\mathcal{I}$ con a mismo, considerado como un "informante" $i_{\top}$ cuya plausibilidad es máxima, y con todo el grupo de informantes "desconocidos" considerándolos como un único informante $i_{\perp}$ al cuál se le asigna una plausibilidad menor que la de cualquier otro informante conocido. Podemos decir entonces que $\mathcal{I}$ es un conjunto reticulado bajo la relación $\mathcal{P}$.

Este orden parcial permite a a aceptar o rechazar información. En principio, cualquier información no conflictiva (es decir, consistente con el conocimiento ya aceptado) puede ser aceptada. Cuando existen conflictos, a puede rastrear los agentes que informaron el conocimiento conflictivo, y descartar la información conflictiva que provenga de los agentes menos confiables. Las definiciones a continuación nos permiten formalizar estas ideas.

DEFINICIÓN 4.1 *Una* estructura de informantes *es un par* $\langle \mathcal{I}, \mathcal{P} \rangle$, *donde* $\mathcal{I}$ *es un* conjunto de informantes $\{i_1, i_2, \cdots, i_n\}$ *y* $\mathcal{P}$ *es un orden parcial sobre* $\mathcal{I}$, *llamado* relación de plausibilidad *denotado con el símbolo* $\geq_{\mathcal{P}}$. $\mathcal{I}$ *contiene un elemento* $i_{\top}$ *tal que* $\forall j \in \mathcal{I}, i_{\top} \geq_{\mathcal{P}} j$, *y un elemento* $i_{\perp}$ *tal que* $\forall j \in \mathcal{I}, j \geq_{\mathcal{P}} i_{\top}$.

Cuando dos o más informantes comunican el mismo literal, es necesario solamente retener el reporte del informante más plausible. Como nuestra estructura de informantes es un reticulado, existe un menor conjunto de cotas superiores para el conjunto de informantes que reportan el literal repetido. a solamente necesita registrar dicho conjunto.

DEFINICIÓN 4.2 *Dada una estructura de informantes, el* conjunto de información $\Im$ *con el que cuenta* a *es definido como sigue. Para cada literal* l *informado, consideremos el subconjunto de los miembros de* $\mathcal{I}$ *que lo informan. A partir de este subconjunto se puede construir un conjunto* C_l *de cadenas maximales en el orden* $\mathcal{P}$. a *necesita registrar solamente el supremo de cada una de dichas cadenas. Es decir, para cada* l *se construye el conjunto* $\Im_l = \{lub\{C_l\} : C_l$ *es una cadena maximal de informantes de* $l\}$. *El conjunto de información* $\Im$ *es el la unión de todos los* $\Im_l$.

EJEMPLO 4.1 *Supongamos que en el conjunto de ítems de información encontramos el subconjunto* $\{\langle i_p, l\rangle, \langle i_q, l\rangle, \langle i_a, l\rangle, \langle i_b, l\rangle, \langle i_c, l\rangle\}$, *y que además en* $\mathcal{P}$ *tenemos* $\{i_p \geq_{\mathcal{P}} i_q, i_a \geq_{\mathcal{P}} i_b, i_a \geq_{\mathcal{P}} i_c\}$. *Entonces* $\Im_l = \{\langle i_p, l\rangle, \langle i_a, l\rangle\}$.

Dado un conjunto de información $\Im$, entonces, ¿cuál es el conjunto $\mathcal{IK}_{\mathcal{C}}$, subconjunto consistente del mismo, que a aceptará? El siguiente procedimiento efectivo describe dicho proceso.

DEFINICIÓN 4.3 *El conjunto* $\mathcal{IK}_{\mathcal{C}}$ de información aceptada *se construye del siguiente modo: dado un conjunto de información* $\Im$, *para cada par de ítems de información* $\langle l, i_p\rangle$ *y* $\langle m, i_q\rangle$ *pueden darse tres casos:*

1. *Si* $\{l, m\} \cup \mathcal{K} \not\vdash \perp$ *entonces aceptar ambos ítems.*
2. *Si* $\{l\} \cup \mathcal{K} \not\vdash \perp$,
   $\{m\} \cup \mathcal{K} \not\vdash \perp$, *pero*
   $\{l, m\} \cup \mathcal{K} \vdash \perp$, *y es posible decidir cuál informante es más plausible bajo* $\mathcal{P}$, *entonces aceptar el reporte del informante más plausible.*

3. *Si* $\{l\} \cup \mathcal{K} \not\vdash \bot$,
   $\{m\} \cup \mathcal{K} \not\vdash \bot$,
   $\{l, m\} \cup \mathcal{K} \vdash \bot$, *pero no es posible decidir cuál informante es más plausible bajo* $\mathcal{P}$, *entonces rechazar ambos reportes.*

Tal como fue presentado en secciones anteriores, el sistema de razonamiento por suposición de Reiter no produce ninguna conclusión en el caso de existir extensiones múltiples. Dado un conjunto de conocimiento ($\mathcal{K}_{\mathcal{N}} \cup \mathcal{IK}_{\mathcal{C}}$, Δ) el agente a puede ahora decidir entre dichas conclusiones, si es que las distintas extensiones se basaban en premisas de agentes cuya plausibilidad es comparable. Es decir, no solamente puede a reestaurar la consistencia de su base de conocimientos por la vía de la plausibilidad, sino que también puede elegir entre las distintas extensiones.

Para describir formalmente el funcionamiento del sistema necesitamos algunas definiciones previas.

DEFINICIÓN 4.4 *Dada una extensión E de una teoría $\langle \mathcal{K}, \Delta \rangle$, y una estructura de informantes $\mathcal{I}$ ordenada bajo $\mathcal{P}$, definimos su* cota de plausibilidad $E_{\mathcal{P}}$ *como el conjunto* $\{lub\{l_k, i_k\}\}$, *donde cada l_k es un miembro del subconjunto de $\mathcal{K}_{\mathcal{C}}$ utilizado para derivar E, y i_k es el informante del cuál a aceptó l_k*

Intuitivamente, la cota de plausibilidad de una extensión es el conjunto de "puntos débiles" por donde puede ser atacado bajo la relación de plausibilidad. Una relación de orden entre cotas de plausibilidad puede establecerse a partir de $\mathcal{P}$.

DEFINICIÓN 4.5 *Una extensión $E^1_{\mathcal{P}}$ es mayor que otra extensión $E^2_{\mathcal{P}}$ si todos los elementos de la cota de plausibilidad de $E^1_{\mathcal{P}}$ son a lo sumo tan plausibles como todos los elementos de la cota de plausibilidad de $E^2_{\mathcal{P}}$ y por lo menos un elemento de la cota de plausibilidad de $E^2_{\mathcal{P}}$ es estrictamente más plausible que todos los elementos de la cota de plausibilidad de $E^1_{\mathcal{P}}$ Denotaremos esta relación como $E^1 >_{\mathcal{P}} E^2$. De ese modo, una extensión E^1 es preferible a otra extensión E^2 si y sólo si $E^1 >_{\mathcal{P}} E^2$.*

5 Ejemplos

El ejemplo paradigmático donde se espera que el sistema propuesto produzca resultados interesantes es el conocido como "Diamante de Nixon" [9]. En el mismo se desea atribuir tentativamente dos propiedades conjuntamente inconsitentes a un individuo. Como ambas en conjunto no pueden asumirse, existen sistemas (llamados "crédulos" [2]) que generan dos extensiones a la teoría, en cada una de las cuáles es válida una de las propiedades, y sistemas "escépticos" que generan una única extensión, en la cuál ninguna propiedad vale.

Como solución general al problema, ninguna de ambas posturas ("crédula" o "escéptica") parece satisfactoria. La solución propuesta por este sistema es, como puede ya adivinarse, chequear la plausibilidad de los informantes que proveyeron la información.

EJEMPLO 5.1 [Diamante de Nixon(*Touretzky[9]*)]
El conocimiento general que a tiene acerca de algunas actitudes políticas de los miembros de los partidos políticos y de algunas minorías religiosas de EEUU se resume en un par de reglas tentativas:

Los republicanos no son pacifistas $\frac{r(X):\neg p(X)}{\neg p(X)}$

Los cuáqueros son pacifistas $\frac{q(X):p(X)}{p(X)}$.

Con respecto a políticos en particular a conoce por la prensa especializada que Nixon es republicano $\langle i_p, r(nixon)\rangle$, *pero también ha visto en algún reportaje que Ms. Nixon, la madre del ex-presidente de los EEUU, asegura que su hijo es un ferviente cuáquero* $\langle i_q, q(nixon)\rangle$. *¿Debe a suspender su juicio con respecto a las actitudes políticas de Nixon? Podemos formalizar esta situación del siguiente modo:*

$$\mathcal{K} = \{\}, \quad \Delta = \{\frac{r(X) : \neg p(X)}{\neg p(X)}, \frac{q(X) : p(X)}{p(X)}\}.$$

En estas condiciones es perfectamente posible para a aceptar los reportes $\langle i_p, r(nixon)\rangle$ *y* $\langle i_q, q(nixon)\rangle$. *Sin embargo, a la hora de decidirse sobre el supuesto* *pacifismo de Nixon,* a *generará una extensión* $E^1 = \{p(nixon)\}$ *y otra extensión* $E^2 = \{\neg p(nixon)\}$. *Si* a *tiene razones para confiar más en la prensa especializada que en la madre de Nixon, entonces* E^2 *será preferida* a E^1 *por la plausibilidad de su base, y entonces* a *concluirá* $\neg p(nixon)$.

Es necesario recalcar que aunque el ejemplo es trivial, la situación a la que llega **a** no es en nada trivial. Su sentido común le sigue dictando que los cuáqueros tienden a ser pacifistas y los republicanos tienden a no serlo. Al mismo tiempo tiene conocimiento sobre un individuo, del cuál conoce (consistentemente) que es cuáquero y republicano. Sin embargo, con este conocimiento sólo llega a la conclusión de que no es pacifista, en base a las posibles extensiones y sus plausibilidades, es decir, Nixon es una excepción a la regla de que los cuáqueros son pacifistas.

Ejemplos como el 5.1, pese a ser sencillos, son de gran importancia, ya que casos más complejos se edifican a partir de situaciones similares, y requieren soluciones similares. Esto sucede en el ejemplo de las "ambigüedades en cascada", donde la situación de ambigüedad del pacifismo de Nixon se propaga hacia otras de sus características.

EJEMPLO 5.2 [Ambigüedades en Cascada(*Horty* et. al. *[3]*)]
El conocimiento general que a *tiene acerca de algunas actitudes políticas ha aumentado con respecto al ejemplo 5.1:*

Los republicanos no son pacifistas	$\frac{r(X):\neg p(X)}{\neg p(X)}$
Los cuáqueros son pacifistas	$\frac{c(X):p(X)}{p(X)}$
Los republicanos son fanáticos del futból	$\frac{r(X):ff(X)}{ff(X)}$
Los fanáticos del futból son belicistas	$\frac{ff(X):b(X)}{b(X)}$
Los pacifistas no son belicistas	$\frac{p(X):\neg b(X)}{\neg b(X)}$
Nixon es republicano	$\langle i_p, r(nixon)\rangle$
Nixon es cuáquero	$\langle i_q, q(nixon)\rangle$

¿Qué puede a concluír con respecto al belicismo de Nixon? En particular tenemos las extensiones $E^1 = \{ff(nixon), b(nixon)\}$ y $E^2 = \{p(nixon), \neg b(nixon)\}$, las cuales, como puede verse, son inconsistentes entre sí. Si, como en el ejemplo anterior, a está en condiciones de asignar mayor plausibilidad a alguno de sus informantes, entonces la extensión en cuestión será la preferida.

En este ejemplo, algunos razonadores llegan solamente a la primera conclusión por considerar que la ambigüedad en $p(nixon)$ "bloquea" la segunda extensión. En este sistema ésto no sucede en ningún caso. Si a confía más en alguno de sus informantes, entonces su decisión con respecto a $p(nixon)$ se reduce al ejemplo 5.1, y en consecuencia su decisión con respecto a $b(nixon)$.

6 Conclusiones y Trabajo en Progreso

Se ha presentado un método que mejora la representación del conocimiento particular de un agente y su habilidad para razonar. La ventaja del mismo consiste esencialmente en restaurar la consistencia, esencial para el razonamiento, de un conjunto de conocimiento particular informado por otros agentes. Esto permite que el sistema de razonamiento llegue a conclusiones basadas en la plausibilidad de los informantes, donde antes un único nivel de credibilidad en la información particular impedía las mismas. Esto fue analizado en los ejemplos.

El sistema tiene posibles mejoras que están siendo investigadas actualmente. Una de ellas consiste en modificar la relación de plausibilidad, fundamentalmente cuando un informante provee información que después se comprueba falsa, o cuando un informante sistemáticamente acierta. Esto constituye un "aprendizaje" que el agente realiza en base a la experiencia.

Con respecto a implementaciones, la lógica de Reiter se caracteriza por la imposibilidad de calcular sus puntos fijos. Un sistema de razonamiento revisable computacionalmente tratable fue presentado en [8]. Actualmente se está trabajando en extender dicha implementación para permitir el manejo de plausibilidad.

Referencias

[1] Roderick M. Chisholm. *Theory of Knowledge.* Prentice Hall, Englewood Cliffs, New Jersey, 1977.

[2] John F. Horty, Richmond H. Thomason, y David S. Touretzky. A Clash of Intuitions: The Current State of Nonmonotonic Multiple Inheritance Systems. En *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, páginas 476–482, International Joint Conference on Artificial Intelligence, Morgan Kaufmann Publishers, Los Altos, CA, 1987.

[3] John F. Horty, Richmond H. Thomason, y David S. Touretzky. A Skeptical Theory of Inheritance in Nonmonotonic Semantic Networks. *Artificial Intelligence*, 43:311–348, 1990.

[4] Raymond Reiter. A Logic for Default Reasoning. *Artificial Intelligence*, 13(1,2):81–132, Apr 1980.

[5] Raynond Reiter. Nonmonotonic Reasoning. En Howard E. Shrobe y AAAI, editores, *Exploring Artificial Intelligence*, páginas 439–481, Morgan Kaufmann Publishers, San Mateo, CA, 1988.

[6] Nicholas Rescher. *Hypotetical Reasoning.* North Holland, Amsterdam, 1974.

[7] Nicholas Rescher. *Plausible Reasoning.* Van Gorcum, Dodrecht, Germany, 1976.

[8] Guillermo R. Simari y Ronald P. Loui. A Mathematical Treatment of Defeasible Reasoning and its Implementation. *Artificial Intelligence*, 53(2-3):125–158, 1992.

[9] David S. Touretzky. *The Mathematics of Inheritance Systems.* Morgan Kaufmann Publishers, Los Altos, CA, 1986.