

## MEDIDAS DE SIMILARIDAD PARA EL PROBLEMA DE LA CUALIFICACION

Javier Vilanova Arias

### I. Introducción

En los últimos años la inteligencia artificial se ha convertido en un ámbito privilegiado de aplicaciones para la lógica. Esta toma un papel determinante en el diseño de máquinas inteligentes que utilicen un sistema de representación del conocimiento declarativo, mediante sentencias en algún lenguaje (frente a una representación meramente procedural, mediante programas)<sup>1</sup>. A este respecto la tarea es doble. Por una parte, es necesario construir un lenguaje lo suficientemente expresivo, capaz de recoger aspectos del conocimiento como los relativos a su variación en el tiempo o a su naturaleza vaga. Por otra, se han de elaborar técnicas de tratamiento y utilización del conocimiento expresado en ese lenguaje en tareas de predicción, planificación, diagnosis, aprendizaje, etc.

En este contexto es donde surge el problema de la cualificación, que se presenta a la hora de estipular las precondiciones de una acción, en la mayoría de los casos excesivamente numerosas para ser procesadas eficazmente, por lo que se hace necesario seleccionar la información pertinente en el contexto de esa acción. Con éste se conecta el problema del marco o, como Shoham prefiere llamarlo, problema de la persistencia<sup>2</sup>, que se presenta en tareas de predicción a la hora de establecer que una serie de hechos permanecerán sin variación durante un período extenso de tiempo, hechos entre los cuales se encontrarán las cualificaciones contextuales de la acción.

Desde el punto de vista adoptado aquí la ejecución de una acción involucra dos fenómenos interconectados, a los que se refieren respectivamente el problema de la cualificación y el problema de la persistencia. El primero se expresa a través de un condicional en el que la parte antecedente incluye las cualificaciones de la acción (cuándo se debe ejecutar y qué condiciones

---

<sup>1</sup> Nilsson (1991, p. 32) expone las ventajas de la representación declarativa del conocimiento, más versátil que la procedural, que permite disociar el componente de la máquina que decide como usar el conocimiento del conocimiento mismo. Estas ventajas, entre otras, son el acceso público, fiabilidad, universalidad, facilidad de comunicación con el usuario, adecuación formal, etc., ya citadas por otros autores como Smolensky y McCarthy.

<sup>2</sup> Para Shoham (1988, pp. 17-18) el problema de la persistencia se presenta al efectuar predicciones extendidas en el tiempo, como otra cara de la moneda con respecto al problema de la cualificación en tareas predictivas.

han de darse para que ello sea posible) y la parte consiguiente la acción en cuestión. El segundo se expresa en un condicional en el que la parte antecedente está constituida por la acción que se ejecuta y la parte consiguiente por las consecuencias previsibles de esa acción sobre el estado del mundo. Desde esta perspectiva nuestra tarea involucra también la revisión de nuestra descripción del mundo, modificando aquellos enunciados afectados por la acción ejecutada y estableciendo cuáles permanecen idénticos, enunciados que a su vez pueden entrar dentro de la serie de precondiciones de acciones posteriores<sup>3</sup>.

Muchas de las búsquedas de solución de estos problemas provienen del campo de la lógica modal. Algo que no debe resultar sorprendente si tenemos en cuenta que ésta se desarrolló con la vista puesta en los condicionales, que recogen precisamente la forma lógica de la ejecución de una acción. Intuitivamente, la idea sobre la que descansan estas aproximaciones es la identificación del universo de mundos posibles con estados de cosas optativos con los que se puede enfrentar el sistema, para compararlos posteriormente en base a los aspectos de tales estados de cosas que resultan relevantes respecto a la instrucción que se deba ejecutar. De este modo se seleccionan las precondiciones necesarias (por ejemplo, cargar una pistola antes de dispararla) de aquellas meramente contextuales o muy evidentes (que el ambiente en que se realiza el disparo contenga oxígeno o que la pistola no sea de chocolate). Es aquí donde aparece la noción de similaridad, como una sofisticación del concepto clásico de accesibilidad intermundana, pero a la que se le pide ahora que proporcione comparaciones más finas que la dicotomía sí-no de la vieja relación de acceso, y que utilice criterios de medida relativos al contexto de operación o a aquellos aspectos del mundo que consideramos más esenciales (su «física»).

## II. Requisitos para una definición de similaridad

Llanamente, el concepto de similaridad puede ser entendido como la explicitación formal del «parecido» o «semejanza» que guardan dos situaciones cualesquiera<sup>4</sup>. Dicho de otro modo, si la relación de accesibilidad intermundana clásica sólo nos indicaba si un mundo podía ser o no «concebible» desde el nuestro<sup>5</sup>, la nueva relación de similaridad nos indicará la «proximidad» de cada uno de ellos, su cercanía al actual, o, por expresarlo en los términos de la relación de acceso, en qué medida tenemos que prescindir de las ideas que tenemos sobre nuestro mundo para poder concebir aquél.

<sup>3</sup> Para una discusión de estos problemas y otros relacionados véase el artículo de Martínez y Sagüillo citado en la bibliografía.

<sup>4</sup> Ruspini (1989b, p. 14) lo explica señalando que la función de similaridad indica en qué medida el conocimiento de la verdad de ciertos enunciados (aquellos relativos al problema que se está estudiando) en un mundo puede ser extrapolado a otros.

<sup>5</sup> Hughes & Cresswell 1968, pp. 77-78.

La cuestión que se nos plantea ahora es la selección de los criterios con que vamos a establecer tal parecido. Por ejemplo, las variaciones en los valores de verdad del conjunto de enunciados que componen nuestra representación del mundo, o bien sólo un subconjunto de ellos que consideremos pertinente para el problema que examina la máquina en ese momento, o bien a través de las relaciones de implicación que se dan entre conceptos en cada mundo, etc. Otro aspecto es el relativo a las características formales de la relación: si va a ser una relación de orden, de equivalencia...; si va a aportar valores numéricos o de otro tipo; si tomará como variables pares de mundos o de conjuntos de ellos, etc.

Para tomar tales decisiones hemos de atender al objetivo primario de la investigación: la construcción de una relación de similaridad entre descripciones de estados de cosas entendidos como mundos posibles, que permita dotar a un sistema de planificación, control, diagnosis, etc., de un criterio de elección de la información pertinente para la toma de decisión que ha de ser efectuada. Criterio que, a su vez, servirá para indicar la actuación pertinente del mundo cuando alguna circunstancia del mismo sea modificada por la ejecución de la instrucción o algún evento similar. Desde este punto de vista, se han distinguido una serie de requisitos básicos que, a ser posible, ha de cumplir la definición, y que se presentan a continuación.

**Requisito 1:** ha de ser una medida no meramente sintáctica (enunciados que se encuentran en cada descripción mundana), sino también semántica, en el sentido de que se ha de revalorizar aquella parte del conocimiento que consideramos más profunda o esencial para nuestro mundo (su «física»).

Esta cláusula puede ser enunciada de otro modo estableciendo que la relación de similaridad se ha de regir por criterios no meramente cuantitativos (calculando el número de enunciados que cambian de valor de verdad en el otro mundo) sino también cualitativos (es decir, haciendo hincapié en aquellos enunciados que recogen elementos esenciales de la estructura de nuestro mundo). Por ejemplo, nuestra medida de similaridad tendría que designar como más cercano al actual un mundo donde hablara mi tía sordomuda que un mundo donde hablara un chimpancé. A su vez, habría de indicarnos más próximo este último que uno donde el animal parlante fuera una rana, y éste que otro en el cual quien tuviera el don de la palabra fuera una silla de montar. En cierto modo, este es el requisito más básico, y que diferencia los conceptos de similaridad más elaborados, establecidos *a posteriori* sobre el universo de mundos posibles, de los conceptos apriorísticos primitivos.

**Requisito 2:** ha de ser contextualmente sesgado, discriminando la información pertinente para la situación concreta que está siendo examinada por el sistema en el momento.

Este requerimiento apunta directamente al problema de la cualificación para el cual se intenta construir la relación de similaridad. Como hemos

dicho, el sistema ha de tomar en consideración únicamente aquellos mundos que resultan relevantes para el problema en cuestión, y que corresponden a los enunciados verdaderamente relevantes de la parte antecedente de la instrucción. Por otra parte, esta condición puede no ser aplicada directamente en la relación de similaridad original, sino en alguna sofisticación posterior de la misma con valor auxiliar<sup>6</sup>.

**Requisito 3:** debe poder ser formulada en una lógica monomótona, preferiblemente lógica borrosa, que la haga operativa en contextos de sistemas de representación del conocimiento parciales y razonamiento aproximado.

La necesidad de una lógica monomótona para resolver el problema de la cualificación resulta evidente si atendemos a los términos en que se plantea éste. El problema de la cualificación exige que el sistema «salte» a la conclusión una vez verificadas las premisas que podríamos denominar «propias» (para el caso del disparo, que la pistola esté cargada y el seguro desconectado por ejemplo), ignorando, a no ser que haya información en contra, las condiciones puramente contextuales o muy evidentes (que no se encuentre metido en una piscina o dentro de un horno, que el caño de la pistola no esté obturado), pero le ha de ser posible volver sobre ellas (por ejemplo, si el disparo no se realiza aunque se hayan satisfecho las premisas propias, y el robot intenta diagnosticar las causas del mal funcionamiento de la pistola) ante información nueva, sacando nuevas conclusiones que pueden invalidar la inferencia anterior. Como se ve, al hacer esto el sistema está violando el principio de monotonía al que se atiene la lógica clásica:

$$\text{Si } A \vdash C, \text{ entonces } A \wedge B \vdash C^7.$$

La preferibilidad de una lógica borrosa se justifica por la gran aplicabilidad de la misma, ya contrastada, en el diseño de sistemas de inteligencia artificial que incorporan conocimiento sentencial (con especial éxito en lo que se refiere a sistemas expertos, programas que simulan el comportamiento racional de un «experto» humano). En efecto, la estructura de la lógica borrosa parece adaptarse perfectamente a los lenguajes informáticos, trabajando eficazmente, por otra parte, en contextos imprecisos o vagos.

**Requisito 4:** ha de aportar una medida de similaridad entre dos mundos posibles cualesquiera, y no un mero criterio para conocer el o los mundos más similares al actual.

Esta estipulación constituye el auténtico caballo de batalla de la meta planteada, ya que, de hecho, no se ha elaborado ningún procedimiento de

<sup>6</sup> Como hace por ejemplo Ruspini (1989a, p. 26) al introducir la similaridad marginal como una extensión de la definición general a partir de un concepto o enunciado.

<sup>7</sup> Como indica Shoham (1988, p. 78), el principio de monotonía se encuentra íntimamente vinculado al importante teorema de deducción que, entre otras cosas, permite el paso de un sistema de deducción natural a un sistema axiomático: Si  $(A \wedge B) \vdash C$  entonces  $A \vdash (B \rightarrow C)$ .

cálculo de la similaridad que aporte tales valores numéricos. Si, además, tenemos en cuenta la elección de la lógica borrosa efectuada en el requisito anterior, la cual trabaja con valores numéricos (en principio los incluidos en el intervalo real entre 0 y 1), resultaría conveniente también que nuestra medida de similaridad aportara valores numéricos. Igualmente resulta conveniente que la medida de similaridad sea ampliable para pares de conjuntos de mundos posibles, de tal modo que se pueda calcular el grado de similaridad entre los conjuntos de mundos que satisfacen a dos enunciados distintos, midiendo así su interdependencia.

**Requisito 5:** debe ser posible su utilización en implementaciones de sistemas informáticos operando en tiempo real y con unas limitaciones materiales de computación.

Este requisito se refiere a los condicionamientos pragmáticos impuestos por la capacidad limitada de procesamiento. El algoritmo para medir la similaridad que se presente ha de ser lo suficientemente simple como para ser ejecutado en un tiempo razonable por parte del sistema que lo incorpore. A este respecto habrá que buscar el debido equilibrio entre la firmeza de las inferencias realizadas y los costes de espacio y tiempo que conllevan.

### III. Enfoques generales en la concepción de la similaridad

Desde los presupuestos establecidos en el apartado anterior vamos a explorar a continuación una serie de enfoques en el tratamiento de la similaridad existentes en la literatura actual en lógica modal. Aunque la lista no es exhaustiva, ni mucho menos, creo que se corresponden bastante bien con los acercamientos que se pueden encontrar en la actualidad. Se ha prescindido, por otra parte, de otros desarrollos, interesantes por sí mismos, que se han dado desde otros ámbitos, como pueden ser la teoría de prototipos o el estudio de la sinonimia lingüística, que aun cuando guardan cierta relación con los enfoques presentados aquí no entran dentro de su perspectiva general<sup>8</sup>.

De este modo distinguiremos, eventualmente, cuatro enfoques generales que constituyen cuatro grupos de conceptos de similaridad<sup>9</sup>.

**Enfoque 1:** aquellas nociones de similaridad, meramente sintácticas, surgidas en los primeros análisis de los condicionales subjuntivos por Stalnaker y Lewis, y que son definidas a priori sobre el universo de mundos posibles.

La investigación llevada a cabo independientemente (e, incluso, polémica-

<sup>8</sup> Véase, por ejemplo, Tversky 1977, o Trillas & Valverde 1981.

<sup>9</sup> Advertamos que el objetivo de este apartado es el examen de tales enfoques desde el punto de vista de los requisitos antes apuntados, y no, por lo tanto, la descripción o el análisis más o menos profundo de los mismos (algo imposible dadas las limitaciones de espacio), para lo cual remitimos a la bibliografía incluida al final del trabajo.

camente) por Stalnarker y por Lewis estaba especialmente dirigida a proporcionar una versión formal del test de Ramsey para la validez de un contrafáctico. Este fue reformulado por Stalnarker del siguiente modo<sup>10</sup>:

— Primero, revisa mínimamente el mundo actual de tal modo que el antecedente se haga verdadero.

— Segundo, evalúa la verdad del consecuente en ese mundo mínimamente revisado. Si es verdadero, el contrafáctico también lo es.

La primera parte del test exige la existencia de una función de selección que asigne a cada mundo  $W$  y a cada sentencia  $A$  un mundo  $f(W, A)$  que sea la revisión mínima de  $W$  para hacer a  $A$  verdadero, es decir, una función que asigna a cada par  $(W, A)$  el mundo más similar a  $W$  donde la sentencia  $A$  es cierta. Posteriormente, Lewis (1972) arremetió contra esa versión a su juicio demasiado simple, a la que achacaba la asunción no justificada de una serie de presupuestos sobre la revisión mínima (unicidad del mundo más próximo, del límite en la relación, del ordenamiento parcial de mundos respecto a la similaridad al mundo actual, del centramiento o accesibilidad de cada mundo por sí mismo) presentando una versión del test de Ramsey que no dependa de tales presuposiciones:

—  $(A \square \rightarrow C)$  es verdadero en  $W$  si y sólo si algún AC-mundo es más similar a  $W$  que cualquier  $AC\bar{C}$ -mundo, si hay AC-mundos accesibles<sup>11</sup>.

Como se ve, la línea de Lewis se encamina, más que por un análisis de la propia noción de similaridad que justifique o en su caso deseche las asunciones mencionadas, por una salvación *a posteriori* de las fallas del concepto a través de una sofisticación de la evaluación del condicional que la utiliza. De ahí que el principal problema de este enfoque sea la vacuidad de la relación de similaridad, al no proporcionar ninguna fórmula o criterio para calcular su valor, por lo que incumple tanto los requisitos 1 (*a posteriori*), como 4. Por otra parte, utilizan una lógica modal monótona (requisito 3), y aunque cumple en cierto modo el requisito 1 (la elección se hace con respecto a una sentencia  $A$ ), lo hace vacuamente al no proporcionar un procedimiento de medición.

**Enfoque 2:** aquellas definiciones que, en estudios posteriores sobre contrafácticos en la línea de Lewis (Ginsberg, Smith), proporcionan una similaridad *a posteriori*, cualitativa y contextualmente sesgada, especialmente dirigida a la resolución del problema de la cualificación.

La idea básica de Ginsberg y Smith es la de sustituir la enumeración exhaustiva de las precondiciones de la acción por una serie de restricciones generales que en un dominio dado puedan interferir sobre las instrucciones posibles. De este modo, la acción se interpreta como una implicación contra-

<sup>10</sup> Sustituyendo el stock de creencias del sujeto cognoscente que aparecía en la versión original de Ramsey por el mundo posible que se adecua a ese stock (Stalnarker 1968).

<sup>11</sup> Lewis 1972, p. 60. Advértase la peculiar notación de Lewis: «A-mundo» denomina un mundo donde la sentencia  $A$  es verdadera, «AC-mundo» un mundo donde las sentencias  $A$  y  $C$  son verdaderas, y « $A\bar{C}$ -mundo» un mundo donde la negación de  $A$  es verdadera.

fáctica cuyo antecedente recoge sólo lo que aquí hemos llamado premisas «propias» de la acción y su consecuente los resultado de la misma. El sistema ejecutará la acción, entonces, cuando exista un mundo máximamente similar al actual en el que se satisface el antecedente y el consiguiente es verdadero, siempre que se satisfagan las restricciones generales del dominio.

Ginsberg define la similaridad primero sólo sintácticamente: cada mundo es visto como un subconjunto del conjunto evidencial  $S$  (conjunto de sentencias que constituyen nuestra descripción del mundo actual<sup>12</sup>), se seleccionan como conjuntos máximamente similares al actual con respecto a la premisa del contrafáctico  $p$  ( $W(p, S)$ ) aquellos subconjuntos de  $S$  que no implican la negación de  $p$  y que no pueden ser extendidos sin hacerla falsa:

$$W(p, S) = \{T \subseteq S / \neg(T \vDash \neg p) \wedge (T \subset U \subseteq S \Rightarrow U \vdash \neg p)\}$$

Posteriormente se introduce conocimiento contextual<sup>13</sup> relevante mediante dos procedimientos. Por una parte, introduciendo un predicado  $B$  de «mal mundo» que discrimina aquellos mundos que incumplen alguna restricción básica del dominio. Por otra, usando una relación de orden parcial  $<$  en los subconjuntos de  $S$  que extiende la inclusión conjuntística aumentando nuestra habilidad para comparar mundos. Los conjuntos máximamente similares al actual  $S$  con respecto a  $p$  son ahora los subconjuntos de  $S$  que: (1) no implican  $\neg p$ ; (2) no incumplen el predicado  $B$ , y (3) si son anteriores con respecto a  $<$  a otro subconjunto de  $S$ , entonces éste último incumple una de las dos condiciones anteriores:

$$W(p, S) = \{T \subseteq S / (T \vDash \neg p) \wedge (\neg B(T)) \wedge (T < U \subseteq S \Rightarrow (U \vDash \neg p) \vee B(U))\}$$

A la luz de esta definición resulta obvio que este enfoque cumple los requisitos 1 y 2. Por otra parte, Ginsberg y Smith utilizan una lógica monótona, presentando el primero una posible traducción de su teoría en clave de una lógica multivaluada (cuyos valores de verdad serían verdadero, falso, e indeterminado o verdadero por omisión) muy sugerente para una posible conversión a la lógica borrosa. Ginsberg mismo presenta aplicaciones de su teoría a la simulación de aspectos de la planificación a través de regresión de metas (entendida como una búsqueda de respuesta a la pregunta: ¿qué premisa implicaría contrafácticamente la satisfacción de las metas?) y la diagnosis automática (entendiendo la diagnosis como una consecuencia del fallo observado), por lo que parece adecuarse eficientemente al requisito 5.

<sup>12</sup> Como veremos, esto va a tener consecuencias importantes, pues provoca el hecho de que los mundos posibles sean «sólo» modelos parciales del mundo actual.

<sup>13</sup> Apresurémonos a decir que lo que Ginsberg entiende por conocimiento contextual recoge aspectos tanto del requisito 1 como del 2, pues puede incluir tanto la información más relevante para el problema que atañe al sistema como la más importante desde el punto de vista de su física. De hecho, desde el momento en que Ginsberg da una medida de similaridad relativa al mundo actual pero también a la premisa del contrafáctico está cumpliendo el requisito 2.

El problema con la definición de Ginsberg y Smith es que no proporciona una medida cuantitativa, sino únicamente un procedimiento para obtener los mundos máximamente similares al actual. Además, el hecho de que los mundos posibles surjan como «debilitamientos» del mundo actual hace de ellos sólo modelos parciales del mismo, lo que impide otras medidas de similaridad que no sean las respectivas al mundo actual, haciendo incompatibles planteamientos más generales, como el de Ruspini.

**Enfoque 3:** una noción de similaridad surgida en la aproximación combinatorial a la modalidad (Amstrong, Mormann) dada a posteriori a partir de relaciones lógicas pero también empíricas entre universales.

La teoría combinatorial de la modalidad se caracteriza por interpretar los mundos posibles como conjuntos de pares ordenados que combinan universales (propiedades, atributos) y particulares (individuos), para posteriormente interpretar la estructura conceptual del mundo en base a la posibilidad o imposibilidad de las distintas combinaciones entre ellos<sup>14</sup>.

En Mormann (1992) esa estructura conceptual viene dada por un grupo laticia formado sobre el conjunto de J-conceptos ordenado por las operaciones « $\vee$ » y « $\wedge$ ». Un J-concepto es un conjunto de universales que inhieren sólo en un conjunto de individuos que sólo poseen esos universales como propiedades (dentro, por supuesto, del mundo combinatorio J en cuestión), algo así como los predicados naturales del mundo, las esencias en sentido escolástico. A partir de ahí se define para cada mundo su conjunto de pares críticos, pares de universales que no pueden inherir en el mismo individuo sin contradicción (por ejemplo, chimpancé y hablante). De este modo, un mundo  $J'$  es más similar al mundo J que otro mundo  $J''$  ( $J' <_J J''$ , siendo una relación de orden parcial) si tiene en común con J todo los conjuntos de pares críticos que  $J''$  tiene, pero no viceversa.

El planteamiento de Mormann posee la ventaja de que los mundos posibles son tan «completos» como el actual, permitiendo, así, calcular la similaridad con respecto a cualquier mundo y no sólo el actual, como es el caso en Ginsberg y Smith. Además, permite comparaciones cualitativas muy finas al establecer una jerarquía entre las propiedades del mundo actual (sus pares críticos) que deben ser salvaguardadas por los mundos más similares<sup>15</sup>. Sus limitaciones surgen del hecho de que sólo funciona con sistemas modales monótonos (vs. requisito 3) no más fuertes que S4, por lo que su campo de aplicaciones es muy limitado<sup>16</sup>. No cumple tampoco los requisitos 2 (dependiente del contexto), ni 4 (no aporta valores numéricos).

<sup>14</sup> Armstrong 1989, Bigelow 1988.

<sup>15</sup> El concepto de Moormann recogería perfectamente las diferencias de similaridad entre los mundos donde hablaran una persona sordomuda, un chimpancé, una rana... del ejemplo puesto al comentario del requisito 1.

<sup>16</sup> Además, el hecho de que el conjunto de individuos no sea idéntico para todos los mundos impide la introducción de la regla de Barcan.

**Enfoque 4:** una medida de similaridad cuantitativa surgida en el intento de Ruspini de construir, desde la lógica de mundos posibles, una semántica para la lógica borrosa.

El mismo Zadeh (1977) entendió la lógica borrosa como un proyecto de construcción de una «teoría de la posibilidad» que recogiera, frente a la teoría de la información de Shannon y Weaver, no sólo las dimensiones cuantitativas sino también cualitativas (su significado) de la información<sup>17</sup>. Basándose en ese enfoque Ruspini idea la construcción de una semántica para la lógica borrosa basada en mundos posibles, entendidos como el conjunto de descripciones alternativas del mundo compatibles con la evidencia disponible.

Para construir tal semántica parte de una función que asigna para cada concepto (o enunciado) y cada mundo un número real entre 0 y 1 que representa la medida en que ese mundo satisface el concepto (no olvidemos que se trabaja con lógica borrosa). La relación de similaridad tendría las propiedades formales de la relación de similaridad borrosa de Zadeh<sup>18</sup>, proporcionando un valor entre 0 y 1 para cada par de mundos posibles. Esta medida de similaridad es utilizada para cuantificar la extensión de la inclusión y la intersección entre subconjuntos de mundos posibles, interpretando así el grado de implicación de un concepto  $p$  por otro  $q$  (similaridad entre el ínfimo de los  $q$ -mundos y el supremo de los  $p$ -mundos), o su consistencia (similaridad entre los supremos de cada mundo), de un modo análogo a como la inclusión y la identidad entre conjuntos sirven para interpretar la implicación y equivalencia lógica clásicas respectivamente. Así, el grado de implicación de  $p$  por  $q$ , que expresa la cantidad mínima de extensión o estimamiento requerido para alcanzar un  $p$ -mundo desde un  $q$ -mundo, viene dado por la similaridad entre el ínfimo de los mundos para los que  $q$  es cierta y el supremo de los mundos para los que  $p$  es cierta.

Estas nociones son utilizadas por Ruspini para definir los operadores de necesidad y posibilidad, a partir de los cuales elabora un sistema formal capaz de recoger resultados tan importantes como el teorema del *modus ponens* generalizado de Zadeh, y toda la arquitectura formal de la lógica borrosa.

La limitación de la teoría de Ruspini, sobre todo, es que aun fundamentando todo su sistema en la relación de similaridad no aporta ningún algoritmo o fórmula para calcular ésta, limitándose a indicar que han de ser

<sup>17</sup> «... the proper framework for information analysis is possibilistic rather than probabilistic in nature, thus implying that what is needed for such an analysis is not probability theory but an analogous —and yet different— which might be called the theory of possibility». Zadeh 1978, p. 194.

<sup>18</sup> En Zadeh (1970) se presentan las relaciones de similaridad como extensiones borrosas de las relaciones de equivalencia clásicas, con propiedades reflexiva, simétrica y maxi-mini-transitiva (o, como en el caso de la utilizada por Ruspini \*-transitiva). Las relaciones de similaridad borrosa de Zadeh resultan excepcionalmente versátiles y poderosas a la hora de establecer clasificaciones, permitiendo generar árboles de partición y clases de similaridad en su conjunto base, a partir de los que Zadeh denomina conjuntos de nivel- $\beta$ , definidos:

$$R_{\beta} = \{ (x, y) / \mu R(x, Y) \geq \beta \}$$

generalizadas a partir de enunciados simples sobre aspectos limitados del comportamiento del sistema. De este modo, si bien da cuenta de los requisitos 2 (a través del concepto de similaridad marginal<sup>19</sup>), 3, 4, y, muy posiblemente sea susceptible de ser aplicada eficientemente en sistemas informáticos (requisito 5), paradójicamente está más cerca de las nociones apriorísticas de Lewis y Stalnarker que de las relaciones cualitativas de los otros enfoques.

### III. Conclusiones

Los distintos enfoques pueden ser comparados a través del cuadro siguiente, que los examina según los cinco requisitos estipulados.

	req. 1	req. 2	req. 3	req. 4	req. 5
Enfoque 1	No	No	No	No	No
Enfoque 2	Sí	Sí	Sí	No	Sí
Enfoque 3	Sí	No	No	No	?
Enfoque 4	?	Sí	Sí	Sí	?

A la luz del cuadro, el concepto de similaridad buscado surgiría naturalmente de una fusión de los enfoques 2 y 4, «llenando» la relación de similaridad borrosa de Ruspini con la fórmula de Ginsberg y Smith. Por supuesto, las diferencias formales y conceptuales entre ambos enfoques invalida toda opción de tal «fusión natural». La estrategia más plausible, dadas estas limitaciones, es intentar una reformulación de alguno de los enfoques introduciendo los aspectos positivos del otro, pero sin quebrantar la consistencia final ni la pureza teórica de la definición.

Posiblemente el mayor escollo para llevar a cabo una reformulación tal es que mientras Ginsberg trabaja con una lógica multivaluada nomonótona que acepta como valores veritativos «verdad», «falsedad» e «indeterminado», Ruspini utiliza una lógica borrosa cuyo conjunto de valores es el intervalo real entre 0 y 1. Por otra parte, no hay que olvidar que en Ginsberg, al contrario que Ruspini, la relación de similaridad es siempre relativa al mundo actual y a una proposición o conjunto de proposiciones; además, es de naturaleza comparativa, es decir, nos proporciona el o los mundos más similares al actual, sin darnos el valor de esa similaridad. Por ello, su utilización en sistemas que incorporen lógica borrosa resulta imposible en su formulación actual.

<sup>19</sup> Ruspini 1989, pp. 31-32, define la similaridad marginal ( $S_x$ ) como una «simetrización» del preorden inducido en el conjunto potencia de mundos posibles por el grado de implicación  $I(I)$ . Así la similaridad marginal entre los conceptos  $p$  y  $q$  va a ser el mínimo de sus respectivos grados de implicación:  $S_x(p/q) = \min ( I(p/q), I(q/p) )$ .

A la vista de estas limitaciones, la opción más plausible es la búsqueda de una formulación de la similaridad en lógica borrosa (con las propiedades formales de la relación de similaridad borrosa de Zadeh que utiliza Ruspini) siguiendo la estrategia utilizada por Ginsberg y Smith. Es decir, observando las variaciones en los valores de verdad en cada mundo, introduciendo conocimiento contextual que discrimine mundos «no interesantes» mediante «predicados-B», e intensificando la capacidad de comparar mundos mediante una relación de orden «<».

No se cierran, sin embargo, otras vías de investigación. Entre ellas, el intento de reconstruir la relación de similaridad de la aproximación combinatorial a la posibilidad (enfoque 3), o la propuesta de Ruspini de interpretar la similaridad a partir de técnicas de clasificación o «clustering»<sup>20</sup>. Por otra parte, puede ser fructífero un análisis de la lógica de la ignorancia cronológica de Shoham, muy cercana al trabajo de Ginsberg y Smith aunque no utilice el concepto de similaridad, y que se enfrenta exitosamente al problema de la cualificación<sup>21</sup>.

## Bibliografía\*

- Ginsberg, M. (1986): «Counterfactuals». *Artificial Intelligence*, vol. 30, pp. 35-79.
- Ginsberg, M. & Smith (1988a): «Reasoning about action I: A possible worlds approach». *Artificial Intelligence*, vol. 35, pp. 165-195.
- (1988b), «Reasoning about action II: The cualification problem». *Artificial Intelligence*, vol. 35, pp. 311-342.
- Hughes, C. E. & Cresswell, M. J. (1968): *An introduction to modal logic*, London, Methuen and Co ltd.
- Lewis (1972): «Counterfactual and comparative possibility». *Journal of Philosophical Logic*, 2, 418.
- Martínez, C. (1991): «Causalidad en I. A.: lógica de la ignorancia cronológica vs. similaridad». *Encuentro de Lógica y Filosofía de la Ciencia*, Madrid.
- Martínez, C. & Sagüillo, J. M.: «From philosophical logic to artificial intelligence: the problem of change».

<sup>20</sup> Ruspini 1989b, pp. 35-36.

<sup>21</sup> Shoham utiliza una lógica modal monótona del tiempo y la acción, dirigida a simular tareas de predicción. Si bien su concepto de modelo preferido puede ser visto como una versión particular del mundo maximalmente similar de Ginsberg, el concepto de modelo cronológicamente máximamente ignorante se aparta de manera más definida del enfoque 2, al incorporar una representación del tiempo como base formal. Además, el modelo cronológicamente máximamente similar ya no es un mundo posible, sino una «interpretación kripkeana» que incorpora un conjunto de mundos posibles y una función de interpretación.

\* Zadeh 1970 y Zadeh 1978 se citan a través del libro recopilatorio: Zadeh, L. A. (1987): *Fuzzy sets and applications: selected papers by L. A. Zadeh*, John Wiley and sons cop.

Lewis y Stalnarker se citan a través del libro recopilatorio: Harper, W. (ed.) (1981): *Ifs: conditionals, belief, decision, chance and time*, Reidel Cop.

- Mormann, T. (1992): «Structural accesibility and similarity of possible worlds». *Journal of Philosophical Logic*, n° 21, pp. 149-172.
- Nilsson, N. J. (1991): «Logic and artificial intelligence». *Artificial Intelligence* 47, pp. 31-56.
- Ruspini, E. H. (1989a): «The semantics of approximate reasoning». *Artificial Intelligence Center*, SRI International, Menlo Park, C. A.
- (1989b), «On the semantics of fuzzy logic». *Artificial Intelligence Center*, Computer and informational sciences division, Menlo Park, C.A.
- Shoham, Y. (1988): *Reasoning about change*, Massachusetts, MIT Press.
- (1989), «Efficient reasoning about rich temporal domains». En el libro recopilatorio R. I. Thomason (ed.): *Philosophical logic and artificial intelligence*, Klawer Academia Publishers.
- Stalnarker (1968): «A theory of conditionals». *American Philosophical Quarterly Monograph Series*, n° 2, Oxford, Blacwell.
- (1975): «Indicative conditionals». *Philosophia*, 5.
- Tversky, A. (1977): «Features of similarity». *Psychological Review*, vol. 84, pp. 433-460.
- Valverde, L. (1985): «On the stucture on F-indistinguishability operators». *Fuzzy sets and systems*, n° 17, pp. 313-328.
- Zadeh, L. A. (1970): «Similarity relations and fuzzy orderings».
- (1978): «Fuzzy sets as a basis for a theory of possibility». *Fuzzy Sets and Systems*, vol. 1, pp. 3-28.

Javier VILANOVA ARIAS  
Universidad de Santiago de Compostela