

CONJUNTOS BORROSOS Y LOGICA DE ZADEH

Griegorz Malinowski

Abstract

As it is generally acknowledged the possibility to apply the apparatus of the traditional formal logic to the analysis of everyday reasoning is very limited. However, recent development of logical studies has considerably extended them. The aim of the article is to account for one of the most interesting but simultaneously most controversial conceptions inspired by logical many-valuedness: the fuzzy set theory and logics based on it.

1. Teoría de conjuntos y lógica de conceptos imprecisos

Zadeh (1965) define un conjunto borroso A de un dominio dado U como un objeto abstracto definido por una función característica generaliza U_A con valores en el conjunto real $[0, 1]$:

$$U_A : U \rightarrow [0, 1]$$

Los valores de U_A se interpretan como grados de pertenencia de elementos de U sobre el conjunto borroso A . Los valores extremos de esta función, «0» y «1», denotan respectivamente: no perteneciente a A y entera pertenencia a A . Si se limita el alcance de U_A a $\{0, 1\}$, resulta en una función característica «ordinaria», y en este sentido todo conjunto «clásico» es un caso especial de un conjunto borroso. Asimismo, es posible detectar la afinidad entre conjuntos borrosos y clases inexactas (Körner, 1966) o conjuntos «n-valuados» (e.g. Urquhart, 1986).

Los conjuntos borrosos han de ser un instrumento para modelar predicados imprecisos de los lenguajes naturales. Zadeh sugiere, por ejemplo, que la propiedad de «ser mucho mayor que 1» definida sobre el conjunto de los reales positivos R^+ puede ser asignada a un conjunto borroso W mediante una función característica arbitraria $R_w : R^+ \rightarrow [0, 1]$ que satisfaga las siguientes condiciones: $R_w(0) = 0$, $R_w(1) = 0$, $R_w(5) = 0.01$, $R_w(100) = 0.95$, $R_w(500) = 1$, etc. Resulta sencillo concluir, en lo que respecta al ejemplo anterior, que sólo los valores $R_w(0)$, $R_w(1)$ y la monotonía de R_w son incuestionables. La selección de los otros valores es arbitraria, no obstante debe estar relacionada con el problema de constituir el contexto del predicado que se esté modelando W .

En la familia $F(U)$ de sub(conjuntos) borrosos de un dominio dado U Zadeh introduce la relación de inclusión poniendo:

$A \subseteq B$ si y sólo si $U_A(x) \leq U_B(x)$ para cualquier $x \in U$, y las contrapartidas de las operaciones de complementación ($-$), unión (\cup) y la intersección (\cap) a través de la siguiente especificación de las funciones características:

$$\begin{aligned} U_{-A}(x) &= 1 - U_A(x) \\ U_{A \cup B}(x) &= \max \{U_A(x), U_B(x)\} \\ U_{A \cap B}(x) &= \min \{U_A(x), U_B(x)\} \end{aligned}$$

Bellman y Giertz (1973) demostraron que $U_{A \cup B}(x)$ y $U_{A \cap B}(x)$ son las únicas funciones continuas no decrecientes que garantizan la compatibilidad de la construcción con el álgebra estándar de conjuntos y el hecho de que $(F(U), \cup, \cap, -)$ es un retículo De Morgan¹ si y sólo si U_{-A} se define como arriba. A pesar de la naturalidad de la propuesta, varios estudios admiten como (más) útiles álgebras de conjuntos borrosos definidas de forma alternativa. Un ejemplo de esto es el álgebra considerada por Zadeh (1976) que incluye las operaciones \cup y \cap definidas mediante los esquemas:

$$\begin{aligned} U_{A \cup B}(x) &= U_A(x) + U_B(x) - U_A(x) \cdot U_B(x) \\ U_{A \cap B}(x) &= U_A(x) \cdot U_B(x) \end{aligned}$$

que son semejantes a las reglas para calcular las medidas de las uniones e intersecciones de los conjuntos de Borel sobre hechos independientes en la teoría de la probabilidad. Estas operaciones satisfacen las condiciones necesarias que se reducen a las clásicas en el caso binario. Sin embargo, ninguna función es idempotente, i.e. $U_A \cup A(x)$ y $U_A \cap A(x)$ coinciden con U_A pero consideradas conjuntamente no son distributivas. Desde este momento, cuando hablemos de un álgebra de conjuntos borrosos nos referiremos a la primera de las construcciones de Zadeh.

Los valores de las funciones características generalizadas pueden ser identificadas con los valores lógicos de proposiciones obtenidas a partir de los predicados que sirven como base para determinar conjuntos borrosos (concretos). Y, consiguientemente, con un álgebra de conjuntos borrosos podemos asociar una lógica multivaluada incontable. La forma más transparente de presentar la relación mutua entre la lógica base y el álgebra de conjuntos borrosos consiste en definir la inclusión y las operaciones algebraicas por medio del predicado conjuntístico « \in » y las constantes lógicas:

$$\begin{aligned} A \subseteq B &= \text{df } \forall x (x \in A \rightarrow x \in B) \\ -A &= \{x : \neg (x \in A)\} \\ A \cup B &= \{x : x \in A \vee x \in B\} \\ A \cap B &= \{x : x \in A \ \& \ x \in B\} \end{aligned}$$

Para el álgebra de Zadeh la elección de una lógica base está en gran medida predeterminada: esta lógica debe estar basada en una matriz de X_1 - elementos, en la cual la negación se expresa mediante la función $1 - p$, la disyunción y la conjunción, respec-

¹ i.e. un retículo distributivo con « \rightarrow » que satisface las condiciones: $--x = x$, $-(x \cup y) = -x \cap y$ y $-(x \cap y) = -x \cup -y$.

tivamente por : $\max \{p, q\}$ y $\min \{p, q\}$, y el cuantificador universal como el ínfimo. La función de implicación no está unívocamente determinada. Sin embargo, evidentemente, debe cumplir la restricción siguiente:

(.) si $p \rightarrow q = 1$, entonces $p \leq q$.

Implicaciones especialmente interesantes para las que cumple (.) son las Lukasiewicz y Gödel (definida sobre $[0, 1]^2$). Puede resultar de interés señalar que de algún modo ambas reciben cierto apoyo desde la teoría de la probabilidad; la primera en virtud de la probabilidad subjetiva, la segunda de la condicional (véase Gaines (1976b)). La implicación de Lukasiewicz y, consiguientemente, su lógica de X_1 -valores, han sido las más aplicadas. Sin embargo, la creencia generalizada entre los investigadores dedicados al estudio de la vaguedad, tanto los teóricos como prácticos, es que sólo una aplicación concreta del álgebra de conjuntos borrosos puede decidir acerca de la forma final de la lógica base (Gaines, 1976a).

En las aplicaciones iniciales de la teoría de conjuntos borrosos se dedicó mucha atención al análisis del razonamiento que utiliza predicados imprecisos, conectivas proporcionales estándar y cuantificadores. De especial interés resultaron ser las paradojas (antinomias) cuyo análisis resultó adecuado y consolidó la razón de ser de la concepción de vaguedad. Esta explicación dió lugar a entender el término «lógica borrosa» como una cierta clase de lógica multi-valuada con un número infinito no enumerable de valores, con la lógica de Lukasiewicz en el *background*. Un caso típico de nodelar un predicado impreciso en esta última lógica es un intento de analizar la paradoja clásica de *un hombre calvo* de Goguen (1969); intuitivamente, podríamos aceptar las dos proposiciones siguientes:

- (z1) *Un hombre con 20.000 cabellos en su cabeza no es calvo.*
- (z2) *Un hombre que tiene un cabello menos que alguien que no es calvo tampoco es calvo.*

Si aplicamos el *modus ponens* 20.000 veces, obtendremos, por (z1) y (z2), la conclusión de que un hombre sin cabello tampoco es calvo. Naturalmente la paradoja surge de (z2) y más específicamente de la inexactitud del predicado «calvo» o equivalentemente «no-calvo». La paradoja desaparece cuando el valor lógico de la proposición «Un hombre con n cabellos no es calvo» se identifica con el grado de pertenencia de un hombre con n cabellos a un conjunto borroso «no-calvo», ya que (z2) tendría un valor lógico menor que 1, digamos $1 - \epsilon$, donde $\epsilon > 0$. Si, por ejemplo, en lógica base utilizamos la implicación de Lukasiewicz, entonces obtendremos una proposición de valor lógico que corresponderá a $1 - 20.000 \epsilon$, es decir, prácticamente falso. (Dejamos al lector la tarea de determinar el valor de ϵ que garantiza la falsedad).

El desarrollo de la teoría de conjuntos borrosos ha sobrepasado todas las expectativas. Para su uso se han adoptado todos los conceptos importantes de teoría de conjuntos, topología, álgebra y cálculo de probabilidad. Es preciso subrayar que fueron las aplicaciones en Ciencias de la Computación y Control Automático las que dieron ímpetu a un desarrollo más extendido de la teoría: han confirmado la utilidad

² Lukasiewicz: $p \rightarrow q = \min(1, 1 - p + q)$. Gödel: $p \rightarrow q = 1$ si $p \leq q$ y $p \rightarrow q = q$ en otro caso.

de la metodología no convencional desarrollada mediante la teoría de conjuntos borrosos (Gottwald, 1981).

2. Lógica borrosa

La lógica borrosa de Zadeh (1975) es un método para modelizar (formalización) razonamientos imprecisos, más específicamente predicados inexactos y términos veritativo-funcionales. Su construcción recoge la creencia de que pensar en términos de conjuntos borrosos es un rasgo típico de la percepción humana. Los ejemplos de razonamiento cuyo análisis pudiera ser posible en virtud de la lógica borrosa son «inferencias» tales como:

Reagan es muy sano
Los hombres sanos viven una larga vida

Reagan vivirá una larga vida.

A Mario le gustan las mujeres altas y embrujadoras
Eva no es muy alta pero si muy embrujadora

A Mario probablemente le guste Eva.

La lógica borrosa pretende formular reglas de inferencia aproximada. Para este propósito intenta formalizar usos lingüísticos coloquiales de ciertas «cercas» aplicadas a conceptos imprecisos tales como «muy», «más o menos», «demasiado», etc. La lógica de Zadeh es una construcción semántica de dos niveles que permite la borrosidad de los predicados, sus cercas y sus valores lógicos. Sus elementos centrales son:

- (1) Un conjunto TV denumerable de valores lógico lingüísticos generados por medio de su elemento «verdadero» con la ayuda de la cerca «muy» y las conectivas lógicas (proposicionales).
- (2) Cercas de predicados y valores lógicos, siendo «muy» uno especial (compárese 1).
- (3) El procedimiento de aproximación lingüística que compensa la carencia de clausura del lenguaje (objeto) y la clausura del conjunto TV sobre las conectivas lógicas.

La lógica borrosa se forma sobre la base de una lógica ordinaria X_1 – valuada con las conectivas \neg , \rightarrow , \vee , \wedge , \equiv y los valores en $[0, 1]$. Su principio constitutivo consiste en identificar los predicados con subconjuntos borrosos de un universo dado y los valores lógicos con subconjuntos borrosos del conjunto de valores de la lógica básica, i.e., $[0, 1]$. La que más frecuentemente se discute es el sistema «fundamental» FL obtenido a partir del LX_1 de Lukasiewicz (véase Bellman, Zadeh, 1977). Nuestras consideraciones ulteriores se confinarán a algunos aspectos de esta construcción.

Se asume que el conjunto TV de valores lógicos lingüísticos de FL es el conjunto de la forma:

$$TV = \{verdadero, falso, no verdadero, no falso, muy verdadero, más o menos verdadero, más bien verdadero, no muy verdadero, no muy falso, \dots\}$$

donde «verdadero» es un conjunto borroso designado de $[0, 1]$, «muy» una cerca fijada, y los restantes elementos se definen a partir de «verdadero», «muy» y de las operaciones determinadas por las conectivas de Lukasiewicz. Obviamente, los «nombres» de todos los valores lógicos son etiquetas convencionales y sus significados propios se derivan de funciones características.

Las cercas son operaciones de un argumento que aplican conjuntos borrosos en conjuntos borrosos. La más básica (una primitiva) es una cerca que refleja el adverbio «muy» y que se denota mediante g. Zadeh supone

$$(g) \quad U_{gA}(x) = (U_A(x))^2$$

y sugiere que otras cercas deberían ser definidas como superposiciones entre g y las conectivas de la lógica básica. El operador «derivado» más importante d se describe como sigue

$$(d) \quad U_{dA}(x) = (U_A(x))^{0.5}$$

Zadeh defiende que «más o menos» es una contraparte lingüística de d. En la literatura relevante g y d se denominan *cercas standard* y la expresibilidad por medio de su uso se ha convertido en un criterio de definibilidad de cerca. Conversamente, la búsqueda directa de otras cercas lingüísticas razonables se basa en el análisis de enunciados compuestos que contienen g. Así, si p es el enunciado:

«El jarrón es hermoso» (*jarrón ε objeto hermoso*),

entonces $dp = p \& \neg gp$ i.e., «El jarrón es hermoso pero no muy hermoso» puede ser condensado en «El jarrón es ligeramente hermoso»³. Aceptando que $\&$ y \neg son conectivas de Lukasiewicz obtenemos

$$U_{\text{ligeramente } A}(x) = \min \{U_A(x), 1 - (U_A(x))^2\}.$$

El procedimiento de aproximación lingüística surge de la carencia de clausura de TV bajo las conectivas lógicas. Más específicamente, si $U_{\text{verdadero}}$ es una función característica del valor denominado «verdadero», entonces, por supuesto, todos los restantes elementos de ese conjunto quedan de ese modo asimismo especificados. Así, por ejemplo, los valores iniciales de TV son definidos mediante las siguientes funciones:

³ La argumentación que apoya la caracterización de «bastante» que aquí se presenta es originaria de Zadeh. Evidentemente, la elección del término lingüístico concreto para este modificador puede ser discutible.

$$\begin{aligned}
U_{\text{falso}}(x) &= U_{\text{verdadero}}(1-x) \\
U_{\text{no verdadero}}(x) &= 1 - U_{\text{verdadero}}(x) \\
U_{\text{muy verdadero}}(x) &= (U_{\text{verdadero}}(x))^2 \\
U_{\text{más o menos verdadero}}(x) &= (U_{\text{verdadero}}(x))^{0.5}
\end{aligned}$$

(x toma valores en el conjunto de valores de la lógica básica, $x \in [0, 1]$). Consecuentemente, todos los valores lógico lingüísticos dependen de la introducción (subjética) del significado de «verdadero». Zadeh denomina *localización* a este rasgo y valores *locales* a los elementos de TV. De este modo obtenido, como resultado de la localización, los sistemas de LF se denominan *lógicas locales*. Para estas lógicas existe un modo común de definir las conectivas lógicas; consiste en identificarlas con las operaciones de álgebra de (subconjuntos) borrosos de $[0, 1]$ y así, en un sentido, con las conectivas de la lógica base. Empleando el «generalizado» \in los significados de \sim , \Rightarrow , \vee y \wedge de LF pueden ser abordados mediante las equivalencias siguientes:

$$\begin{aligned}
x \in [\sim A] &= \neg(x \in [A]) \\
x \in [A \Rightarrow B] &= (x \in [A] \rightarrow x \in [B]) \\
x \in [A \vee B] &= (x \in [A] \vee x \in [B]) \\
x \in [A \wedge B] &= (x \in [A] \wedge x \in [B]),
\end{aligned}$$

donde \neg , \rightarrow , \vee , \wedge , \equiv son conectivas de Lukasiewicz y $[Z]$ es un conjunto borroso representando, o pudiendo representar, (sobre el lado izquierdo) un valor lógico lingüístico relevante. Estos postulados, sin embargo, no garantizan la clausura del conjunto TV° para prácticamente ninguna de las lógicas locales LF° -si, por ejemplo, α y β son enunciados que tienen valores «más bien verdadero» y «no muy falso y no muy verdadero», respectivamente, entonces el valor del enunciado $\alpha \ \& \ \beta$ generalmente no es un valor lingüístico de LF° . Una *aproximación lingüística* es un procedimiento heurístico que asigna valores lógico-lingüísticos a cualquier proposición. Por razones obvias resulta imposible especificar ni tan siquiera un principio general de ese procedimiento; lo que puede decirse meramente es que en una aplicación concreta de una lógica borrosa consiste en la búsqueda para un enunciado del valor más cercano de TV° . Asumiendo que LA es un predicado de aproximación se puede dar una caracterización ecuacional de las conectivas LF como sigue:

$$\begin{aligned}
[\sim A] &= LA \quad x: \neg(x \in [A]) \\
[A \Rightarrow B] &= LA \quad x: (x \in [A] \rightarrow x \in [B]) \\
[A \vee B] &= LA \quad x: (x \in [A] \vee x \in [B]) \\
[A \wedge B] &= LA \quad x: (x \in [A] \wedge x \in [B])
\end{aligned}$$

De igual modo el concepto de aproximación se aplica a los predicados y a las cercas.

Como ya se ha mencionado, la lógica borrosa pretende formular las reglas de razonamiento (o inferencia) aproximado. Una regla tanto básica como estándar de esa clase es la regla *composicional de inferencia*

U₁ es F
U₁ y U₂ son G

U₁ y U₂ son F
U₂ y U₃ son G

U₂ es LA { F*G }

U₁ y U₃ son LA { F*G }

donde U₁ y U₂ son objetos, F y G predicados (propiedades o relaciones) * una operación de composición relacional; en el caso inicial * se define como sigue:

$$U_{F*G}(U_2) = \sup_x (U_F(x) \& U_B(x, U_2)),$$

y LA { F * G } es una aproximación lingüística a la relación borrosa (monaria) F*G. La siguiente es una ejemplificación del primer esquema:

a es un número pequeño
a y b son aproximadamente iguales

b es un número más o menos pequeño

En reglas más complicadas, las premisas son enunciados compuestos de varios grados de complejidad e.g. si u₁ is F, then u₂ es G, y pueden ser cuantificados mediante *cuantificadores borrosos* i.e., expresiones como la mayoría, muchos, varios, pocos, etc. En la formulación de otras reglas de inferencia borrosa se emplean operaciones especiales sobre conjuntos borrosos -aquí un buen ejemplo es la *Regla de Modus Ponens Composicional*:

u₁ es F
si u₁ es G, entonces u₂ es H

u₂ es LA { F * (G' + H) }

en donde G' + H se denomina la suma ligada de cilindricación de G' (el complemento de G) y H, (véase Bellman, Zadeh, 1977).

Lo que constituye la base de la aplicación práctica de las reglas de inferencia borrosas es la asignación de subconjuntos borrosos de un universo dado a predicados borrosos; en el caso de relaciones (no monarias), subconjuntos borrosos de los productos cartesianos relevantes de sus dominios. Este procedimiento, a menudo denominado como *restricción borrosa*, obviamente excede a la lógica. Sin embargo, dentro del ámbito de una lógica es posible e incluso recomendable formular principios generales de restricción que resulten de la peculiaridad de los aparatos formales.

La concepción «borrosa» de Zadeh ha encontrado su lugar entre los métodos aceptados en Inteligencia Artificial. Mantiene esta posición en virtud de aplicaciones efectivas como, por ejemplo, en diagnosis médica (véase Turner, 1984).

Bibliografía

- Bellman, R.E.; Giertz, M., «On the analytic formalism of the theory of fuzzy sets», *Information Sciences*, 5, 1973, pp. 149-156.
- Bellman, R.E.; Zadeh, L.A., «Local and fuzzy logics», w: Dunn, Epstein (eds.), 1977, pp. 105-165.
- Giles, R., «A non-classical logic for physics», *Studia Logica*, 33, 1974, pp. 397-416.
- Goguen, J.A., «The logic of inexact concepts», *Synthese*, 19, 1969, pp. 325-373.
- Gottwald, S., «Fuzzy-mengen und ihre Anwendungen. Ein Überblick», *Elektronische Informationsverarbeitung und Kybernetik*, 17, 1981, pp. 207-235.
- Körner, S., *Experience and theory*, Routledge and Kegan, Paul, London, 1966.
- Gaines, B.R., «Foundations of fuzzy reasoning», *International Journal of Man-Machine Studies*, 8, 1976a, pp. 623-668.
- Gaines, B.R., «General fuzzy logics», *Proceedings of the 3rd European Meeting on Cybernetics and Systems Research*, Viena, 1976b.
- Turner, R., *Logics for Artificial Intelligence*, Ellis Horwood Ltd., 1984.
- Urquhart, A., «An interpretation of many-valued logic», in *Handbook of philosophical logic*, vol. III, (eds. D. Gabbay, F. Guenther), D. Reidel, Dordrecht, Holland, 1986, pp. 71-116.
- Zadeh, L.A., «Fuzzy sets», *Information and Control*, 8, 1965, pp. 338-353.
- Zadeh, L.A., «Fuzzy logic and approximate reasoning», *Synthese*, 30, 1975, pp. 407-428.
- Zadeh, L.A., «A fuzzy-algorithmic approach to the definition of complex or imprecise concepts», *International Journal of Man-Machine Studies*, 8, 1976, pp. 249-291.

Griegorz MALINOWSKI
Universidad de Lódz