Resolución de ecuaciones, unificación, automatización del conocimiento y matemática educativa

Luis Carlos Cachafeiro

1. Introducción

LOS cambios producidos en nuestra sociedad a partir de las investigaciones en la Teoría del Conocimiento e Inteligencia Artificial tienen una amplia repercusión en la vida diaria de las personas y, a consecuencia de ello, en el propio currículum educativo, siendo las Matemáticas una de las materias más condicionadas a adaptarse a las diferentes necesidades y medios que dichos cambios producen.

Muchos de los resultados matemáticos clásicos, algunos de los cuales aparecen explícitamente en los actuales curricula de Secundaria y Bachillerato, tienen repercusiones destacadas en los avances en Inteligencia Artificial. Por ello esta disciplina no puede considerarse ajena a las programaciones actuales de las asignaturas de Matemáticas. En este artículo, observaremos la estrecha relación existente entre la resolución de ecuaciones y la unificación, operación esta que surge de las investigaciones realizadas en el campo de la prueba automática y que es utilizada directa o indirectamente en buena parte de las investigaciones en Inteligencia Artificial.

La finalidad última de la Inteligencia Artificial (IA) es la de producir programas y máquinas que realicen procesos auténticamente inteligentes. Las primeras investigaciones en IA surgen en los años 50 y sufrieron una aparente parada tras comprobar que esos primeros experimentos (e.g. programas que imitaban una conversación, iniciales probadores de teoremas, máquinas capaces de aprender si bien de una forma muy primitiva) distaban mucho de ser inteligentes. Esas investigaciones ignoraban muchos de los problemas reales de cualquier actividad inteligente y que suelen ser muchísimo más complejos que los procesos simples de deducción formal. De esta forma quedó probado que, por una parte, los humanos desconocíamos el funcionamiento de nuestra propia inteligencia y que por otra parte aunque se conociera ese funcionamiento, podría ser muy difícil de imitar mediante máquinas tanto por la enorme cantidad de información que se debe emplear como por hallarse organizada de forma muy compleja. Los enormes esfuerzos dedicados desde entonces a estas investigaciones, ya han producido avances notorios en muchos campos del conocimiento como en el proceso de imágenes, prueba automática, procesamiento del lenguaje natural, traducción automática entre otros (Graubard 1993).

En una aplicación inteligente puede suponerse la existencia de una base de conocimientos que contiene los datos disponibles (concretos o flexibles) y

tarbiya, nº 15. 1997. Págs. 79-88

Tabla 1

ALGUNOS TIPOS DE DATOS EN IA

Tipos de datos	Ejemplos
TÉRMINOS Expresiones básicas de la información	*f(x,g(a,y))* *Persona(N(Pepe),Edad(x),Altura(173))*
SUSTITUCIONES Adjudican valores a las variables	$z \leftarrow g(a)$ $x \leftarrow$ *32 años*
CLÁUSULAS Relacionan literales y permiten expresar datos y reglas	X es el hijo de Y si Y es el padre de X *Hijo (X,Y)* $\Leftarrow$ *Padre (Y,X)*
EXPRESIONES EN GRAMÁTICAS DE UNIFICACIÓN Construcciones jerarquizadas e interdependientes de ciertas estructuras sobre otras.	x_0 [categoría: Frase contenido: (1)] x_1 [categoría: Nombre contenido: (2)] x_2 [categoría: Verbo contenido: (1) [número: (2) sujeto: (3) modo declarativo]]

de un sistema de inferencia que es el que permite obtener nuevos resultados a partir de los previos. Podemos suponer que una parte de los datos expresan nuestra experiencia previa mientras que otra parte corresponde a la descripción de una situación actual que se desea transformar de una forma inteligente o idónea. También se precisa un método de aprendizaje que permitirá incorporar como nuevos conocimientos los resultados de anteriores cálculos o experimentaciones. Uno de los mayores problemas es disponer de estructuras de datos apropiadas para conservar, mantener y utilizar apropiadamente la información. En la tabla 1 pueden observarse algunos de los datos empleados en IA. Las expresiones en gramáticas de unificación se emplean especialmente en los campos de procesamiento del lenguaje natural y traducción automática (Knight 1989).

Los sistemas de inferencia permiten, a partir de

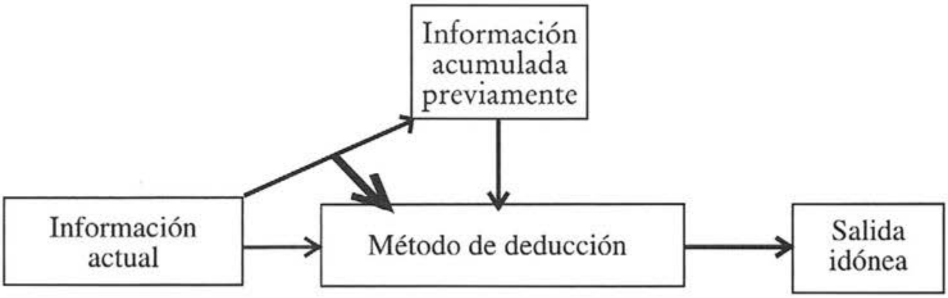

Figura 1
ACTUACIÓN DE UN MECANISMO INTELIGENTE

una información acerca del estado actual de un sistema dado, proporcionar una salida adaptada a ese estado. Esta salida dependerá tanto de la información disponible como del método de inferencia empleado. Para proporcionar respuestas idóneas, es primordial poder obtener nuevas informaciones a partir de los elementos ya conocidos. En Computación, una operación esencial para ello es la unificación.

2. Unificación de Términos

Esta operación permite contrastar dos informaciones flexibles (e.g. que incluyen variables) para obtener nuevos datos y adjudicar valores a las variables. Formalmente, si s y t son dos expresiones, la sustitución σ es unificador de s y de t si $\sigma(s) = \sigma(t)$.

Podemos clarificar el significado de esta operación y su cálculo con el siguiente ejemplo simplificado:

Ejemplo 1

Supongamos que a dos criptógrafos se les suministra un determinado texto para descifrar. Cada uno de ellos, independientemente, consigue conocer una parte del texto, por lo que colocan variables subindicadas para señalar palabras no descifradas. Concretamente el primero escribe:

Cada	x_2	*tiene su*	*propia*	*cadena*	*privada de*	x_9	*y de*
x_{12}	*de*	x_{14}	*que son*	*inaccesibles*	*para el otro*	x_{21}.	*Cada*
x_{23}	*desconectado*	*parece*	*tener una*	x_{28}	x_{29}	*separada.*	

El segundo obtiene el siguiente texto:

x_1	x_2	*tiene su*	*propia*	x_6	*privada de*	*recuerdos*	*y de*
experiencias	*de*	*aprendizaje*	*que son*	*inaccesibles*	*para el otro*	x_2.	*Cada*
x_2	x_{24}	*parece*	*tener una*	*«mente*	*propia»*	*separada.*	

De esta forma, al contrastar ambos textos surgen nuevos datos que no eran expresamente conocidos por ninguno de los dos criptógrafos. Los valores que son inmediatamente adjudicados son los siguientes:

$x_1 \leftarrow$ *Cada*	$x_{12} \leftarrow$ *experiencias*	$x_{24} \leftarrow$ *desconectado*	
$x_6 \leftarrow$ *cadena*	$x_{14} \leftarrow$ *aprendizaje*	$x_{28} \leftarrow$ *«mente*	
$x_9 \leftarrow$ *recuerdos*	$x_{21} \leftarrow x_2$	$x_{23} \leftarrow x_2$	$x_{29} \leftarrow$ *propia»*

quedando libre x_2. Por lo tanto cualquier asociación que se realice a mayores para x_2^{+} proporciona un texto «correcto» en el sentido de que es posible que se trate del original. Pero, por otra parte las anteriores asociaciones son las mínimas posibles y, ciertamente, las únicas «seguras». Así, haciendo $x_2 \leftarrow$ *individuo* se proporciona una adjudicación de valores correcta, dado que coinciden los dos textos, pero innecesaria ya que es información extra que se introduce y que no se encuentra realmente presente[1].

Ejemplo 2

La unificación de $f(x, x, a)$ con $f(y, b, y)$ resulta en fallo por el conflicto de los símbolos a y b.

La unificación de $f(x, g(x))$ con $f(y, y)$ resulta también en fallo pero en este caso porque la variable y no puede reemplazarse simultáneamente por x y por $g(x)$. En este último caso, si bien x y $g(x)$ no pueden igualarse, la sustitución $x \leftarrow g(x)$ es admitida en PROLOG como unificador porque permite expresar soluciones recursivas[2].

1 La adjudicación original de la que procede el texto es $x_2 \leftarrow$ *hemisferio* y este texto se encuentra en e (Johnson Laird 90).

2 El término resultante de la unificación es un término $f(g(g(...g(x),g(g(...g(x)))))$ con un número infinito (numerable) de símbolos g en cada argumento.

La unificación permite detectar si una información actual puede contrastarse favorablemente con alguno de los datos que corresponden a experiencias previas. Es el prototipo de operación de contraste porque permite el intercambio de datos entre las dos fuentes de información a contrastar. La unificación es utilizada expresamente en la demostración automática de teoremas y en la programación declarativa, e indirectamente en todos los sistemas de IA que hacen uso de la programación lógica. Existen otras formas de contraste como el *matching* y *pattern-matching* que pueden verse como formas particulares de unificación. Una sustitución s es un matching de s y de t si σ sólo actúa sobre s, σ verifica $\sigma(s)= t =\sigma(t)$ y por lo tanto σ también es unificador de s y t.

Una forma «natural» de contraste puede verse en el reconocimiento auditivo de palabras. En este reconocimiento la audición aparece como el resultado de un proceso: las células del oido interno hacen llegar al cerebro una secuencia de impulsos eléctricos que reproducen las características del sonido recibido. En el cerebro se produce un contraste de esta información, relativamente pobre (i.e. poco estructurada), con los modelos previos de sílabas o palabras, modelos que son necesariamente flexibles (pues en otro caso sólo se reconecería la expresión o habla de una persona concreta). Como consecuencia de este contraste se produce la adjudicación de valores (e.g. sílabas, palabras o frases cortas) que proporciona la compresión sensorial del sonido recibido. Dada la velocidad con que se realiza este proceso, generalmente esta última fase no llega a ser detectada, haciéndonos creer que ni siquiera existe y que la percepción es totalmente automática (Johnson-Laird, 1990).

Aunque la unificación fue definida y empleada explícitamente por el lógico J. A. Robinson (1964) en el campo de la prueba automática de teoremas, sorprendentemente no fue hasta primeros de los años 80 cuando se relacionó explícitamente esta operación con la transformación de ecuaciones. Martelli y Montanari (1982) buscando una forma de acelerar la unificación dieron una serie de reglas de transformación que transformaban el problema inicial hasta obtener el resultado, dado como: *Fallo, Forma resuelta* o *Forma factorizada*. En la forma factorizada o triangular las incógnitas no se encuentran totalmente despejadas lo que permite reducir en gran medida el coste de la unificación.

Se pueden observar muchas semejanzas entre la forma de realizar la unificación de Martelli y Montanari y el método de Gauss de resolución de sistemas de ecuaciones lineales. En ambos casos se parte de un sistema de ecuaciones que debe transformarse mediante ciertas reglas, incluyendo las de Borrado y Fallo y finaliza bien cuando se obtiene una forma resuelta (en cada ecuación aparece una variable despejada y su valor correspondiente) o una forma triangular (en cada ecuación aparece una variable despejada y una expresión que puede tener variables despejadas en otras ecuaciones).

En la tabla 2 se muestran una serie de ejemplos comparativos entre el método de Gauss de resolución de ecuaciones lineales y el método de Martelli y Montanari de unificación de términos.

3. Unificación Ecuacional

Como se mostró previamente, para que los términos s y t sean unificados por σ se precisa la igualdad de todos los símbolos de $\sigma(s)$ y de $\sigma(t)$. Esta forma «dura» de igualdad se conoce como igualdad sintáctica y exige que s y t se encuentren dados de una misma forma o notación.

Sin embargo nuestro concepto de igualdad no es generalmente tan estricto. Así, un mismo significado puede expresarse mediante dos frases completamente diferentes (e.g. «*Pedro es mayor que Juan*» o también «*Juan es menor que Pedro*»). Esta forma de igualdad en la que unas expresiones pueden ser

Tabla 2

EJEMPLOS COMPARATIVOS DE UNIFICACIÓN Y RESOLUCIÓN DE ECUACIONES

	Transformación de sistemas	Forma resuelta	Forma triangular
Ejemplo Martelli y Montanari	$f(x,a)=f(b,y)\Rightarrow\begin{cases}x=b\\a=y\end{cases}$	$\begin{cases}x=a\\y=f(z)\end{cases}$	$\begin{cases}x=f(y,a)\\y=g(z,b)\end{cases}$
Ejemplo Gaus	$\begin{cases}x+y=5\\x-y=3\end{cases}\Rightarrow\begin{cases}x+y=5\\2y=2\end{cases}$	$\begin{cases}x=4\\y=1\end{cases}$	$\begin{cases}x+y=5\\y=1\end{cases}$ ó $\begin{cases}x=5-y\\y=1\end{cases}$

iguales a otras sin que sean necesariamente coincidentes, se conoce como igualdad ecuacional o igualdad semántica y se expresa como $=_E$ o simplemente como =. En el lenguaje matemático se utiliza la igualdad ecuacional en todos los campos de la matemática inclusive en la matemática escolar. Así, cuando se trabaja con las expresiones 2/3 y 4/6 se utiliza la igualdad ecuacional siendo ambas representaciones diferentes del número [2/3]. En la aritmética, todo número tiene una representación más simple llamada forma normal (en el caso anterior ésta es 2/3). Pero éste no es el caso del álgebra, donde $x+3$ no es más simple que $3 + x$.

Sean s y t dos expresiones. Si una sustitución σ iguala semánticamente estas expresiones [i.e. $\sigma(s) =_E \sigma(t)$] se dirá que σ es un unificador ecuacional de s y t y que éstos términos son unificados ecuacionalmente por σ. La unificación ecuacional de los términos s y t en la teoría E, puede interpretarse como la resolución de la ecuación $s = t$ en esta teoría, ya que se trata de obtener los valores de las variables que permiten construir una igualdad $\sigma(s) =_E \sigma(t)$.

Ejemplo 3

Se supone que para cualquier valor de u y de v, se verifica la igualdad $f(u, u, v) = f(v, u, v)$.

En estas condiciones, la unificación ecuacional de los términos $f(x, x, a)$ y $f(y, b, y)$ no resulta en fallo, puesto que la sustitución que lleva $y \leftarrow a$ y $x \leftarrow b$ es un unificador ecuacional, ya que proporciona las expresiones $f(b, b, a)$ y $f(a, b, a)$ que son iguales en la teoría considerada.

La unificación ecuacional es mucho más potente que la unificación pero también es mucho más compleja ya que ahora dos objetos iguales pueden diferir mucho a nivel de símbolos, y dependerá de la propia teoría que un problema dado se pueda resolver o no de una forma simple, o que ni siquiera existan métodos eficientes de resolución (Siekman, 1989, Cachafeiro, 1994). Los conocimientos matemáticos previos suelen proporcionar información muy relevante acerca de las propiedades de la teoría a considerar. Así, las posibilidades de transformar una expresión polinómica en otra equivalente es muy diferente si se considera en Z, en Q, en R o en C (e.g. en Z no existe ningún valor de x para el que la expresión 3x+2=0 pueda convertirse en una igualdad en esa teoría).

Ejemplo 4

Unificar ecuacionalmente $2x+4$ con $3+x$, esto es, resolver $2x+4 = 3+x$ en las teorías Ø, N y Z.

En la teoría Ø (en la que no hay ecuaciones diferentes de la identidad sintáctica), no existen valores de la variable que conviertan esa expresión en una identidad y el problema no tiene solución. En N el problema $2x+4 = 3+x$ tampoco tiene solución. En Z esta ecuación tiene por solución $x = -1$. Esto significa que, al aplicar la sustitución $x \leftarrow -1$ a los términos $2x+4$ y $3+x$ se obtiene $2\cdot(-1)+4$ y $3+(-1)$ que son iguales en Z (representaciones del número 2). Por lo tanto esta sustitución es un unificador ecuacional (en Z) de $2x+4 = 3+x$.

Por ello, la unificación ecuacional es heredera y comparte con la investigación matemática los problemas y resultados concernientes a la resolución algebraica de ecuaciones, que es una de las primeras y últimas fuentes de producción de resultados matemáticos. Como ejemplo de ello se puede mencionar, por una parte, la matemática de la época babilónica y, por otra, al resultado de Matiyasevitch (1970) que demuestra la no decibilidad del 10° Problema de Hilbert.

Por otra parte, muchos métodos actuales de unificación ecuacional utilizan directamente resultados de la resolución de ecuaciones. Como ejemplo de ello se pueden mencionar: el método de Boole de resolución de ecuaciones (Martin y Nipkow 1989), la obtención de bases de soluciones de ecuaciones diofánticas (Boudet *et al.* 1990) y el resultado de Matiyasevitch y su aplicación al estudio de la decibilidad de la unificación ecuacional en otros problemas y teorías (Bockmayr 1987).

Como reconocimiento de que las etapas importantes de la investigación de resolución de ecuaciones también lo son para la Unificación Ecuacional en la tabla 3 se muestran algunas de ellas (Boyer 1986).

La unificación ecuacional aprovecha todos estos resultados conocidos y los extiende a nuevos campos como: la resolución de ecuaciones en teorías que no habían sido estudiadas hasta ahora, la resolución de problemas ecuacionales arbitrarios en una teoría dada o en grupos de teorías muy generales etc. Con este amplio e histórico bagaje y renovados métodos, la unificación ecuacional, que ya es usada explícitamente en algunos sistemas (e.g. demostración automática de teoremas y programación lógica y funcional) será una de las herramientas más importantes para la automatización del conocimiento.

Tabla 3

ALGUNAS ETAPAS RELEVANTES DE LA RESOLUCIÓN DE ECUACIONES

Época	BABILONIA II° milenio a.C.	GRECIA siglo V a.C.	DIOFANTO siglo III
Soluciones en	Q+	R+	Z+
Tipo de ecuación	1° y 2° grados	1° y 2° grados	1° y 2° grados
Generó	El número racional	El número real	Interés por la aritmética
Época	SCIPIONE del FERRO siglo XVI	ABEL siglo XIX	MATIYASEVITCH 1976
Soluciones en	R	R	Z
Tipo de ecuación	3° grado	grado n>4	polinómica de grado arbitrario
Generó	El número complejo	El álgebra moderna	?

4. Implicaciones en el currículum y conclusiones

Las investigaciones en unificación ecuacional concretamente y en teoría del conocimiento e IA en general, modificarán en uno u otro sentido el currículum actual de las Matemáticas en todos los niveles educativos y de otras disciplinas. Se pueden considerar algunas direcciones como las más probables de ser influidas por estas investigaciones:

— El análisis del papel histórico, evolución y rol actual de los conocimientos matemáticos. Este análisis es útil tanto en la matemáticas como en otras materias (e.g. filosofía, historia de la ciencia, ética). Permitirá comprender mejor el significado de la resolución de ecuaciones como antesala de nuevos conceptos e ideas acerca del conocimiento.

— La incorporación de nuevas teorías y problemas de resolución ecuacional (e.g. teorías con un operador conmutativo). Ello puede aparecer combinado con nuevos métodos de resolución de ecuaciones y el uso de soporte visual e informático para la corrección y la autocorrección.

— El estudio general de los métodos de transformación de sistemas en una visión integrada de todo un proceso. Esto incluye tanto la simplificación de expresiones numéricas o algebraicas, como inecuaciones, derivadas etc. Esta integración permite clarificar el tipo de operación e incidir más en las diferencias de unos y otros sistemas.

Para concluir incidimos en el significado de la unificación y de la unificación ecuacional como herramientas claves para realizar el contraste de datos, operación ésta que se encuentra en prácticamente todos los desarrollos de la Inteligencia Artificial. Asimismo destacamos la intensa relación entre estas operaciones y la resolución de ecuaciones. Para la unificación mostramos cómo el mecanismo de Martelli y Montanari tiene muchos puntos en común con el método de Gauss de resolución de ecuaciones. La relación entre la unificación ecuacional y la resolución de ecuaciones es aún mucho más profunda, ya que en ambos casos se persigue resolver un problema ecuacional en una teoría dada. Por ello, las investigaciones en unificación ecuacional pueden verse como una continuación de las ya realizadas en la resolución de ecuaciones, que ha sido uno de los campos más fructíferos de las Matemáticas. Se mencionaron también algunas partes del currículum que pueden verse en poco tiempo ampliadas o modificadas como consecuencia directa o indirecta de las investigaciones en el campo de la unificación.

BIBLIOGRAFÍA

BOCKMAYR (1987). A Note on a Canonical Theory with Undecidable Unification and Matching Problem. *Journal of Automateed Reasoning 3*, 379-381.

BOUDET, A., CONTEJEAN, E., DEVIE, H. (1990). A new AC unification algrithm with a new algorithm for solving diophantique equations. *IEEE Symposium on Logic in Computer Science* 289-299.

BOYER, C. B. (1986). *Historia de la Matemática.* Madrid: Alianza Editorial.

CACHAFEIRO CHAMOSA, L.C. (1994). *Algoritmos Xerales de Unificación Ecuacional.* Tese Dotoral. Universidad de A Coruña. Abril 1994.

GRAUBARD, S. R. (comp.) (1993). *El Nuevo Debate sobre la Inteligencia Artificial. Sistemas simbólicos y redes neuronales.* Barcelona: Editorial Gedisa.

JOHNSON-LAIRD, P.N. (1990). *El ordenador y la mente. Introducción a la ciencia cognitiva.* Barcelona: Editorial Paidós.

KNIGHT, K. (1989). Unification: A Multidisciplinary Survey. *ACM Computing Surveys 21,* 92-124.

MATIJASEVIC Y. (1970). Diophanine representation of recursively enumerable predicates. *Actes Congr\ès International des Mathématiciens 1,* 235-238.

ROBINSON, J.A. (1965). A Machine-Oriented Logic Based on the Resolution Principle. *Journal of Association Computing Machinery (JACM),* 12, 32-41.

MARTELLI, A., MONTANARI, U. (1982). An efficient unification algorithm. *Transaction on Programming Languages and Systems,* 4(2), 258-282.

MARTIN, U., NIPKOW, T. (1989). Boolean Unification — The Story So Far. *Journal of Symbolic Computation* 7, 275-293.

SIEKMAN, J. (1989). Unification Theory : a Survey. *Journal of Symbolic Computation,* 7, 207-274.

Resumen

En este trabajo se describen las operaciones de unificación y de unificación ecuacional, justificando su importancia como operaciones fundamentales en el campo de la automatización del conocimiento. Se muestra la estrecha relación de ambas con la resolución de ecuaciones y se consideran las implicaciones que puedan tener en el currículum especialmente en el de las matemáticas.

Palabras clave: inteligencia artificial, resolución de ecuaciones, currículum de matemáticas.

Abstract

In this paper, we describe the unification and equational unification operations justifying their importance as operations in the field of the automatization of knowledge. We show the close relationship of both operations with the equational resolution and we also consider some implications in the curriculum specially in the mathematicas study programmes.

Key words: artificial intelligence, equational resolution, mathematical curriculum.

Luis Carlos Cachafeiro Chamosa

I.B. Pontepedriña

Amor Ruibal sn Santiago de Compostela

email: cichamos@udc.es