Capítulo 1

INTRODUCCIÓN

1.1. Historia

El Análisis de Conceptos Formales (FCA) constituye una herramienta formal para el análisis de datos que permite la extracción de conocimiento a partir de un conjunto de objetos y de las propiedades que cumplen dichos objetos.

La teoría en su forma actual se remonta al grupo de investigación dirigido por Rudolf Darmstadt Wille, Bernhard Ganter y Peter Burmeister, donde se originó el análisis de conceptos formales en la década de 1980. La base matemática, sin embargo, ya se había creado por Garrett Birkhoff en la década de los 30 como parte de la teoría general de retículos. Antes del trabajo del grupo de Darmstadt, ya existían enfoques similares de varios grupos franceses, y sus fundamentos filosóficos apuntan principalmente a Charles S. Peirce y al pedagogo Hartmut von Hentig.

En su artículo "Reestructuración de la Teoría de Retículos" (de 1982, con el que se inicia el FCA como disciplina matemática) Rudolf Wille parte de un descontento con la teoría de retículos en particular y con las matemáticas puras en general.

La producción de resultados teóricos - a menudo alcanzados por medio de elaboradas "gimnasias mentales" - eran impresionantes, pero las conexiones entre dominios vecinos, o incluso entre partes de una misma teoría se estaban debilitando. La reestructuración de la teoría de retículos es un intento de revitalizar las conexiones entre dominios mediante la reinterpretación de la teoría de la manera más concreta posible con el fin de propiciar una mejor comunicación entre los teóricos de retículos y los usuarios potenciales de dicha teoría. Este objetivo se remonta a Hartmut von Hentig, quien en 1972 pidió una reestructuración de las ciencias con el fin de conseguir una mejor enseñanza y de hacer a la ciencia más asequible, tanto con el fin de conocerla mejor como para poder criticarla de una forma más cercana (es decir, sin necesidad de un conocimiento especializado).

El FCA corrige el punto de partida de la teoría reticular en el desarrollo de la lógica formal en el siglo XIX, donde un concepto, como predicado unario, se había reducido a su extensión. El objetivo fue trabajar con una visión de los conceptos menos abstracta, haciendo uso también de la intensión, una orientación que provenía de la lingüística y de la lógica conceptual clásica.

FCA tiene como objetivo aclarar los conceptos siguiendo la máxima de Charles S. Peirce de desplegar las propiedades observables y elementales de los objetos. En su filosofía tardía, Peirce supone que el pensamiento lógico tiene como objetivo percibir la realidad por medio de la triada: concepto, juicio y conclusión. En este sentido, Wille dice:

El objetivo y el significado del FCA como teoría matemática sobre conceptos y sus jerarquías es apoyar la comunicación racional entre seres humanos mediante el desarrollo matemático de estructuras conceptuales apropiadas que se puedan manipular con *la lógica.*

1.2. Contextos y conceptos

El FCA tiene como entrada una tabla con una relación entre un conjunto de objetos y un conjunto de propiedades (atributos).

Esta entrada es lo que se denomina **contexto formal** y la relación que se define en dicha tabla determina qué objetos cumplen qué propiedades.

Un concepto formal se define como un par formado por un conjunto de objetos (**extensión**) y un conjunto de atributos (**intensión**) de forma que la extensión está formada por todos los objetos que comparten los atributos dados y la intensión por los atributos compartidos por los objetos dados.

Por lo que, un **contexto formal** se puede definir como una tripleta *K=(X,Y,I)*, donde X es un conjunto de objetos, Y es un conjunto de atributos e I una relación binaria, $I \subseteq X \times Y$, que muestra qué objeto posee qué atributo. Formalmente, se puede considerar como un grafo bipartito que refleje las relaciones entre los conjuntos X y Y, o también como una tabla con los objetos ocupando las filas y los atributos en las columnas, y de forma que un valor booleano en la celda *(x,y)* significa que el objeto x tiene el atributo y.

	necesita agua para vivir	vive en el agua	vive en la tierra	necesita clorofila	dicotiledónea	monocotiledónea	puede moverse	tiene extremidades	lactantes
sanguijuela	×	×					×		
brema	×	×					×	×	
rana	×	×	×				×	×	
perro	×		×				×	×	×
maleza acuática	×	×		×		×			
caña	×	×	×	×		×			
judía	×		×	×	×				
maíz	×		×	×		×			

Contexto Formal

1.3. Implicaciones y reglas de asociación con FCA

Una vez formalizados matemáticamente los conceptos, puede resultar relativamente sencillo trasladar las relaciones lógicas que encontramos entre atributos... aunque debemos tener en cuenta que las relaciones lógicas que se obtengan son aquellas que vengan confirmadas por nuestro contexto, que podría representar un conocimiento concreto en un instante de tiempo, pero que podría variar al añadir un mayor número de objetos o atributos. Por ejemplo, que en el ejemplo siguiente podamos decir que "ser animal de jungla" implica "ser mamífero" se debe a que todos los objetos que verifican la propiedad "ser animal de jungla" verifican "ser mamífero", pero en el momento que introduzcamos un objeto nuevo que sea de la jungla (por ejemplo, un manglar) y no sea mamífero, esa implicación deja de ser cierta. Por supuesto, podemos trabajar con varios atributos simultáneamente.

Consecuentemente, en FCA podemos formalizarlo de la siguiente forma: dados A y B subconjuntos de atributos, diremos que se tiene la implicación $A \to B$ si se verifica $A' \subseteq B'$, es decir, todos los objetos que tienen cada atributo de A también tienen cada atributo de B (observa que es coherente con la implicación intuitiva que dimos en el apartado anterior).

Con esta definición, las implicaciones obedecen las reglas de Armstrong (reflexiva, aumentativa y transitiva) comunes en las dependencias funcionales que se dan entre los atributos de una base de datos:

Reflexividad: $B \subseteq A \quad A \to B$

Aumento: $A \to B \quad A \cup C \to B \cup C$

Transitividad: $A \to B, B \to C \quad A \to C$

A partir de la definición de implicación y de las propiedades básicas que verifica, podemos definir un cálculo lógico que nos permitirá realizar sistemas de deducción completos sobre el contexto actual. En cierta forma, hemos pasado de tener un conocimiento por ejemplos a disponer de un conocimiento abstracto que introduce sistemas de razonamiento más elaborados en nuestro mundo, partiendo únicamente de las observaciones concretas que hemos realizado, es decir, hemos aprendido reglas generales a partir de ejemplos.

1.4. Aplicaciones

Desde su introducción ha sido aplicado en campos tan variados como la minería de datos, minería de textos, gestión del conocimiento, web semántica, desarrollo de software, biología, etc.

La doctora Karell Bertet, investigadora de la Universidad de La Rochelle con la que colabora el director del proyecto, ha desarrollado una librería java, java-lattices, para la generación, representación y manipulación de los Conceptos Formales, etc.
Dicha librería implementa la generación de retículos. Se pueden generar:

- Un retículo dados los nodos y sus relaciones.
- Álgebras booleanas de 2^n elementos.
- Retículos de permutaciones.
- Retículos aleatorios de n nodos.
- Retículos de conceptos a partir de un contexto o de un conjunto de implicaciones.
- Cálculo de implicaciones y de bases de implicaciones.
- Etc.