

Terminología

Heurística y minería



La autora y Carlos González fotografiados por Luis Argüelles al acabar un proceso de minería de datos en las instalaciones de una mina de carbón, abril de 2004

La heurística¹ -técnica de la indagación y el descubrimiento según el DRAE- es la principal herramienta de resolución de problemas en inteligencia artificial. En “Computer Science as Empirical Inquiry: Symbols and Search” (Communications of the ACM vol. 19.3, 1976) Allen Newell and Herber A. Simon, ganadores del premio Alan Turing del año 1975, definen la informática como una ciencia empírica que estudia los fenómenos que rodean a los ordenadores. El ordenador no es solo hardware sino junto al programa una máquina viva, un sistema físico de símbolos que resuelve problemas mediante búsquedas heurísticas.

Para resolver un problema con un computador hay que generar estructuras de símbolos que lo representen y determinar cuáles son las que representan la solución. Una vez establecidas, el procedimiento de resolución tiene que buscar las soluciones entre todas las estructuras, y la heurística acelera la búsqueda.

Resolver un problema exige reconocer las soluciones si las encontramos, es decir saber qué buscamos. En contraposición, existen procedimientos en inteligencia artificial que permiten aflorar conocimientos ocultos no conocidos a priori; por ejemplo aflorar información contenida pero no explícita en grandes bases de datos. El conjunto de estos procedimientos, que a su vez recurren a técnicas diversas, se denomina “minería de datos” o más recientemente “descubrimiento del conocimiento”, término mucho más adecuado en mi opinión, aunque yo por el carácter provocador al que no renuncio en esta sección, haya aludido al primero en el título.

Conseguir nuevo conocimiento es un problema filosófico que ya está planteado en el diálogo de Platón sobre la virtud: Menón le pregunta a Sócrates si es posible indagar en lo que no se conoce y si en el supuesto de que se encontrase, sería posible reconocerlo. El saber es un camino de búsqueda del alma en su interior, el alma recuerda lo que sabe porque lo ha conocido en una vida anterior y descubre que ya lo sabía, responde Sócrates.

La cuestión que plantea Menón se puede ligar al concepto de descubrimiento del conocimiento y la respuesta de Sócrates al hecho de que bases de datos creadas para albergar información conocida, pueden reencarnarse en otras y recordar que contienen otros conocimientos. Es como si las técnicas de descubrimiento del conocimiento pudieran dar respuesta a Menón de acuerdo con las enseñanzas de Sócrates.

¹ También, en algunas ciencias, manera de buscar la solución de un problema mediante métodos no rigurosos como por tanteo, reglas empíricas, etc.

Ya basta. Acabo proponiendo una traducción al español de las expresiones clave de las listas de IFAC, IEEE, ACM y AMS en las que aparece la palabra heurística y las expresiones “minería de datos” o “descubrimiento del conocimiento”, siguiendo convenciones de números anteriores. Tomo la lista de la IFAC como referencia ordenada en inglés alfabéticamente y distingo con (IEEE) (ACM) o (AMS) las claves que aparecen también en la listas del IEEE, ACM o AMS respectivamente. Las palabras clave que no aparecen en la lista de IFAC las distingo con un asterisco (*).

Claves en las que aparece la palabra heurística

- | | |
|--|--------------------------------------|
| 1. Approximation methods and heuristics* (AMS) | Métodos de aproximación y heurística |
| 2. Heuristic methods *(ACM) | Métodos heurísticos |
| 3. Heuristic programming | Programación heurística |
| 4. Heuristic researches | Búsquedas heurísticas |
| 5. Heuristic strategies* (AMS) | Estrategias heurísticas |
| 6. Heuristics (AMS, UNESCO)) | Heurística |
| 7. Neural and heuristic methods | Métodos heurísticos y neuronales |

La clave minería de datos solo aparece en la lista del ACM dentro del apartado de aplicaciones de las bases de datos (H2.8). La clave descubrimiento del conocimiento no aparece en ninguna lista. Es curioso observar que en la lista de IEEE no aparece ninguna de las tres claves.

Hasta el próximo número.

Teresa de Pedro
Investigadora Científica
Instituto de Automática Industrial, Agencia Estatal CSIC
tere@iai.csic.es