Big Data Analytics: A Cognitive Perspectives

Yingxu Wang, Chongqing University of Science and Technology, School of Electrical and Information Engineering, Chongqing, China
Jun Peng, Chongqing University of Science and Technology, School of Electrical and Information Engineering, Chongqing, China

ABSTRACT

Big data are pervasively generated by human cognitive processes, formal inferences, and system quantifications. This paper presents the cognitive foundations of big data systems towards big data science. The key perceptual model of big data systems is the recursively typed hyperstructure (RTHS). The RTHS model reveals the inherited complexities and unprecedented difficulty in big data engineering. This finding leads to a set of mathematical and computational models for efficiently processing big data systems. The cognitive relationship between data, information, knowledge, and intelligence is formally described.

KEYWORDS

Big Data, Big Data Engineering, Big Data Science, Cognitive Foundations, Cognitive Informatics, Denotational Mathematics, Recursively Typed Hyperstructures

INTRODUCTION

Data are the most fundamental and pervasive cognitive objects in the brain that link the real-world entities and attributes to mental abstractions via sensation and quantification. Data are an abstract representation of the quantity of realistic entities or abstract objects against specific measurement scales (von Neumann, 1946; Ullman & Widom, 1997; Tucker, 1992; Lewis & Papadimitriou, 1998; Sternberg, 1998; Chicurel, 2000; Chapra & Canale, 2002; Jacobs, 2009; Hassanien et al., 2015; Wang, 2003, 2007, 2015a, 2016c).

Big data are extremely large-scaled heterogeneous data in quantity, complexity, retain, retrieval, semantics, cognition, distribution, maintenance, and processing costs. Big data are pervasively manipulated across contemporary science disciplines such as computer science, information science, cognitive informatics, web-based computing, cloud computing, social networks and computational intelligence (Wang, 2015a). The inherent complexity of big data and the exponentially increasing demands on big data have create unprecedented problems in all aspects and phases of big data engineering. The challenges stem from not only the oversized magnificent and complexity of datasets beyond classical handling capacity of theories and technologies, but also the extended domain of big data out of the traditional domain of real numbers (R) (Wang, 2016c).

Basic characteristics of big data are unstructured, heterogeneous, monotonous growing, mostly nonverbal, hybrid, unclear semantics, decay in consistency, and increase in entropy over time (Wang, 2006, 2016b). Big data plays an indispensable role not only in a wide range of engineering applications, but also in the cognitive mechanisms of human sensation, quantification, qualification, estimation, memory, and reasoning (Jacobs, 2009; Snijders et al., 2012; Wang, 2015a; Wang et al., 2016c).
The taxonomy of cognitive objects represented in human brains can be classified into four forms known as data, information, knowledge, and intelligence in a hierarchical structure from the bottom up according to their levels of abstraction (Berkeley, 1954; Turing, 1950; Shannon, 1956; von Neumann, 1958; McCarthy et al. 1955; McCulloch, 1965; Debenham, 1989; Bender, 2000; Hassanien et al., 2015; Wang, 2009a, 2010, 2014a, 2015a, 2015c, 2016a, 2016d). It is recognized that almost all fields and hierarchical levels of human activities generate exponentially increasing data, because data plays an indispensable role in fundamental cognitive mechanisms of humans such as sensation, quantification, qualification, estimation, memory, reasoning, and knowledge generation.

This paper presents the cognitive foundations of big data in human reasoning and engineering applications. In the remainder of this paper, the cognitive quantification of big data is explored in Section 2. The computational generation of big data and the type theory for big data modeling are described in Section 3. A set of mathematical models of big data systems is introduced in the form of recursively typed hyperstructures beyond \( \mathbb{R} \) in Section 4. The hierarchical properties and transformability of big data among other cognitive objects are elaborated in Section 5. The cognitive foundations of big data analytics provide a theoretical framework of big data science and its engineering applications.

**The Cognitive Foundations of Big Data Science**

Big data are pervasively generated by human cognitive processes, formal inferences, system quantifications, mathematical operations, computing manipulations, and information processing. It is recognized that data are not tangible physical entities in the realistic world, because they are abstract quantifications of realistic entities, attributes, and relations. The cognitive foundations of big data can be elaborated in the facets of their semantical, quantificational, and type properties.

**The Cognitive Model of Data**

A conceptual model of data as a formal concept can be formally described according to the methodology of concept algebra (Wang, 2015c).

**Definition 1.** The conceptual model of data, \( C_d \), is a hyperstructure of attributes \( (A) \), objects \( (O) \), as well as their internal \( (R^i) \), input \( (R^i) \), and output \( (R^o) \) relations with respect to external concepts in the context of knowledge \( K \), i.e.:

\[
C_d \{ \text{data}: A, O, R^i, R^o, R^e \}
\begin{align*}
A &= \{ \text{cognitive\_object*}, \text{abstraction}, \text{quantity}, \text{fact}, \text{figure}, \text{unit} \} \\
O &= \{ \text{sensory}_d, \text{observed}_d, \text{experimental}_d, \text{measured}_d, \text{typed}_d, \\
    &\quad \text{derived}_d, \text{statistical}_d \} \\
R^e &= O \times A \\
R^i &\subseteq \mathcal{R} \times C \\
R^o &\subseteq C \times \mathcal{R}
\end{align*}
\]

As given in Definition 1, the primitive attribute of \( C_d \{ \text{data} \} \) is identified as a cognitive\_object that represents a collective super concept where the concept belongs to. Other attributes in \( A \) represent the subfacets of intentions of the concept. The conceptual structure of data in Equation 1 is a Cartesian product of international relations \( R^e = O \times A \) between the extension and intension of the concept. It is noteworthy that both sets of attributes and objects in the formal concept are mutually constrained,
Abstract Intelligence: Embodying and Enabling Cognitive Systems by Mathematical Engineering
www.igi-global.com/article/abstract-intelligence/175662?camid=4v1a

Development of an Ontology for an Industrial Domain
www.igi-global.com/chapter/development-ontology-industrial-domain/27314?camid=4v1a