Journal of Network and Computer Applications ■ (■■■■) ■■■–■■■

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

journal homepage: www.elsevier.com/locate/jnca

Semantic relation computing theory and its application

Yunchuan Sun [a], Cheng Lu [b], Rongfang Bie [b,*], Junsheng Zhang [c]

[a] Business School, Beijing Normal University, Beijing, China
[b] College of Information Science and Technology, Beijing Normal Univerisity, Beijing, China
[c] Institute of Scientific and Technical Information of China, Beijing, China

ARTICLE INFO

Article history:
Received 16 October 2013
Received in revised form
11 June 2014
Accepted 23 September 2014

Keywords:
Orthogonal basis
Semantic link
Semantic link network
Semantic relation
Semantic relation basis
Reasoning rule

ABSTRACT

Semantic relations among objects are primary semantic factors, which play the most important role for human and smart systems /machines to understand and control the situation in the context of connected systems. However, few existing works focus on the study of semantic relations from the mathematical view, though it would be the basis for further research on semantics. Existing research works focus on the representation of semantic relations and reasoning with relations, but seldom concentrate on the semantic relation computing including accurate reasoning, integrity checking and redundancy checking. In this paper, we aim at exploring algebraic computing approach of semantic relations. A mathematical computing theory for establishing semantic relations – semantic relation space theory – is proposed for the first time, including conceptions of semantic relation basis, orthogonal basis, and some basic operations for semantic relations. The proposed theory can be used to represent the semantic relations among objects in an accurate way and to deduce implicit relations in the connected smart systems, especially for automatic reasoning and autonomous computing. Furthermore, integrity and consistency issues among semantic relations are also discussed based on the theory. A case study in scientific research domain has shown the feasibility and effectiveness of the proposed theory.

1. Introduction

A cyber-physical system (CPS) is a new complex embedded system combining computing, communication and control technologies. The computing process and physical process realize the interaction and deep fusion in the open environment and achieve the integration of the open embedded computing, network real-time communication with the precise remote control and other advanced features (Sowa, 2008). As a complex system for integrating information space and physical space, the CPS uses embedded sensors to obtain relevant data information in the physical world. Through the interconnected network system of information transmission, it realizes the use of information in the world for intelligent information storage and processing, and then it could make judgments and decisions. So the CPS realizes the integration of the information space and physical space and seamless interaction.

The foremost task of cyber-physical systems is to represent objects and the relations between two different objects. The construction of relations is the foundation to realize the intelligent applications for the cyber-physical systems. So the identification of semantic relations between objects is an important research problem. Most of the existing models are focusing on the semantic relation representation and simple reasoning over the semantic relations. However, few research works focused on the mathematical modeling and theory on semantic relations. This paper aims to explore the theory of semantic relation computing and the applications.

1.1. Motivation

Semantic relations among different objects play a primary role for human and intelligent systems to grasp the situation accurately in the complicated environment. Information scientists have been always trying their best to describe the complex relations among physical objects in the virtual cyber space and made a long list of achievements in the past decades.

Semantic network is a famous model which represents knowledge among different conceptions with the semantic relations (Sowa, 2008), which has been widely used in artificial intelligence and machine translation, such as building a translation approach to portable ontology specifications (Gruber, 1993). WordNet[1] is a successful application of the semantic network of English words which can record the various semantic relations between synonym sets of English words. As a concept model for designing database, the

* Corresponding author.
E-mail addresses: yunch@bnu.edu.cn (Y. Sun), 201221210025@mail.bnu.edu.cn (C. Lu), rfbie@bnu.edu.cn (R. Bie), zhangjs@istic.ac.cn (J. Zhang).

[1] http://en.wikipedia.org/wiki/WordNet.

http://dx.doi.org/10.1016/j.jnca.2014.09.017
1084-8045/

entity-relationship model (ER model) which has been widely used during developing information systems currently, describes the relationship between two or more entities by attaching a semantic label on the relations (Chen, 1976). Network model for database is conceived as a flexible way of representing objects and their relationships (Bachman, 1973). Its distinguishing feature is that the schema, a graph in which object types are nodes and relationship types are arcs, is not restricted to be a hierarchy or lattice. Meanwhile, the relational data model represents relations among entities with the unique form of normalized tables (Codd, 1969, 1970). Resource Description Framework[2](RDF) uses triples with three components (namely subject, predicate and object) to represent the relationships between things by extending the linking structure of the current Web, where predicate denotes the relationship. RDF is a standard model for data interchange on the Web with the support of the features facilitating data merging. Its linking structure forms a directed, labeled graph, where the edges represent the named links between two resources (nodes) (Lassila and Swick, 1999). Linked Data aims to build semantic links between data on the semantic web according to the internal relationship among data (Heath and Bizer, 2011). Semantic link network (SLN) is a loosely coupled semantic data model for managing web resources, and it focuses on the semantic relations between resources with semantic links (Zhuge, 2012).

Semantic relations among objects are regarded as the most important elements during mapping the physical world into the cyber world (Li et al., 2012; Qian et al., 2011). These models provide efficient ways of representation for semantic relations and allow users to freely define semantic relations according to the practical applications; however, none solid theory on semantic relation computing has been studied, and there are few works on the study of the relationships between semantic relations among objects from the mathematical view though it plays a primary role in semantic relation computing and reasoning.

1.2. Related work

There are some research works on SLN focusing on the construction, management and reasoning of semantic relations. A set of reasoning rules among semantic link types can be used to derive out implicit semantic relationships among resources (Zhuge, 2009). SLN schema provides a theory of normalized forms to manage and query in the complicated network by studying of the internal relationships among semantic link types (Zhuge and Sun, 2010). An SLN schema, represented as a triple $\hat{S}$(*ResourceTypes*, *LinkTypes*, *Rules*), specifies resource types $\{rt_1, rt_2, \ldots, rt_k\}$, semantic link types with the form of $rt_i \Longrightarrow^{\alpha} rt_j$, and reasoning rules on semantic links (Zhuge and Sun, 2010). An instance of semantic link network can be built and regulated in a normal way with the schema. Semantic relation discovery from large volumes of Semantic Web data is a fundamental technology in analytical domains, such as business intelligence and homeland security. An approach is proposed to employ constituent dependency information for tree kernel-based semantic relation extractions between entities (Yang et al., 2012).

Reasoning is one of the primary features of the semantic link network (Zhuge and Sun, 2010; Sun et al., 2013). Rule-based reasoning is the main form of reasoning in semantic link network though analogy reasoning and inductive reasoning (Zhuge, 2012; Zhang and Sun, 2010) are also supported. In SLN model, semantic relations between resources are represented by semantic links attached semantic factor. By using the basis of semantic relations, all semantic links between resources can be represented with a semantic link matrix. The potential semantic links can be deduced by matrix computing through rule-based reasoning (Zhuge, 2009; Sun et al., 2012).

Beside of semantic relation reasoning, recently there has been increasing interest in predicting links between objects in complex networks, which can be helpful in many data mining tasks. Link prediction in complex networks has attracted increasing attention from both physical and computer science communities. Link prediction has been studied in a co-authorship network as an instance of social network to predict the interactions among researchers (Liben-Nowell Kleinberg, 2007). Algorithms (the random-walk-based methods and the maximum likelihood methods etc.) can be utilized to extract missing information, identify spurious interactions, infer the new interactions likely to occur in the near future, evaluate network evolving mechanisms, and so on (Zhang and Sun, 2010). In addition, numerous methods based on tensor factorization also have been proposed in link prediction (Ullman, 1976).

Link similarity has been more and more used in link prediction. Zeng et al. (2012) proposed a novel tensor-based prediction method, which is designed through two steps: First, tracking time-dependent network snapshots in adjacency matrices which form a multi-way tensor by using an exponential smoothing method. Second, applying common neighbor algorithm to compute the degree of similarity for each nodes (Zeng et al., 2012). Zhuge (2007) developed approaches to link prediction based on measures for analyzing the "proximity" of nodes in a network, in which experiments on large co-authorship networks suggest that information about future interactions can be extracted from network topology alone, and that fairly subtle measures for detecting node proximity can outperform more direct measures (Zhuge, 2007).

Link similarity is also widely used in community discovery. Inspired by the fact that elements in the same community are more likely to share common links, Wu et al. (2012) propose a method to find local community structure by analyzing link similarity between the community and the vertex which explores community structure heuristically by giving priority to vertices which have a high link similarity with the community. Based on citation semantic link network scheme, Chen and Wang (2012) propose a method to calculate the similarity of citation semantic link network by which researchers can discover semantic community of various research fields in citation network, and foresee developing directions and hot spots of various research fields.

There are lots of researches on graph pattern matching, in which subgraph isomorphism is most widely discussed (Lü and Zhou, 2011). But subgraph isomorphism only concentrates on the structure matching, and neglects the semantic matching. Recent works on graph pattern matching, such as tree pattern queries on graph-structured data, usually choose to avoid queries with semantic information (Spiegel et al., 2012).

This paper aims at exploring algebra theory of semantic relations between objects in the CPS. Our contributions include: (1) a space theory for semantic relations is proposed, which involves in formal definition of semantic relations and relavant operations, semantic relation basis and orthogonal basis; (2) a matrix-based computing approach for semantic relations is also proposed based on the space theory; (3) a case study in scientific research domain indicates the feasibility and effectiveness of the proposed theory and approach.

The rest of this paper is organized as follows. In Section 2, the semantic relation space theory is proposed by introducing semantic relation operations and conceptions including semantic relation basis and its orthogonalization. In Section 3, we present the methods and algorithms for rule-based semantic relation computing. In Section 4, integrity and consistency issues involved in semantic relations are discussed. In Section 5, a case study in scientific research field is presented to illustrate the applications of the proposed theory. Finally, in Section 6, we draw the conclusion.

[2] http://www.w3.org/TR/PR-rdf-syntax.

Please cite this article as: Sun Y, et al. Semantic relation computing theory and its application. Journal of Network and Computer Applications (2014), http://dx.doi.org/10.1016/j.jnca.2014.09.017