Sensors & Transducers, Vol. 159, Issue 11, November 2013, pp. 167-173

www.sensorsportal.com

Sensors & Transducers

http://www.sensorsportal.com

Conceptual Analysis of Node Application Program of Semantic Reasoning Network

Shi Yun Ping
Xi'an University of Posts & Telecommunications, 710121, China
E-mail: syp@xupt.edu.cn

Received: 3 September 2013 /Accepted: 25 October 2013 /Published: 30 November 2013

Abstract: In this paper, we employed Semantic Reasoning Tight Method, Random Process character subset and Global Optimization Search to rank Node Application Program. The Semantic Reasoning Network is demonstrated on a dataset come from text. Here, the characteristics are treated as criteria, to classify Node Application Program. We introduce the Related Work on Conceptual Analysis and the Terms. From the experiments, we conclude that Tight method based Semantic Reasoning Network performs better than other two techniques comparable to the Conceptual Analysis done in text. *Copyright © 2013 IFSA.*

Keywords: Node application program, Operation analysis (OA), Semantic reasoning Network, Random process, Terms.

1. Introduction

Services are tendered and availed in almost all the business and industries. The growth and proliferation of IT across industries and business appear to have fuelled the requirement as well as the delivery of services profoundly. Delivering services has been an attractive business proposition for many industries lately. The latest development in the systems is a new paradigm, called Node Application Program [1]. Node Application Program heralded another significant mile stone in the history of IT. Earlier, Internet catered mostly to the business to Customer (B2C) category of the users on the web. As against this, Node Application Program enable B2B interaction as well Web. They are independent of platform and natural languages, which is suitable for accessing from heterogeneous environments. With the rapid introduction of Node Application Program technologies, researchers focused more on the functional and interfacing aspects of Node Application Program, which include HTTP and XML-based messaging. They are used to communicate by employing pervasive standards based web technologies. Node Application Program is based on XML and three other core technologies: WSDL, SOAP, and UDDI. WSDL is a document which describes the services' location on the web and the functionality the service provides. Information related to the web service is to be entered in a UDDI registry, which permits web service consumers to find out and locate the services they required. With the help of the information available in the UDDI registry based on the Node Application Program, client developer uses instructions in the WSDL to construct SOAP messages for exchanging data with the service over HTTP characteristics [2].

In this study, we address this problem of efficiently identifying a set quality characteristics by employing Semantic Reasoning Networks vs. Tight Semantic Reasoning, Random Process character subset and Global Optimization search. Tight

Semantic Reasoning is special form of Semantic Reasoning Network that is widely used for Conceptual Analysis [3] and clustering [4], but its potential for general probabilistic modeling (i.e., to answer joint, conditional and marginal queries over arbitrary distributions) remains largely unexploited. Tight Semantic Reasoning represents a distribution as a mixture of components, where within each component all variables are assumed independent of each other. The Random Process character subset of a variable Y, (MB(Y)), by definition, is the set of variables such that Y is conditionally independent of all the other variables given MB(Y). A Random Process Character subset Directed Connection Graph (MB DCG) is a Directed Connection Graph over that subset of variables. When the parameters of the MB DCG are estimated, the result is a Semantic Reasoning Network, defined in the next section. Recent research by the Machine Learning Community [5-7] has sought to identify the Random Process character subset of a target variable by filtering variables using statistical decisions for conditional independence and using the MB predictors as the input features of a classifier. However, learning MB DCG classifiers from data is an open problem [8]. There are several challenges: the problem of learning the graphical structure with the highest score (for a variety of scores) is NP hard [8] for methods that use conditional independencies to guide graph search, identifying conditional independencies in the presence of limited data is quite unreliable and the presence of multiple local optima in the Global Optimization search Enhanced Random Process character subset Classifier space of possible structures makes the search process difficult.

Conceptual Analysis using the Random Process character subset of a target variable in a Semantic Reasoning Network has important properties. It specifies a statistically efficient prediction of the probability distribution of a variable from the smallest subset of variables that contains all of the information about the target variable, it provides accuracy while avoiding over fitting due to redundant variables and it provides a classifier of the target variable from a reduced set of predictors. The TS/MB procedure proposed in this paper allows us to move through the search space of Random Process character subset structures quickly and escape from local optima, thus learning a more robust structure.

2. Quality Issues in Node Application Program

OA plays an important role in finding out the performance of Node Application Program. Earlier, OA has been used in networking and multimedia applications. Recently, there is a trend in adopting this concept to Node Application Program [9]. The basic aim is to identify the OA characteristics [10-11] for improving the quality of Node Application Program through replication services [10], load distribution, and service redirection [4]. To measure the OA of a web service, characteristics like Response Time, Throughput, Availability, Reliability, Cost, and Response Time are considered.

2.1. OA Characteristics

Users of Node Application Program are not human beings but programs that send requests for services to web service providers. OA issues in Node Application Program have to be evaluated from the perspective of the providers of Node Application Program (such as the airline-booking web service) and from the perspective of the users of these services (in this case, the travel agent site) [6]. There are other models available related to the quality of Node Application Program issues. An OA model represented in Table 1 shows that Operation Analysis model of Node Application is based on internal characteristics, which are independent of the service environment, and external characteristics that are dependent on the service environment. The characteristics of the model in Table 1 are almost similar to the characteristics of QWS Dataset used in this paper.

Table 1. Operation Analysis attributes of Node Application.

OA Factor	Internal attributes	External Attributes
Reliability	Correctness	Availability and consistency
Performance	Efficiency	Load management
Integrity	-	Security
Usability	I/O attributes	-

2.2. Description of QWS Dataset

QWS dataset consists of different web service implementations and their characteristics. The Conceptual Analysis is measured based on the overall quality rating provided by all the characteristics. The functionality of the Node Application Program can be helpful to differentiate between various services. The characteristics G1 to G10 are used as explanatory variables and the attribute G11 is used as the target variable. However, characteristics G12 and G13 are ignored as they do not contribute to the analysis.

The Node Application Program [1, 2] in the QWS dataset are classified into four categories, such as:

1) Platinum (high quality);
2) Gold;
3) Silver;
4) Bronze (low quality).

The Conceptual Analysis is measured based on the overall quality rating provided by VSRE. It is grouped into a particular web service based on Conceptual Analysis. The functionality of the Node Application Program can be helpful to differentiate between various services.

3. Overview of Semantic Reasoning Networks

A Semantic Reasoning Network is a directed Connection graph model that represents conditional independencies between a set of variables. It has two constituents: One is a network graphical structure which is a directed Connection graph with the nodes of variables and arcs of relations. The other is the conditional probability table associated with each node in the model graph. Machine learning techniques are able to estimate the structure and the conditional probability table from the training data. Based on the Semantic Reasoning probability inference, the conditional probability can be estimated from the statistical data and propagated along the links of the network structure to the target label. By setting a threshold of confidence, the final probability value can be used as the indication for the Conceptual Analysis decision. The Semantic Reasoning formula can be mathematically expressed as below:

$$P(H_J \mid E) = \frac{P(E \mid H_J) \times P(H_J)}{\sum_{i=1}^{n} P(E \mid H_i) \times P(H_i)} \quad (1)$$

$$j = 1,2,\cdots,n$$

According to the basic statistical theory, e.g., the Chain Rule and independency relation derived from the network structure, the joint probability of E can be calculated by the production of local distributions with its parent nodes,

$$P(E) = \prod_{i=1}^{n} P(E_i \mid Parent \quad of \quad (E_i)) \quad (2)$$

In the above formulas, E denotes a set of variable values, i.e. E= {E_1, E_2, ···, E_n}. H is termed as hypothesis.

H is called the prior probability and P(H|E) is called posteriori probability of H given E. If E_i has no parent nodes, Parent Of (E_i) is equal to P(E_i).

3.1. Tightly Method Semantic Reasoning

Tight Semantic Reasoning models are so named for their "naive" assumption that all variables Xi are mutually independent given a "special" variable C. The joint distribution is then given Tight by

$$P(C, X_1, \cdots, X_n) = P(C) \prod_{i=1}^{n} P(X_i \mid C)$$

The single variable conditional distributions P(X_i|C) can take any form (e.g., multinomial for discrete variables, Gaussian for continuous ones). When the variable C is observed in the training data, naive Semantic Reasoning can be used for Conceptual Analysis, by assigning test example (X_1, ···, X_n) to the class C with highest P(C|X_1, ···, X_n) [2]. When C is unobserved, data points (X_1, ···, X_n) can be clustered by applying the EM algorithm with C as the missing information, each value of C corresponds to a different cluster, and P(C|X_1, ···, X_n) is the point's probability of membership in cluster C. Tight Semantic Reasoning models can be viewed as Semantic Reasoning Networks in which each X_i has C as the sole parent and C has no parents. A naive Semantic Reasoning model with Gaussian P(X_i|C) is equivalent to a mixture of Gaussians with diagonal covariance matrices. While mixtures of Gaussians are widely used for density estimation in continuous domains, naive Semantic Reasoning models have seen very little similar use in discrete and mixed domains. However, they have some notable advantages for this purpose. In particular, they allow for very efficient inference of marginal and conditional distributions. To see this, let X be the set of query variables, Z be the remaining variables, and k be the number of mixture components (i.e., the number of values of C).

A slightly richer model than naive Semantic Reasoning which still allows for efficient inference is the mixture of trees, where, in each cluster, each variable can have one other parent in addition to C. The basic graph of Semantic Reasoning Network is presented in Fig. 1.

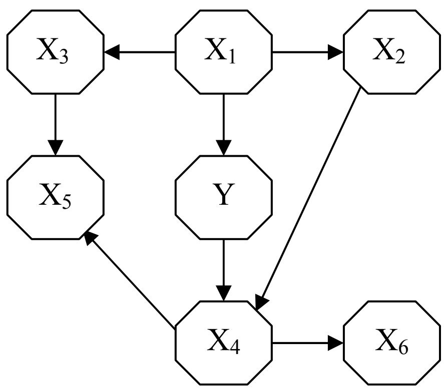

Fig. 1. The Semantic Reasoning Network.

3.2. Random Process Character Subset

The Random Process condition implies that the joint distribution P can be factorized as a product of conditional probabilities, by specifying the distribution of each node conditional on its parents.

In particular, for a given DCG S, the joint probability distribution for X can be written as

$$P(X) \models \prod_{i=1}^{n} P_i(X_i \mid Pa_i) \quad , \tag{3}$$

where Pa_i denotes the set of parents of X_i in S; this is called a Random Process factorization of P according to S. The set of distributions represented by S is the set of distributions that satisfy the Random Process condition for S. If P is faithful to the graph S, then given a Semantic Reasoning Network (S, P), there is a unique Random Process character subset for Y consisting of PaY, the set of parents of Y, chY, the set of children of Y, and Pa chY, the set of parents of children of Y.

For example, consider the two DCGs in Fig. 1 and Fig. 2. The factorization of P entailed by the Semantic Reasoning Network (S, P) is

$$\begin{aligned} &P(Y, X_1, \cdots, X_6) = P(Y \mid X_1) \cdot P(X_4 \mid X_2, Y) \\ &\cdot P(X_5 \mid X_3, X_4, Y) \cdot P(X_2 \mid X_1) \cdot P(X_3 \mid X_1) \\ &\cdot P(X_6 \mid X_4) \cdot P(X_1) \end{aligned} \tag{4}$$

The factorization of the conditional probability P(Y |X_1, ···, X_6) entailed by the Random Process character subset for Y corresponds to the product of those (local) factors in equation (2) that contain the term Y.

$$\begin{aligned} &P(Y \mid X_1, \cdots, X_6) = C_0 \cdot P(Y \mid X_1) \\ &\cdot P(X_4 \mid X_2, Y) \cdot P(X_5 \mid X_3, X_4, Y) \end{aligned} \tag{5}$$

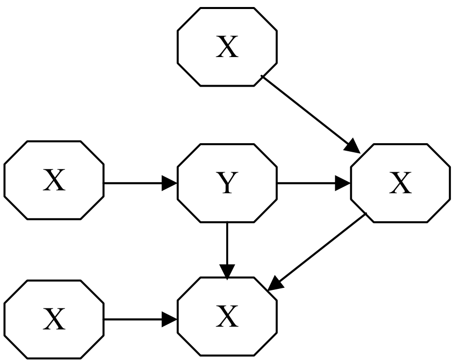

Fig. 2. A Random Process DCG for Y.

3.3. Global Optimization Search

A Heuristic is an algorithm or procedure that provides a shortcut to solve complex decision problems. Heuristics are used when you have limited time and/or information to make a decision. For example, some optimization problems, such as the travelling salesman problem, may take far too long to compute an optimal solution. A good heuristic is fast and able find a solution that is no more than a few percentage points worse than the optimal solution. Heuristics lead to good decisions most of the time, but not always. There have been several Meta-heuristic applications recently in Machine Learning, Evolutionary Algorithms, and Fuzzy Logic problems. Global Optimization Search is a meta-heuristic strategy that is able to guide traditional local search methods to escape the trap of local optimality with the assistance of adaptive memory. Its strategic use of memory and responsive exploration is based on selected concepts that cut across the fields of artificial intelligence and operations research.

Global Optimization list (TL) is given by

$$\mathrm{TL} = \{s^{-1} : s = s_i, i > k - t\} \, , \tag{6}$$

where k is the iteration index and s^{-1} is the inverse of the move s; i.e., $s^{-1}(s(x)) = x$. In words, TL is the set of those moves that would undo one of those moves in the t most recent iterations. t is called the Global Optimization tenure.

The use of TL provides the “constrained search” element of the approach, and hence the solution generated depends critically on the composition of TL and the way it is updated. Global Optimization search makes no reference to the condition of local optimality, except implicitly where a local optimum improves on the best solution. A best move, rather than an improving move is chosen at each step. This strategy is embedded in the Optimum function.

4. Related Work on Conceptual Analysis and Term

4.1. Metrology

The domain of knowledge referred to as metrology forms the foundation for the development and use of analysis instruments and analysis processes in the sciences and in engineering (Fig.3).

While metrology has a long tradition of use in, for example, physics and chemistry, it is rarely referred to in the concept analysis literature. A notable exception in the Semantic Reasoning Network literature is NIST (National Institute of Standards and Technology), which investigated “the underlying question of the nature of IT metrology” in 1996, and identified “opportunities to advance IT metrology.” NIST proposed, for instance, “logical relationships between metrology concepts,” consisting of four steps to follow to obtain measured values: defining quantities/attributes, identifying units and scales, determining the primary references, and settling the secondary references. In addition, in 1999, Gray discussed the applicability of metrology, and the necessity of applying it, from the concept analysis point of view: “We are still perhaps on the eve of giant steps in the new century for information technology. We will still need better analysis and more uniformity, precision, and control to achieve these giant steps." Since then, metrology has been used for the design of the COSMIC analysis method, and is also addressed in [2].

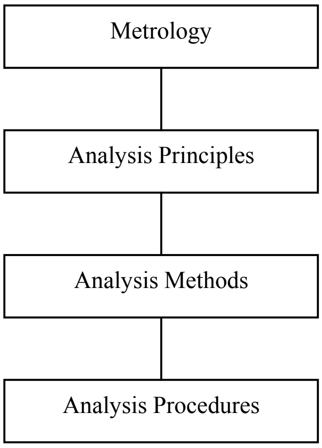

Fig. 3. Analysis foundations.

4.2. Analysis Definitions for the Practical View

While in the Semantic Reasoning Network literature, analysis is often defined as a mapping between two structures, this does not give sufficient information about how to measure in practice. It was pointed out in [2] that it is necessary to move beyond the theoretical definition of the mapping to an operational procedure, as described in the vocabulary of the VIM [1] and modeled with a transition through three levels for a practical view (Fig.4).

An analysis principle forms the scientific basis of analysis. For concept entities (products), the analysis principle involves the model(s) used as a basis to describe the concept that is related to a concept to quantify, and which can be quantified by an analysis method. The idea is that modeling, as a central notion in concept assemble, should be considered at the same level as scientific principles in other sciences and in engineering [2].

4.3. Base Quantity and Analysis Method

To adequately quantify a concept, a analysis method is required, which itself must include a coherent set of definitions and analysis rules, as well as a base unit specific to the analysis method as described in the VIM [1]—Figure 2:

A base unit is "an analysis unit that is adopted by convention for a base quantity" [1]. There is only one base unit for each base quantity.

An analysis method is a generic operational description, i.e. a description of a logical sequence of operations for performing an analysis activity, for moving on from the concept to quantify to the value representing the analysis result [5].

An analysis procedure is a set of operations, described specifically and used in the performance of particular analysis according to a given method [6].

An analysis method should be implemented concretely by some concrete operations achieved through measuring instruments and/or practical operations: selection, counting, calculation, comparison, etc. This description of a analysis according to one or more analysis principles and to a given analysis method is called the analysis procedure, which is more specific, more detailed, and more closely related to the environment and to the measuring instruments (e.g. tools) than the method, which is more generic.

Note that the term metrics is avoided in the definitions above: although it is widely used in Semantic Reasoning Network, its use causes ambiguity, and possibly confusion, by suggesting erroneous analogies; therefore, this term is not used in this text.

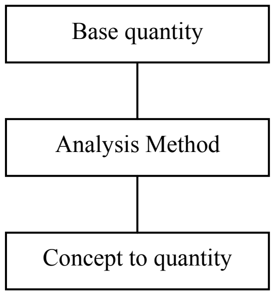

Fig. 4. A base quantity, as defined in the VIM.

4.4. Vocabulary Issues in ISO 15939

In 2002, the ISO documented and adopted a generic model for the analysis process in concept organizations in ISO 15939 (revised in 2007). Specifically, ISO 15939 "identifies the activities and tasks that are necessary to successfully identify, define, select, apply, and improve concept analysis within an overall project or organizational analysis structure" [6]. It also provides "definitions for analysis terms commonly used within the concept assemble," using the VIM as its base, although with some tailoring of the term to facilitate it acceptance within the Semantic Reasoning Network community.

In ISO 15939, a base measure is "a measure defined in terms of an attribute and the method for quantifying it" [5].

To obtain a base measure in practice, an analysis method must be applied to an attribute of an entity (i.e. an object which is itself a model of an object). In the VIM, a analysis method is defined as "a generic description of a logical organization of operations used in a analysis" [5], while in ISO 15939, this definition has been tailored as follows: "a logical sequence of operations, described generically, used in quantifying an attribute with respect to a specified scale" [7]. Both definitions consider "a logical sequence/organization of operations," and from this perspective they are similar.

In ISO 15939, the attribute is a property of an entity. In [2], the entity refers to "the concept to quantify," which should be related to a base unit [1]. The expression "base unit" cannot refer directly to the expression "base measure", since a base unit is a part of an analysis method with rules and conventions designed to obtain a "base quantity." In other words, a base quantity is a combination of a number from the numerical world and a base unit established by convention. For example, by international convention in the SI, the base quantity of length is composed of a number associated with the base unit "meter". To date, there has been little work done to de-fine base units in Semantic Reasoning Network, including base units for the analysis of the quality of concept. In concept analysis, the COSMIC functional size analysis method in ISO 19761 is unique, in the sense that it has explicitly defined its base unit, referred to as "a data movement of a single data group," and its corresponding analysis symbol, "CFP". With this definition, a COSMIC analysis can be expressed as a base quantity in the metrology sense with a number of base units (for example, 15 CFP, 27 CFP, etc.).

In Semantic Reasoning Network, the term time may refer to the number of months representing the base quantity for expressing the concept of effort (as in productivity: concept delivered per work effort unit, measured in person-months), or it may refer to the concept of duration (often measured in calendar-months). The interpretation of the analysis unit "month" will differ, depending on the concept to be represented and measured (e.g. person-months for the concept of human effort, and calendar-months for the concept of project duration).

For the analysis of concept quality, the analysis units of quality concepts, like faults, errors, defects, failures, etc., also need to be explicitly, and uniquely, defined. There are other similarities and differences in the terms used in the ISO 15939 and VIM vocabularies:

1. The concepts of fault, error, and defect:

a) Can be associated with the term "attribute", because they are properties or characteristics of an entity (i.e. concept);

b) But are difficult to quantify, even when there is a consensus on the corresponding base units (1 defect and 1 error, for example): it is not certain whether or not their identification is unique (reproducibility and repeatability problems).

2. The definitions of these "attributes" are generally high-level definitions without corresponding explicit analysis methods.

3. ISO 24765 (System and Semantic Reasoning Network Vocabulary) provides the following definitions:

a) An error is defined as "a human action that produces an incorrect result, such as concept containing a fault";

b) A failure is defined as "an event in which a system or system component does not perform a required function within specified limits";

c) A fault is defined as "a manifestation of an error in concept";

d) A defect is defined as "a problem which, if not corrected, could cause an application either to fail or to produce incorrect results.

Each of these definitions refers to another definition (e.g., an error contains a fault, and a fault is a manifestation of an error), which adds to the difficulty of quantifying concepts like these, and, based on these definitions, of obtaining accurate analysis results.

5. Results and Discussions

We employed Tight method Semantic Reasoning, Random Process character subset and Global Optimization search techniques to classify Node Application Program. We note that the average accuracy of Tight method Semantic Reasoning classifier is 86.4 %, followed by Global Optimization search of 82.45 % and Random Process character subset of 82.2 %. In this context, we employed Back propagation trained neural network to find the importance of different characteristics in Node Application Program. We found that we found that WRRF plays a vital role for classifying the Node Application Program. We excluded the VSRE from dataset and experimented. We observe that the average accuracy of Tight method Semantic Reasoning is 76.2 % and Random Process Character subset is 66.3 % and Global Optimization search is 72.2 %. As Semantic Reasoning Network is a very good classifier to classify Conceptual Analysis type of problems. In this context, the result obtained by Semantic Reasoning classifier is not superior to the results to classify the accuracy of Node Application Program. As Semantic Reasoning Network has not been applied to classify Node Application Program, we employed this approach in the present study.

6. Conclusions

We presented Tight method Semantic Reasoning, Random Process character subset and Global Optimization search to classify Node Application Program. We observed that Tightly method Bays approach predicted better accuracy than Random Process character subset and Global Optimization search. Secondly, Semantic Reasoning belief network is employed first time to classify web services in the present study. Future directions include more exploration of Random Process character subset approach for rule generation and quality of characteristics to decide the Conceptual Analysis of Node Application Program.

References

[1]. H. Lin, R. Juang and D. Lin, Validation of an improved local-based algorithm using GSM measurement data, *IEEE Trans. Mobile Computing*, Vol. 4, September - October 2005, pp. 530–536.

[2]. H. Buddendick, G. Wölfle, S. Burger and P. Wertz, Simulator for performance analysis in UMTS FDD networks with HSDPA, in *Proceedings of the 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Barcelona, Spain, September 2004.

[3]. S. Burger, H. Buddendick, G. Wölfle and P. Wertz, Relationship dependent CDMA verticality in system level simulations, in *Proceedings of the 61st IEEE Vehicular Technology Conference (VTC),* Stockholm, Sweden, May 2005.

[4]. L. Blazevic, J. L. Boudec and S. Giordano, A relationship-based routing method for mobile Ad Hoc networks, *IEEE Trans. Mobile Computing*, March - April 2005, pp. 97–110.

[5]. Ruo Hu, Channel Access Controlling in Wireless Sensor Network using Smart Grid System, *Applied Mathematics & Information Sciences*, No. 6-3S, November 2012, pp. 813-820.

[6]. Ruo Hu, Stability Analysis of Wireless Sensor Network Service via Data Stream Methods, *Applied Mathematics & Information Sciences*, No. 6-3S, November 2012, pp. 793-798.

[7]. Hu Ruo, New Network Access Control Method Using Intelligence Agent Technology, *Applied Mathematics & Information Sciences*, 2, 2013, pp. 44-48.

[8]. A. J. H. Peddemors, M. M. Lankhorst, J. de Heer, Presence, relationship, and instant messaging in a context-aware application framework, in *Proceedings of the 4th International Conference on Mobile Data Management (MDM)*, Melbourne, Australia, 2003.

[9]. CellSpotting.com, http://www.cellspotting.com

[10]. World-tracker.com, http://www.world-tracker. com/