Research on Rulebase Metamodel Construction and Management

Hong Zhu and Shuang Wang

China National Institute of Standardization, Beijing, China
zhuhong@cnis.ac.cn

Abstract. With the implementation of big data strategy in many countries of the world, the data organization and analysis technology has become the main methods for data application and data governance capability. Since data is always organized by specific rules, how to organize information and construct the consolidated rules of data has become an important issue of big data application. However, at present, there is no mature theory of rule and rulebase construction, lack of high-level general rulebase model, and there is no efficient evaluation criteria system for the quality of rule base, which affects the establishment and high-quality management of rulebase. This paper studies how to model and management the rulebase of information resources. First, this paper analyses types of rules and rule relationships according to their functions and semantic properties. Then we propose the requirements for harmonization of rules in the same context. Second, we propose the metamodel of rulebase, and describe the classes and associations of rulebase metamodel. We introduce the internal identifier and composite identifier for rules. Thirdly, we define six measurements for quality control and propose the process of rulebase version control.

Keywords: Rulebase · Rule · Metamodel · Information resource · Standardization

1 Introduction

Many countries regard big data as an important strategic resource, and take big data strategy as a national strategy. Countries expect to use data to promote national economic development, improve social governance, and enhance government service and regulatory capacity. For example, when novel coronavirus outbreak recently, China has achieved data sharing between different systems such as logistics system, transportation system and medical system through big data technology. Since data is always organized by specific rules, how to organize information and construct the consolidated rules of data has become an important issue of big data application. Especially in standardization area, a lot of international standards are written using a rule-based approach, and these rules will also be repeatedly referenced and reuse by other standards. Therefore, it is necessary to study the types of rules and methodologies of rules organization and management for information resources, so as to realize the application of rule-based method and the organization and reasoning of knowledge.
The concept of *rulebase* is defined as the consolidated lists of the rules and associated guidelines which interwork and identified in ISO/IEC 15944 standards [1]. Rulebase is a kind of database, which is also used in knowledge management, data governance, expert system, ontology and so on.

The current research on rules and rulebase mainly includes the rules relationships, formal expression of rules, the method to construct a rulebase in terminology and information areas [2–11]. However, there are few researches on the categories of rules, rulebase metamodel, methodology of establishment and management of rulebase. This paper will proposes the categories of the rules and their relationships in information resources, then propose the metamodel for rulebase, and finally analyze the quality control and version control methodology of rulebase.

## 2 Types of Rules of International Standards

Through the research and analysis of existing rule-based information resources, this paper propose that rules can be classified as requirement, recommendation, permission, possibility and capability types by their functions and semantic properties.

### 2.1 Requirement Type

The rules of requirement type propose objective and verifiable criteria to be met. If the requirements are consistent with the documents, deviation from these standards is not allowed. The preferred verbal forms are “shall/shall not”. In order to effectively distinguish document requirements from external constraints, “must” and “may not” cannot be used instead of “shall” and “shall not”.

### 2.2 Recommendation Type

The type of recommendation for a rule is considered to give appropriate advice, possible options or action plans without mentioning or excluding other options or action plans. The verbal forms of rule as recommendation are “should/should not”.

### 2.3 Permission Type

The type of permission for a rule is to express consent or opportunity to do something. The verbal forms of rule as permission are “may/may not”. “Can”, “might”, “possible” or “impossible” cannot be used instead of “may” and “may not”.

### 2.4 Possibility and Capability Type

The rule of possibility type is to express the expected or imaginable material or causal results. Ability type rules express the ability, fittingness, or quality necessary to accomplish or accomplish a specific task. The verbal forms of rule as possibility and capability are “can/cannot”. Do not use word “may” instead of “can” in the context.
3 Relationships Between Rules

The rules in information resource are organized according to various relationships. The relationship between rules can be classified as equivalence relationship, associated relationship and custom relationship:

**Equivalence relationship.** Rules which having the same content, reusing in different information resources and specify the same requirements are in equivalence relationship.

**Associated relationship.** Rules which have the same subject, and are usually in the same context of the information resource are in the associated relationship.

**Custom relationship.** Other relationships of rules as needed, such as sequential relationship, causal relationship etc.

**Harmonization of rules in the same context**

The most important, in the same context, rules shall meet the following requirements:

— Consistency: there is no rule semantic conflict with others;
— Simplicity: there are no duplicate rules or redundant rules in the same context;
— Integrity: all rules in the same context can constrain all requirements of the subject.

4 Rulebase Metamodel

4.1 Overview of Rulebase Metamodel

Lots of information resources including international standards are using rule-based method. The rulebase is the consolidated lists of the rules and associated guidelines which interwork and identified. The rulebase would be very useful for data organization, reasoning and reusing applications.

The metamodel is the data model that specifies one or more other models, such as data models, process models, ontologies, etc. The metamodel is independent of the specific database technology, application and hardware and software environment. Therefore, the metamodel need not be physically implemented exactly as specified. The purpose of the rulebase metamodel is to describe rulebase, rules and various relations which might hold among rules. The metamodel of rulebase is a conceptual mode, which can provide basic design requirements for specific rulebase construction.

According to the research and applications of rulebase in information technology, this paper propose a rulebase metamodel as illustrated in Fig. 1. The metamodel is specified using the Unified Modelling Language (UML).
4.2 Classes in the Rulebase Metamodel

Rulebase class

*Rulebase* is a class each instance of which models a rulebase, a set of rules with associated guidelines structured according to the relations among them. A minimal *Rulebase* could be simply a sets of rules in the same context which specify the same subject. The most elaborated *Rulebase* is the full consolidated set of rules from all information resources. A *Rulebase* has 5 attributes. A *Rulebase* shall/may participate in *rulebase_rule_membership*, *rulebase_rule_source*, *rulebase_reference*, *rulebase_importation* and *rulebase_relation_inclusion* associations.

Rule class

*Rule* is a class each instance of which models a rule. *Rules* are statement governing conduct, procedure, conditions and relations. A *Rule* has 8 attributes. A *Rule* shall/may participate in *rulebase_rule_membership*, *rulebase_rule_source*, *rule_guideline* and *rule_relationship* associations.

Guideline class

*Guideline* is a class each instance of which models a guideline. *Guidelines* are non-mandatory information leading to a compliant solution for the related requirement. *Guidelines* are always used to provide guidance for the specific *Rule*. A *Guideline* has 3 attributes. A *Guideline* shall participate in *rule_guideline* associations.
Rule_relation class

*Rule_relation* is a class each instance of which models a *relation* among *rules*, a sense in which rules may be connected via constituent relation roles. A *Rule_relation* has 2 attribute. A *Rule_relation* shall/may participate in *rulebase_relation_inclusion* and *rule_relationship* associations.

### 4.3 Associations in the Rulebase Metamodel

**rulebase_reference association**

The *rulebase_reference* association specifies the reference of zero or more referenced *Rulebase* by zero or more referencing *Rulebase*.

**rulebase_importation association**

The *rulebase_importation* association specifies the importation of zero or more imported *Rulebase* by zero or more importing *Rulebase*. Such importation specifies that all *Rules* and associated *Guidelines* included in the imported *Rulebase* are also to be included in the importing *Rulebase*.

**rulebase_rule_membership association**

The *rulebase_rule_membership* association specifies the inclusion of one or more *Rules* in one or more *Rulebases*.

**rulebase_rule_source association**

The *rulebase_rule_source* association specifies the *Rulebase* that is the source of a *Rule*. Each *Rule* shall have exactly one *rulebase* specified as its source.

**rulebase_relation_inclusion association**

The *rulebase_relation_inclusion* association specifies the inclusion of zero or more *Rule_relations* in one or more *Rulebases*.

**rule_guideline association**

The *rule_guideline* association specifies the *Rulebase* that is the source of a *Rule*. Each *Guideline* shall have a *rule_guideline* association with exactly one *Rule*.

**rule_relationship association**

The *rule_relationship* association specifies the use of one or more *Rules* in one or more *Rule_relations*.

### 4.4 Identifiers for Rules

Within these rulebases, each of the rules shall be assigned a unique identifier independent of the information resource which is its source. According to different application purpose, there are two kinds of identifiers: internal identifier and composite identifier.

The internal identifier of rules could be an internal sequence number. Any subsequent, i.e., new, entry to the consolidated set of rules shall be assigned the next available sequential number. Once an internal identifier is given, it is considered always in effect and cannot be reassigned if it has retired.
Composite identifiers are mainly used for external reference aims. The compound identifier of a rule is mainly composed of two data elements with different functions, including: 1) the unique identifier assigned by information resource authority organization; 2) the internal identifier of the rule in the information resource referenced.

5 Management and Maintenance of Rulebase

5.1 Quality Control

At present, there is no research on the quality of rulebase. This paper propose the six important quality control requirements as follows:

1) **Normalization**: Check whether the description methods of rulebase and rules meet the requirements of metamodel of this paper;
2) **Accuracy**: Check whether the statement of rules in the rulebase is consistent with the source;
3) **Completeness**: Check if all relevant rules are included in the rulebase;
4) **Consistency**: Check whether there is conflict between rules;
5) **Conciseness**: Check rules for redundancy or logical loops;
6) **Validity**: Check whether all rules in rulebase are from existing valid information resources.

5.2 Version Control

The most important function of version control is to track changes of configuration item (e.g. rulebases, rules, guidelines). It faithfully recorded information such as when and who changed the contents of the configuration item. Besides, another important function of version control is parallel development. Rulebase development is often a collaborative work of many people. Version control can effectively solve the problem of version synchronization and development communication between different developers, and improve the efficiency of collaborative development.

This paper propose the basic process of version control as follows:

1) **Create configuration item.** According to the configuration management plan, project members create configuration items within the scope of their tasks in the configuration library. At this time, the status of the configuration item is draft, and its version number format is initialized.
2) **Modify configuration item with draft status.** Using the check in/check out function of the configuration management system, project members modify the configuration items in the “draft” state and the version number format.
3) **Technology approval.** Technical approval passes.
4) **Official release.** After the configuration item passes the technical approval, Configuration item status changes from draft to release, and the version number format is changed.
5) **Revise.** The modification of configuration items in the status of “official release” must be carried out in accordance with the version control process.
6 Conclusions

Base on the existing research theory of rule and rulebase, this paper studies how to model and management the rulebase of information resources. First, this paper analyses types of rules and rule relationships according to their functions and semantic properties. Then we propose the requirements for harmonization of rules in the same context. Second, we propose the metamodel of rulebase, and describe the classes and associations of rulebase metamodel. We introduce the internal identifier and composite identifier for rules. Thirdly, we define six measurements for quality control and propose the process of rulebase version control.

Acknowledgement. This work was supported by the Dean Fund Project of China National Institute of Standardization (Grant No. 532020Y-7495) and Key Technical Research on the Knowledge Service Platform for Information Classification and Coding Standards(No.532018Y-5953).

References

1. ISO Reference - ISO/IEC 15944–1:2011 Information technology — Business operational view — Part 1: Operational aspects of open-edi for implementation
2. Choi, L.: Packet classification method through hierarchical rulebase partitioning: US (2009)
3. Azuraliza Abu, B., Zuriyah, K., Salwani, A.: Predictive models for dengue outbreak using multiple rulebase classifiers (2011)
4. Ouchiyama, H., Huang, R., Ma, J.: An evolutionary rulebase based multi-agents system. In: Emergent Intelligence of Networked Agents 2007. pp. 203–215. Springer, Berlin, Heidelberg (2007)
5. Jörg, V.: Rulebase construction: fuzzy rulebase systems and gridded data problems[M]. In: Artificial Intelligent Methods for Handling Spatial Data (2019)
6. Bin, Z., Shenglong, D., Jingcheng, W.: Getting complete rulebase from recorded data[C]. In: International Conference in Fuzzy Logic & Technology. DBLP (2009)
7. Barratt, M.D., Basketter, D.A., Chamberlain, M., et al.: Development of an expert system rulebase for identifying contact allergens[J]. Toxicol. Vitro 8(4), 837–839 (1994)
8. Dittrich, K.R, Gatziu, S., Geppert, A.:The Active Database management system manifesto: a rulebase of ADBMS Features[C]. In: Rules in Database Systems, Second International Workshop, Rids 1995, Glyfada, Athens, Greece, September. ACM (1995)
9. Başçıftçi, F., Avuçu, E.: An expert system design to diagnose cancer by using a new method reduced rule base[J]. Comput. Methods Programs Biomed. 157, 113–120 (2018)
10. Li, W.: Research and application of planning inspection and mapping system based on rule base[J]. urban geotechnical investigation & surveying (2018)
11. Asad, M.M., Hassan, R., Sherwani, B., et al.: Knowledge base and rule base safety management systems development strategies and applications for petrochemical industries: a brief outline[j]. Iop Conf. 530, 012049 (2019)