Application of Semantic Technology in Rational Use of Antibacterial Agents

Yu Siwei
Zhongnan Hospital
WuHan University
WuHan, P.R.China 430072
manfisy@163.com

Xu Lei
School of Information Management
WuHan University
WuHan, P.R.China 430072
xlcupid@whu.edu.cn

Huang Zhisheng
Department of Computer Science
Vrije Univerisity of Amsterdam
Amsterdam, the Netherlands 1081hv
huang.zhisheng.nl@gmail.com

***Abstract*—Rational use of antibiotics is a very urgent problem to be addressed for the management of the current Chinese hospital system. However, it involves complex knowledge analysis and management process. Introducing semantic technology for the rational use of antibiotics provides a new technical means for automatic and semi-automatic supervision. In this paper we analyze the medical semantic data organization form, propose a semantic-based system structure of rational use of antibiotics. The implementation of this platform verifies the feasibility of semantic technology used in this field.**

Keywords—semantic technology, antibacterial, Medication guideline; Rational use of drugs; semantic rules, reasoning

I. Introduction

Rational use of antibiotics is an urgent problem to be addressed for the management of the current Chinese hospital system. It is not enough to monitor the rational use of antimicrobial drugs effectively relying on manual intervention only. In many process, we could use knowledge organization methods to bring the monitoring knowledge into the medical information system management process. Automatic monitoring of the rational use of antimicrobial drugs involves complex process of knowledge analysis and management. Semantic technology introducing into the automatic monitoring process for the rational use of drugs provides a new and effective mean.

All kinds of cyber sources processing and analysis technology and other techniques related to semantic web can be referred as semantic technology. RDF/OWL is the core of semantic technology and the main task of semantic technology is to build a knowledge model in a specific domain effectively. The core characteristics of semantic technology are its effective expression and organization ability of knowledge as well as the logical reasoning ability. Semantic technology can be introduced to the rational use of antimicrobial drugs. In this process, medical data of patients can be expressed and organized as semantic formats, and the medication guideline of antibiotics can be extracted as semantic rules. We use rule-based reasoning over massive medical records to monitor the process of drugs use and prescribing. It can avoid the occurrence of the abuse of antibiotic drugs effectively and reduce the frequency of occurrence of medical accidents and disputes.

In order to solve the difficulties during the process of antibacterial drugs use currently, we used APDG[1] to generate patients data and analyzed the rules in medication guideline. We have developed a Semantic Enabled System for Rational Use of Antibiotics(SeSRUA) and a series of related auxiliary tools such as SToGRUA(Semantic Transformer of Guidelines for Rational Use of Antibiotics) and so on.

II. Patient Data Collection and Semantic Storage

In the process of medical information system research and development, lots of patient data is need to carry out system testing and performance evaluation and other processing. However, patient data is strictly controlled in management, because it is limited by law due to data privacy protection and other reasons. The patient data are generally not public, and even in a hospital, the real patient data use is strictly limited. When we need to take these systems or tools with specific patient data for demonstrations, this problem will become more serious.

The Advanced Patient Data Generator(APDG) provides an effective solution for these situation. APDG system can generate all kinds of patient data automatically according to scientific evidence of a specific disease. Scientific evidence related to this disease includes medical literature (such as PubMed) and statistical data from cyber medical source records about the disease, as well as various correlation data between patient data variables. This makes the generated patient data looks like the real patient data and it avoids the restrictions of real patient data use.

We have generated ten thousand chronic bronchitis medical records in Hubei area using APDG as basic data of SeSRUA for data testing and demonstrations. It achieved our requirements for system demo and meets the purpose of patient privacy data protection. The generated patient data are in the form of triples and it is managed by the platform of Larkc[2] which is a massive semantic data reasoning platform.

III. RULE ANALYSIS OF GUIDELINE

Doctors mainly refer to expressions in the Clinical Application Guideline on the Antimicrobials (hereinafter referred to as the guideline). In order to make computer understand the contents in guideline and intelligently judge whether the use of antibiotics is reasonable or not, this paper classified the guideline knowledge representation into three different levels, namely "abstract rules", "specific rules" and “semantic rules” knowledge representation.

A. Abstract Rules

The guideline provides information to guide us. When make guidelines, the guide-makers made them “practitioners-centered”, therefore, they are written in natural language, which cannot be understood by computer. For this, the related rules in natural language in guideline required to be extracted and form-converted.

Rules description in guideline can be preliminarily translated into ”if ... then ...” form, which can be understood by computer and this is called “abstract rules”. For example, “according to the patient's symptoms, signs, and blood, urine and other laboratory examination results, the patient was preliminarily diagnosed as bacterial infections and through the the pathogeny examine the patient was preliminarily diagnosed as bacterial infections with indications for use of antimicrobial drugs”. First, make sure that the content is directive so that it can be convert into “if ... then ...” form. The converted abstract rules are as follows: According to the patient's symptoms, signs, and blood, urine and other laboratory examination results, the patient was preliminarily diagnosed as suspected bacterial infections. And then do as follows:

- **If** the patient has blood laboratory test results and has been preliminary diagnosed as bacterial infections and pathogenic diagnosed as bacterial infections, **then** he indeed has indication to use antibiotics;
- **If** the patient has urine laboratory test results and has been preliminary diagnosed as bacterial infections and pathogenic diagnosed as bacterial infections, **then** he indeed has indication to use antibiotics.

B. Specific Rules

Specific rules are relative to abstract rules. It means to make entailed knowledge of abstract rules in the guideline, like hyponymy relations and cause-effect relationship between the sentences, more concrete. The extracted abstract rule contains three different rules:

- Rule 1. what should not do;
- Rule 2. how to do;
- Rule 3. do as the doctor asked.

It is obvious that the first two rules are instructive, while the third one is not so instructive, namely, specific reference to the index value is not clearly stated in the rule. Participation of experts in the field was then needed to make the rule more concrete and computer understandable. The specific processed rules are called specific rules.

With the participation of medical experts, based on the index values on hospital inspection reports and “and” or ”or” relationship between the values, the abstract rules then can be specified, realizing the converting from abstract rule to specific rule. Specific rules is basic for the description of logical inference relationship in computer programming.

C. Semantic Rules

Semantic rule is kind of rule conversion that carried out on the basis of specific rule, It means change the description rule of nature language into semantic technology-based description rule which can be handle by computer directly.

IV. SYSTEM DESIGN

SeSRUA, a semantic technology-based medical information management system, is built on massive semantic data processing platform ----LarKC. OWLIM is used to be the basic data layer of LarKC. The platform has a pluggable architecture in which it is possible to exploit techniques and heuristics from diverse areas such as databases, machine learning, cognitive science, the Semantic Web, and others. LarKC provides a number of pluggable components: retrieval, abstraction, selection, reasoning and deciding. In LarKC, massive, distributed and necessarily incomplete reasoning is performed over web-scale knowledge sources.

The plug-ins in LarKC can be used to combine the processing workflow for different application requirements, and this can build a SPARQL server for different applications (SPARQL endpoint). Users can use web browser to visit the data in JSON form, which return from the SPARQL server. And the user interface of SeSRUA then transforms these JSON data into corresponding visual data and displays them in a user-friendly interface. Therefore, any SeSRUA users will be able to use it even if he/she knows nothing about semantic technology.

The system architecture of SeSRUA is shown below:

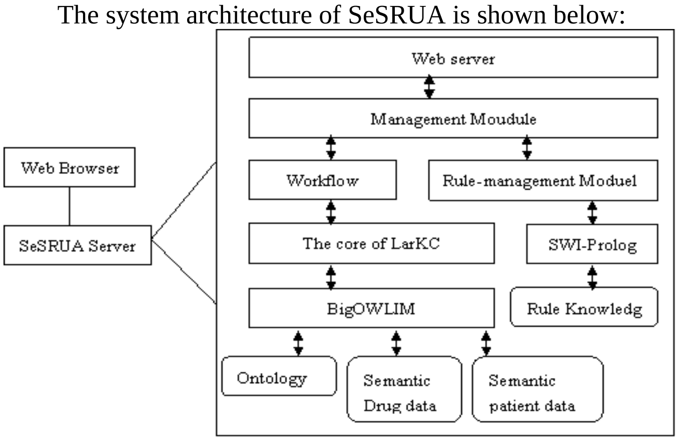

Fig. 1. System Architecture of SeSRUA

The logic programming language Prolog[3] is a rule-based language for knowledge represenation. It is also convenient to be used to formalize the dynamic workflow and to realize the automatic monitoring and dynamic management for the rational use of antimicrobial agents. In SeSRUA, we use Prolog to convert some knowledge in the guideline for rational use of antibiotics into the semantic data.

The following figure 2 shows the graphic user interface of SeSRUA. The figure shows the result of monitoring over a patient on her routine blood test.

Fig. 2. SeSRUA User Interface and Monitoring Result

The SeSRUA system is expected to make full use of the hospital information, further reduce the clinicians' professional burden management of antimicrobials, improve the hospital's pharmacy information management, reduce the incidence of adverse reaction and medical disputes, and ultimately benefit patients.

V. RULE TRANSFORMATION

SToGRUA is a SeSRUA based-system design for semantic data integration. The semantic rules generated by SToGRUA will be used in the monitoring system of SeSRUA.

A. *From Abstract Rule to Specific Rule*

Abstract rules in the guideline contains many instructional rules for using antibiotics rationally and the rule expression form is "if ... then ...". And this form has to be further transformed in the following two steps:

- In the abstract rules, some are specified, some are not. If we extract the indication abstract rules in the guideline, the abstract rules will be specified and do not need to be transformed. We need the participation of medical experts for they can turn the abstract medical test results into specific medical test values. And based on corresponding values on the medical inspection reports, we can then transform the abstract rules into the specific format of "if ... then ..."
- Summarizing the fuzzy words used in the abstract rules expression, through which determine the fuzzy operator subset, then transform the fuzzy words used in the abstract rules expression into fuzzy words existing in the fuzzy operator subset. The rules format after specific conversion is the basis of computer conducting conversion into semantic rules

B. *From the Specific Rule to Semantic Rule*

1) Semantic design converted from specific rule: The specific rule for the guideline contains the knowledge that can be used directly in detecting the data on the electronic medical records. And a considerable part of the knowledge involves indications and contraindications in the antibacterial instructions. For instance, "If infection caused by hemolytic streptococcus, then penicillin should be adopted" and "If infection caused by Streptococcus pneumoniae, then penicillin should be adopted" and so on so forth. However, in existing commonly used semantic medicine ontology (e.g. Drugbank[4]), although the entire ontology adopt the regular form of semantic triples, the description of a single drug's indications normally use a specific predicate. For example, put many indication descriptions right after the word "indication" and then use nature language to describe it, see as the fellowing triples:

<http://www4.wiwiss.fu-berlin.de/drugbank/resource/drugs/DB00007>
<http://www4.wiwiss.fu-berlin.de/drugbank/resource/drugbank/indication>
"For treatment of prostate cancer, endometriosis, uterine fibroids and premature puberty".

Obviously, it is about indications. As the corresponding information hidden in the natural language, indications can not be used directly in the automatic control system for rational antibiotics use. We have to transform them into finer semantic description. We introduced a particular predicate---- "indications" to represent the collection of all the indications. And all the indications belong to this collection.

Meanwhile, for each disease description involves its causes description, we introduced the cause predicate Caused By and the additional parameters predicate Modifier to cover various conditions for each Indication. Therefore, a complete indication semantic rules format should be similar to the followings:

<http://whu.edu.cn/medicine#121216D6369005>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://whu.edu.cn/medicine#DrugBook>.
#The above triple description set an internal concept identifies for certain drugs (e.g. penicillin)

<http://whu.edu.cn/medicine#121216D6369005>
<http://wasp.cs.vu.nl/apdg#Concept>
<http://www.ihtsdo.org/SCT_6369005>.
#The above triples map the internal concept identifies of certain drugs to the corresponding concept in the international general medical ontology SNOMED[5]

<http://whu.edu.cn/medicine#121216D6369005>
<http://whu.edu.cn/medicine#Indications>
<http://whu.edu.cn/medicine#id_121216g6369005>.
#The above triples defines a collection of indications

<http://whu.edu.cn/medicine#id_121216g6369005>
<http://whu.edu.cn/medicine#Indication>
<http:/whu.edu.cn/medicine#id_121216g6369005_1>.
#The above triples defines the first indication description in the corresponding indication collection

<http://whu.edu.cn/medicine#id_121216g6369005_1>

<http://whu.edu.cn/medicine#Symptom> "感染".
#The above triples defines the disease (or symptoms) corresponding to the indications

<http://whu.edu.cn/medicine#id_121216g6369005_1>
<http://whu.edu.cn/medicine#CausedBy> "溶血性链球菌".
#The above triples defines the cause of the indication

<http://whu.edu.cn/medicine#id_121216g6369005_1>
<http://whu.edu.cn/medicine#Modifier> "indication".
#The above triples defines the additional parameter description of the indications

<http://whu.edu.cn/medicine#id_121216g6369005>
<http://whu.edu.cn/medicine#Indication>
<http://whu.edu.cn/medicine#id_121216g6369005_2>.
#The above triples defines the second indication description in the corresponding indication collection, and as for the symptoms and causes of the disease, similar approach will be adopted.

In the system, we adopted the description rule: logic programming language Prolog-based rule and logic rule of SPARQL sever interaction.

2) Realization of specific rule's semantic transformation:

a) Prolog----logic programming language: The specific rules have a unified format "If ... then ...". It is a fixed format of "premise and conclusion" and natural language liked. Hence, they can be automatically transformed into semantic rules by computer program without syntactic parsing by using the complicate natural language processing system. A natural transformation device is realized through logic programming language Prolog.

b) SToGRUA: Using Prolog, we created the SToGRUA（Semantic Transformer of Guidelines for Rational Use of Antibiotics), and the relationship between SeSRUA and SToGRUA is shown in Fig.3.

First, we load the specified guide knowledge as a text file into the Prolog system, then it will be parsed by the SToGRUA system with DCG(Definite Clause Grammar). And finally, a corresponding semantic triples will be got. The main process is just as the following logic program:

```
start :- initial_id_number(N),
    working_on_drug(Drug),
    working_on_ontology(Ontology),
    file_name(File),
    open(File, read, In),
    output_file_init(Drug, Ontology, Out),
    set_stream(In, encoding(utf8)),
    rules_processing(In, Out, N, M),
    close(In),
    close(Out),
    format('~w rules have been processed.~n', [M]),
    true.
```

The predicate rules_processing is designed as follows. It uses the following rules to call the DCG rule corresponding to "cause a result, the state , and the drug name", and by write_ntriple to generate its corresponding semantic triples .

```
rule_processing(Line, Out, N):-
    phrase(rule(Cause, Result, Status, Drug), Line, Rest),
    atom_codes(Cause1, Cause),
    atom_codes(Result1, Result),
    atom_codes(Drug1, Drug),
    atom_codes(Rest1, Rest),
    write_ntriples(Out, N, Cause1, Result1, Status, Drug1,
Rest1),
    true.
```

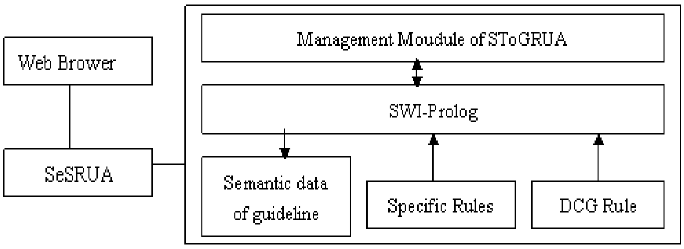

Fig. 3. The Relationship between SeSRUA and SToGRUA

c) DCG rule parsing in SToGRUA system: The DCG parsing rules we made is as fellows: the corresponding parameters will be got by matching condition and result in specific rules; and then we get the causes and symptoms from condition and get status and drug name parameter from result.

```
rule(Cause, Result, Status, Drug) -->
    condition(Cause, Result),
    separators,
    conclusion(Status, Drug).
```

We use DCG rules like the following one to parse conditions of different specific rules:

```
condition(Cause, Result) -->
    causal_condition_header,
    cause(Cause),
    lead_to,
    result(Result).
```

We use DCG rules like the following one to parse results of different specific rules:

```
conclusion(indication, Drug) -->
    conclusion_header,
    indication_operator,
    drug(Drug).
```

VI. CONCLUTION

In this paper, we used semantic technology to monitor rational use of antimicrobial drugs. APDG is used to generate patient data to avoid the limitation of the real data use and the types of rule from guideline were analyzed. We have designed the system framework of the rational use of antimicrobial agents of semantic technology system and introduced the basic technology of module in it. Then a systematic approach to transform the guidelines for rational use of antibiotics in natural language into corresponding semantic data has been elaborated.

The generated patient data and rules are semantic data that can be used in monitoring system of rational use of antibiotics.

The development of the semantic data processing platform makes it possible to judge and monitor the reasonability of antibiotics use in terms of various patient data in electronic medical records. Hence, processing and managing the knowledge of rational antibiotic use automatically and semi-automatically on the basis of semantic technology realized.

REFERENCES

[1] G Huang Z,van Harmelen F,den Teije A,et al.Knowledge-based patient data generation[R].Vrije university,2013.

[2] Fensel D, van Harmelen F, Andersson B, et al. Towards LarKC: a platform for web-scale reasoning[C]//Semantic Computing, 2008 IEEE International Conference on. IEEE, 2008: 524-529.

[3] Jan Wielemaker, Zhisheng Huang, and Lourens van der Meij, SWI-Prolog and the Web, Journal of Theory and Practice of Logic Programming, 8(1):1-30, 2008.

[4] C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, Djoumbou Y, Eisner R, Guo AC, Wishart DS， DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Knox. Nucleic Acids Res. 2011 Jan;39(Database issue):D1035-41. PMID: 21059682.

[5] ABOUT SNOMED CT.[2011-7-5].http://www.ihtsdo.org/index.php?id=snomed-ct0.