Abstract—Checking the consistency of knowledge items between textbook and clinical practice is very important for traditional Chinese medicine (TCM). According to the textbook of internal medicine of TCM, rheumatoid arthritis (RA) is a disease with four main syndromes (also called pattern) and Feng-Shi-Re-Bi is a minor one and the associated literature data are not rich. Bisimulation is an equivalence relationship/method in formal methods which can be used to compare the consistency of system's description and its behaviors. In TCM, knowledge items from textbook can be taken as system's description, and knowledge items from clinical practice can be taken as the behaviors of the system. For RA's Feng-Shi-Re-Bi syndrome, employing bisimulation method, the consistency of knowledge items between textbook and clinical practice have been checked. As a result, on Feng-Shi-Re-Bi syndrome, most knowledge items in textbook can be simulated by the clinical practice e.g., syndrome, symptom, herbal formula, herbal medicine. What's more, in clinical practice, there are also variations based on basic rules of traditional Chinese medicine. In brief, bisimulation is a proper method for consistency checking between TCM textbook and clinical practice.

Index Terms—consistency checking, text mining, bisimulation, rheumatoid arthritis, traditional Chinese medicine.

I. INTRODUCTION

Checking the consistency of knowledge items between textbook and clinical practice is very important for traditional Chinese medicine (TCM). In the clinical practice of TCM, facing different patients of rheumatoid arthritis, different doctors may prescribe different Chinese herbal formulae according to their different therapeutic methods and understanding of patient’s private healthy constitution [1], [2]. Although effective in many clinical cases, however, TCM is very flexible and hard to master [1], [3] according to the nature of its personalized medical service. Under this situation, checking the consistency of knowledge items between textbook and clinical practice is very important for both clinical reference and medical research [2].

Rheumatoid arthritis (RA) is a chronic, systemic inflammatory disorder that may affect many tissues and organs, but principally attacks flexible (synovial) joints [4]–[7]. About 1% of the world’s population has rheumatoid arthritis, women three times more often than men. With several thousand years of clinical practice, some effective herbal formulae have been constructed for its treatments [8]–[10]. According to the textbook of internal medicine of TCM [1], RA is a disease with four syndromes (also called pattern) according to the syndrome classification which is based on the different sets of subjective symptoms of TCM. These syndromes are Feng-Han-Shi-Bi, Feng-Shi-Re-Bi, Tan-Yu-Bi-Zu, and Gan-Shen-Yin-Xu. Among them, syndrome Feng-Shi-Re-Bi is a minor one in RA which is believed to be caused by factors of Feng(wind), Shi (dampness), and Re (heat) [11].

In checking the consistency of knowledge items of syndrome Feng-Shi-Re-Bi in RA between textbook and clinical practice, formal method bisimulation is employed [2], [12]–[14]. Bisimulation is an equivalence relationship/method informal methods which can be used to compare the consistency of system’s description and its behaviors [15], [16], [21]. In this study, knowledge items from textbook can be taken as system’s description, and knowledge items from clinical practice which are extracted from clinical literatures with text mining technique [18] of data slicing algorithm [19], [20] can be taken as behaviors of the system. Focusing on RA’s syndrome of Feng-Shi-Re-Bi, employing bisimulation method, we checked the consistency between textbook and clinical practice.

As a result, most of the items in textbook on syndrome
Feng-Shi-Re-Bi can be bisimulated by the clinical practice. However, in clinical practice, there are also variations based on basic rules of traditional Chinese medicine, e.g., syndrome, symptom, herbal formula, and herbal medicine. What’s more, these variations are all based on the basic rules of traditional Chinese medicine according to specified clinical situations.

In brief, bisimulation is a proper method to check the consistency between TCM textbook and clinical practices. By checking the consistency of knowledge items on syndrome Feng-Shi-Re-Bi between textbook and clinical practices, some new knowledge items can be extracted and demonstrated which might be a good reference for both clinical practice and medical research.

II. MATERIAL AND METHODS

A. Data Collection

The knowledge items of textbook is taken from the Internal Medicine of traditional Chinese medicine which is the kernel classic textbook of traditional Chinese medicine [1]. Focused on RA’s syndrome Feng-Shi-Re-Bi, there are also two syndromes associated with it, i.e., Shi-Re-Zu-Luo and Shi-Re-Yu-Zu. Both of them are similar with syndrome Feng-Shi-Re-Bi.

The knowledge items of clinical practice are exacted from literature database of SinoMed (http://sinomed.cintcm.ac.cn/index.jsp). The dataset were downloaded from SinoMed with query term of “rheumatoid arthritis” on November 22, 2012. This dataset contains 15,900 records of literatures on clinical practices or theoretical research on rheumatoid arthritis. In this dataset, each record/paper is tagged with an unique ID. These records contain the title, keywords, and abstract of published papers [17], [18].

B. Data Mining

Data slicing algorithm is a powerful tool in text mining associated rules in traditional Chinese medicine [17], [19], [20]. There are several steps in this algorithm. The first step is to extract information from dataset which is based on the keyword lists. In this step, keyword lists of syndrome, symptom, herbal formula, and Chinese herbal medicine are extracted and tagged with unique literature ID. The second step is to construct networks of syndrome, symptom, herbal formula, and Chinese herbal medicine. As for the network construction, syndrome, symptom, herbal formula, and herbal medicine are connected to each other with the same literature ID tagged.

C. Network Construction

1) Network of Co-occurrent: To construct the network of syndrome-symptom on Feng-Shi-Re-Bi syndrome which can be demonstrated in Step 1 and Step 2 in Fig. 1, the first step is to find literatures containing Feng-Shi-Re-Bi syndrome. Then, within these literatures, another process for filtering symptoms will be executed. The symptoms filtered out will be tagged with literature ID. As each literature ID is associated with Feng-Shi-Re-Bi syndrome naturally, thereafter, the network of syndrome-symptom on Feng-Shi-Re-Bi syndrome can be constructed on the co-occurrence of Feng-Shi-Re-Bi syndrome and symptoms.

2) Network of Rule-based: To construct the network of herbal formula-Chinese herbal medicine which can be demonstrated in Step 3 and Step 4 in Fig. 1, it is based on the rules which are defined by the composition of herbal formulae. Each herbal formula is composed strictly with a set of Chinese herbal medicines. For example, formula Baihu-Jia-Guizhi-Tang, it is composed by Zhumu, Guizhi, Jingmi, Zhiganao, and Shigao. In this paper, the network of herbal formula-Chinese herbal medicine is the mixture of co-occurrence approach and rule-based approach.

D. Bisimulation

1) Definition: In theoretical computer science, a bisimulation is a binary relation between state transition systems, associating systems which behave in the same way in the sense that one system simulates the other and vice-versa. Intuitively two systems are bisimilar if they match each other’s moves. In this sense, each of the systems cannot be distinguished from the other by an observer [14]–[16], [21].

Definition 1: Given a labeled state transition system(S, \( \rightarrow \)), a bisimulation relation is a binary relation R over S (i.e., \( R \subseteq S \times S \)) such that both \( R^{-1} \) and R are simulations. Equivalently, R is a bisimulation if for every pair of elements \( p,q \in S \) with \((p,q) \in R\), \( \forall \alpha \in \Lambda \):

- \( \forall \alpha \in \Lambda \):
  - \( \forall p \in S \): \( p \xrightarrow{\alpha} p' \) implies that there is a \( q' \in S \) such that \( q \xrightarrow{\alpha} q' \) and \((p',q') \in R\);
  - \( \forall q \in S \): \( q \xrightarrow{\alpha} q' \) implies that there is a \( p' \in S \) such that \( p \xrightarrow{\alpha} p' \) and \((p',q') \in R\);

Given two states p and q in S, p is bisimilar to q, written \( p \sim q \), if there is a bisimulation R such that \((p,q) \in R\).

The bisimilarity relation \( \sim \) is an equivalence relation. Furthermore, it is the largest bisimulation relation over a given transition system.

2) Bisimulation process: In the bisimulation process, there are four steps in the consistency checking on knowledge items of textbook and text mining which can be demonstrated in Figure 1:

- **Step 1**: This is the initialization step of bisimulation. In this step, either from textbook, or from text mining, they both started from the root point of “rheumatoid arthritis”;

- **Step 2**: This is the second step of bisimulation. In this step, syndromes of Shi-Re-Zu-Luo and Shi-Re-Yu-Zu which are similar with Feng-Shi-Re-Bi, are mined out from clinical literatures;

- **Step 3**: In this step, the bisimulation is focused on symptoms associated with Feng-Shi-Re-Bi symptom, herbal formula, and herbal medicine. Symptoms in bold font are described in textbook on Feng-Shi-Re-Bi. Symptoms of normal font are those mined out from clinical literatures. Thus, in clinical practice, symptoms varies according to different personal conditions;
• **Step 4**: This step of bisimulation is focused on herbal medicines included in herbal formula listed in step 3. In this step, herbs in the middle part of bold font are those which can be filtered out from both textbook and clinical literatures. On the country, herbs on the left part of normal font are those mined out from clinical literature alone; herbs on the right part of normal font are those filtered out from textbook.

• **Step 5**: The last step of bisimulation is focused on herbal formula’s adjustments. In this step, herbal medicines are also listed in three columns. These herbal medicines are not included in the herbal formula of Baihu-Jia-Guizhi-Tang which is the one described in textbook. However, these medicines are widely and frequently adopted for the treatment of RA’s syndrome of Feng-Shi-Re-Bi. This step demonstrates the nature of TCM’s personal medicine which requires the changes of herbal medicine’s composition with respect to the personalized healthy condition.

### III. RESULTS

By merging the results of textbook and text mining, RA’s bisimulation on Feng-Shi-Re-Bi can be demonstrated in Table I.

From Table I, it is clear that results of text mining clinical literature are more complex than the results of textbook. There are 13 items discovered by text mining while 10 validated by textbook knowledge items. Thus, the simulating rate from textbook to text mining is 10/13 which is 76.9%. This phenomenon is reasonable because knowledge items from textbook are much briefer while the clinical practice is much more complicated.

In order to get a more intuitive and clearly view, the results from textbook and text mining can be demonstrated in five sequential steps which can be demonstrated in Fig. 1.

In Figure 1, the bisimulation process is divided into 5 steps. Step 1 is the bisimulation on syndrome of Feng-Shi-Re-Bi which is explained in textbook, specifically. In this step, syndromes of Shi-Re-Zu-Luo and Shi-Re-Yu-Zu are mined out from clinical literature. Both of them are similar with Feng-Shi-Re-Bi on syndrome elements of Shi (dampness) and Re (heat). Step 2 is the bisimulation on symptoms. In this step, symptoms of bold font appear in the result of both text mining and textbooks. Step 3 demonstrates the bisimulation on herbal formula associated with syndrome of Feng-Shi-Re-Bi. Herbal formula of Baihu-Jia-Guizhi-Tang is the common one. In clinical practice, Si-Miao-Yong-An-Tang and Danggui-Niantong-Tang are also mined out for the treatment of RA’s Feng-Shi-Re-Bi syndrome. Step 4 and step 5 demonstrate the bisimulation of herbal medicines. Herbs of bold font in the middle part are those which can be filtered out from both textbook and clinical literatures. On the country, herbs on the left part of normal font are those mined out from clinical literature alone; herbs on the right part of normal font are those filtered out from textbook.

### IV. CONCLUSION AND DISCUSSION

For the normalization of traditional Chinese medicine, it is very important to check the consistency of the knowledge items between textbook and clinical practices with respect to the flexibility of TCM. Knowledge items of textbook can be extracted by manually reading the validating. However, knowledge items of clinical practice distributed among millions of published literatures which can only be extracted through the technique of text mining. What’s more, the consistency checking can be relied on the simulating rate between these two aspects through the formal method of bisimulation. By checking the consistency between knowledge items from textbook and text mining, the associated results were demonstrated in Table I.

<table>
<thead>
<tr>
<th>Category</th>
<th>Keyword</th>
<th>Text Mining</th>
<th>Textbook</th>
<th>Bisimulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syndrome</td>
<td>Shi-Re-Zu-Luo</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Feng-Shi-Re-Bi</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Shi-Re-Yu-Zu</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Symptom</td>
<td>Ulcer</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Tetanus</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Morning stiffness</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Movement limitation</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Joint swelling</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Red swelling</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Red swollen</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
<tr>
<td></td>
<td>Joint pain</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Press pain</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td>Numb</td>
<td>√</td>
<td>√</td>
<td>×</td>
</tr>
</tbody>
</table>

and Figure 1. By analyzing the consistency, we have the conclusion that clinical practices implies more knowledge items than those existed in textbook. However, for knowledge items, neither from textbook, nor from clinical practice, could cover each other completely in the general sense.

1) **Major syndrome together with similar syndromes can be found for bisimulation**: For RA’s syndrome of Feng-Shi-Re-Bi, it can be simulated by similar syndrome of Shi-Re-Bi-Zu and Shi-Re-Yu-Zu. This can be demonstrated by syndrome in Table I and step 1 in Figure 1.

2) **Most the symptoms on syndrome Feng-Shi-Re-Bi can be bisimulated**: For all the nine symptoms
discovered by text mining, seven can be checked and validated in textbook. For syndrome Feng-Shi-Re-Bi, there are symptoms of morning stiffness, movement limitation, joint swelling red swollen, joint pain, and press pain. However, in clinical literature, three other symptoms are tightly related with these symptoms. They are ulcer, tetanus, and numb. This can be demonstrated by symptom in Table I and step 2 in Figure 1.

3) **Textbook formula can be simulated together with supporting formulae:** For syndrome Feng-Shi-Re-Bi, it is clearly stated in textbook that formula Baihu-Jia-Guizhi-Tang targeting this syndrome. What’s more, by text mining clinical literature, the supporting formulae targeting the co-occurrence syndromes are also mined out. Although they decreased the bisimulation rate, they are indicating the meaningful clues for clinical practice. This can be demonstrated by step 3 in Figure 1.

4) **Kernel herbal medicine can be simulated together with supporting ones:** As herbal formula is composed by herbal medicines, then, based on the analysis on formula, it is clear that Kernel herbal medicine can be simulated together with supporting ones. This can be demonstrated by steps 4 and 5 in Figure 1.

**ACKNOWLEDGMENT**

This work was partially supported by National Science Foundation of China (No. 30825047, 30902003, 30973975, and 81072982). China Postdoctoral Science Foundation (No. 20110940553). Gansu Science Foundation (No.1208RJZA249).

**REFERENCES**


[18] Sam Schmidt, Peter Vuillermin, Bernard Jenner, Yongli Ren, Gang Li, Yi-Ping Phoebe Chen, Mining Medical Data: Bridging the Knowledge Divide, Proceedings of eResearch Australasia, 2008.

