

PAPER • OPEN ACCESS

## A Chinese medicine formula homology algorithm

To cite this article: Ruixiang Wang 2019 *J. Phys.: Conf. Ser.* **1168** 052051

View the [article online](#) for updates and enhancements.

You may also like

- [Convolution-based estimation of organ dose in tube current modulated CT](#)  
Xiaoyu Tian, W Paul Segars, Robert L Dixon et al.
- [A scanner-specific framework for simulating CT images with tube current modulation](#)  
Giavanna Jadick, Ehsan Abadi, Brian Harrawood et al.
- [Charge-multiplicity and collision-energy dependence of  \$p\_{\perp}\$  spectra from  \$p\$ - \$p\$  collisions at the relativistic heavy-ion collider and large hadron collider](#)  
Thomas A Trainor



**ECS**  
The  
Electrochemical  
Society  
Advancing solid state &  
electrochemical science & technology

**DISCOVER**  
how sustainability  
intersects with  
electrochemistry & solid  
state science research

# A Chinese medicine formula homology algorithm

Ruixiang Wang<sup>1,a</sup>

<sup>1</sup>Department of Management, Shandong University of Traditional Chinese Medicine, No. 4655, Daxue Road, Jinan; 250355, China

<sup>a</sup> Corresponding author: xinxiguanliyu@163.com

**Abstract.** Traditional Chinese Medicine (TCM) has a history of application in China for thousands of years. The essence of TCM diagnosis and treatment of diseases is to judge the cause of the disease according to the patient's symptoms combination and TCM theory, and then give the prescription according to the cause. Doctors often give different prescriptions because of their different experience and understandings of the theory of Chinese medicine, which seriously hinders the standardization and industrialization of TCM. The clinical practice of TCM shows that an effective clinical prescription has its origin, and the ultimate source is the ancient classic prescription. According to this point of view, based on the fuzzy theory, the paper proposes a homology algorithm for TCM prescriptions and has been implemented on the python platform. The algorithm is able to find the classic formulas that are most closely related to the clinical formula based on the membership function. Because research on classical prescriptions is the most adequate, the results of this algorithm can help doctors to understand the mechanism of the clinic formula.

## 1. Introduction

Although the science based on experiment and formal logic does not accept the theory of TCM, TCM is still widely used in China. TCM includes prescriptions, acupuncture, massage, fumigation, and cupping. The prescription is the main method used to treat certain diseases for TCM [1][2]. The TCM prescription is two or more herbs on the basis of a certain method of processing decoction, according to a certain proportion of the combination. Once the target of the disease is determined, the drug of modern medicine is determined. Faced with exactly the same kind of disease, different TCM doctors often give different prescriptions. The explanation for this is that the patient's physique is different, the etiology and pathogenesis of this patient are different, and the prescription is of course different. Viewed from the perspective of TCM theory, the above explanation is reasonable. According to the above logic, the prescription should be unique in the face of the same diagnosis based on TCM theory, but in medical practice, different doctors prescribe completely different prescriptions. In fact, the most important way to improve the clinical level of a Chinese medicine doctor is the accumulation of clinical cases. He must have learned one or more prescriptions from the previous doctors at first, and then explored his experience in the process of continually treating the disease. In the Chinese medicine world, it is hard to think of a doctor who will stop learning the ancients' prescriptions. When a TCM doctor faces a patient, he will inevitably think of the classic formulas he has learned.

On April 16, 2018, the Chinese government announced the catalogue of the first batch of classic prescriptions. It involves 37 ancient medical books, spanning 6 dynasties of the Han, Tang, Song, Jin, Yuan and Qing Dynasties, with a total of 100 prescriptions. These 100 prescriptions were exempted from clinical trials. In this paper, a homology algorithm of TCM formula was proposed and implemented systematically. The algorithm temporarily uses these 100 government-approved classic



formulas as similar targets for other prescriptions. In the system, when you give a TCM formula, according to the algorithm, you will get a few classic formulas with the highest similarity to the query formula. At present, there are cases of excessive use of herbs in clinical practice of Chinese medicine. The results of this study can reduce clinical herbs to a minimum without reducing the efficacy. The results of this study can avoid the repetition of experimental research in prescription, or to lead the experimental research of prescription into a controllable research paradigm. Since these 100 prescriptions are only the first list proposed by the Chinese government, as the list continues to expand, the accuracy of system matching will continue to increase.

## 2. Fussy set in TCM

In the classical collection, whether a statement belongs to a certain set is determined, which is in some ways consistent with people's cognition, which means that the parameters of a system are determined. However, this way of thinking cannot truly reflect certain phenomena and concepts in reality. Zadeh pointed out that when the complexity of a system is improving, our ability to accurately describe the problem is declining. Fuzzy theory uses membership degree as a measure of the extent to which a state belongs to a fuzzy concept.

Definition 1: Set  $X$  to be a collection of objects  $x$ , then the fuzzy set  $\tilde{A}$  is a set of ordered pairs.

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\} \quad (1)$$

In definition 1,  $\mu_{\tilde{A}}(x)$  is the membership function maps  $X$  to the membership space  $M$ , and  $\mu_{\tilde{A}}(x)$  is a real number from 0 to 1. The greater  $\mu_{\tilde{A}}(x)$ , the greater the probability that  $x$  belongs to  $\tilde{A}$ . When  $\mu_{\tilde{A}}(x)$  is equal to 0 or 1, for  $x$ , the fuzzy set becomes a classic set[3].

Since Zadeh proposed the concept of "fuzzy set", the idea of fuzzy mathematics has been widely used in different fields including TCM. The application of fuzzy theory in the field of TCM is mainly reflected in three aspects. The first aspect is reflected in the judgment of symptoms. For example, Chinese medicine has a "red flush" symptom that is of great value in the diagnosis. However, different doctors have different understandings of the symptom. In the face of the same patient, some doctors think that this symptom exists, others will hold the opposite opinion. Some Chinese medicine researchers have proposed using fuzzy mathematics to deal with this uncertainty in the perception of symptoms[4][5]. The second aspect is reflected in the judgment of the contribution of different symptoms to the patient's syndrome. The first expert system of Chinese medicine, Youbo Guan diagnostic system is based on fuzzy mathematics. The degree of membership is defined as a division, the total of all typical symptoms probabilities of a syndrome as a denominator, the total of the symptoms probabilities of the current patient as a numerator. If the threshold is specified, the syndrome its membership degree exceeding the threshold is the result of the patient's TCM diagnosis[6]. The third aspect is to use the theory of fuzzy sets to determine which components of a TCM formula are the main components, which components are responsible for increasing the efficacy of the main components, and which components are responsible for reducing the side effects of other components based on targets of diseases[7].

Some concepts in the diagnosis and treatment of TCM, such as "symptoms", "syndromes" and "prescriptions", have features of fuzzy sets, which are the main reason why fuzzy mathematics can be applied to Chinese medicine. In this paper, we use the "classical formula" as a fuzzy concept, we use the membership parameter in fuzzy math to judge the closeness of a prescription and a classic prescription, and then judge the source of the current prescription according to the comparison of different membership degrees.

### 3. Concepts and algorithm

TCM prescriptions can be seen as a set of herbs. In this paper, each classical recipe is regarded as a fuzzy set. Whether a prescription is subordinate to the fuzzy set is determined by the membership function. We determine which classical formulas are closely related to the prescriptions we are querying based on membership.

#### 3.1. Basic definitions

Definition 1: If  $P$  is a prescription,  $a_i$  is an herb that constitutes  $P$ , then  $P$  can be regarded as a set containing all  $a_i$ .

$$P = \bigcup_{i=1}^m \{a_i\} \quad (2)$$

Definition 2: If  $C$  is a classic formula,  $b_j$  is an herb that constitutes  $C$ , then  $C$  can be regarded as a collection containing all  $b_j$ .

$$C = \bigcup_{j=1}^n \{b_j\} \quad (3)$$

Definition 3: If  $|C|$  is defined as the number of herbs of classic formula  $C$ ,  $|P|$  is defined as the number of herbs in current prescription  $P$ , then  $\phi(P, C)$  is defined as the distance of  $P$  from classic prescription  $C$ .

$$\phi(P, C) = \frac{|P| + |C| - |P \cup C|}{|C|} \quad (4)$$

Definition 4:  $N$  classic recipes are defined as  $n$  fuzzy sets, the membership function of a formula belonging to a fuzzy set is the distance of it from classic prescription.

Definition 5: When the threshold is set to  $\lambda$ , then a classical prescriptions whose distance from the current prescription  $P$  is greater than or equal to  $\lambda$  is called the targeted classical prescription of  $P$ .

#### 3.2. Targeting classical formulas search algorithm

(1) Set all the classic prescriptions that have a targeted relationship with a specific prescription  $P$  to form a set  $A$ . The initial value of  $A$  is an empty set.

(2) Setting  $\Omega$  is a collection that contains all the classic formulas and  $\lambda$  is the threshold.

(3) WHILE( $\Omega \neq \{\}$ ) {

(4) IF( $a \in \Omega$ ) { IF( $\phi(P, a) \geq \lambda$ )

(5)  $A = A \cup a$ ;

(6)  $\Omega = \Omega - a$ ; }

(7)  $A$  is the collection of all classical prescriptions with a targeted relationship with  $P$ , the greater the membership value, the closer the targeting relationship with  $P$ .

### 4. Algorithm implementation and discussion

The essence of the system is a query based on the classic recipe database. The user submits the prescription. The system gives all the classic prescriptions that meet the conditions according to the algorithm and the threshold submitted by the user.

#### 4.1. Database source

As mentioned in the first part, we temporarily consider the 100 classic prescriptions announced by the Chinese government as all classic prescriptions. The data model in the classic formula database includes four attributes, namely the classic formula name, the source of the prescription, the composition of the prescription and the method of preparation of the prescription. The system is an open system. Any truly effective classic formula can be added to the database at any time.

#### 4.2. Algorithm development tools

This research uses PYTHON as a system development tool, using MYSQL as a database development tool, and using DJANGO's MVC framework as a web-side display platform.

#### 4.3. Problems and solutions

The main problem is the storage problem of a large number of herbs. There is at least tens of thousands of herbs in the world, but there are only more than one hundred kinds of herbs involved in the classic prescriptions. The herb in the query formula may be present in the classic formula, or all the classic formulas may not have this herb. To solve this problem, we set up an herb database, first add the herbs involved in the classic prescriptions. Attributes of herb include herb Chinese name, herb pinyin, herb English name, herb composition and other information. We have set up the automatic addition function of herb. When an herb in a query formula does not exist in the database, the herb automatically enters the herb database with the information of the herb Chinese name, and other information of the herb may be vacant, or may be added later according to need.

#### 4.4. An application example

The research object in [8] is a prescription created by the author, called "Ziyin Qingwei Soup". The prescription is used to treat oral ulcers. The prescription is the result of continuous exploration by the author in many years of clinical practice and has good clinical efficacy. The authors performed very detailed animal experiments, and after rigorous statistical analysis, it is proved that the prescription has a scientific basis for the efficacy of oral ulcers. The prescription contains ten herbs, namely, "Rehmannia", "Radix Ophiopogonis", "Radix Scrophulariae", "Moutan Bark", "Chinese Herbaceous Peony", "Plaster", "Gardenia Jasminoides Ellis", "Campanulaceae", "Licorice" and "Mint".

When this formula was put into the system, we found that the most closely related to it was a classic formula called "Shaoyao Gancao Soup" with a degree of membership of 1. "Shaoyao Gancao Soup" contains two herbs, "Licorice" and "Chinese Herbaceous Peony". After the enquiry through academic journals, we found a few records of the treatment of peptic ulcers such as oral ulcers with "Shaoyao Gancao Soup". These articles show that "Shaoyao Gancao Soup" has a positive effect on the treatment of peptic ulcer.

The above query results show that these two prescriptions are effective in the treatment of oral ulcers. Obviously, the herbs of "Ziyin Qingwei Soup" is more than "Shaoyao Gancao Soup". The next question is whether the effect of "Ziyin Qingwei Soup" is much greater than "Shaoyao Gancao Soup". We can design an animal experiment in which animals that have been modeled by oral ulcers are divided into two groups, one using "Shaoyao Gancao Soup" and one using "Ziyin Qingwei Soup", and then statistically analyzing whether there is a significant difference in the efficacy of the two prescriptions for oral ulcers. If the results of the experiment are that there are no significant differences between the two prescriptions in the treatment of oral ulcers, then at least two aspects of value, on the one hand, found the most simplified form of prescription for the treatment of oral ulcers, avoiding waste of herbs and further experiments cost, on the other hand, can promote the standardization and industrialization of traditional Chinese medicine prescriptions. If the experimental results confirm that the efficacy of "Ziyin Qingwei Soup" is significantly better than "Shaoyao Gancao Soup", it means that the remaining eight herbs in "Ziyin Qingwei Soup" enhance the efficacy, providing an idea for the next experiment. If there is a more surprising result, the effect of "Shaoyao Gancao Soup" is obviously stronger than "Ziyin Qingwei Soup", then we naturally propose a logical scientific hypothesis - some components of "Ziyin Qingwei Soup" neutralize some of the active components in "Shaoyao Gancao Soup", which also provides an idea for further experimental research.

#### 4.5. Discussion

Different from modern medicine, different TCM doctors give different prescriptions in the face of the exact same diagnosis. So far, this diversified treatment has been regarded by many Chinese doctors as a feature and advantage of TCM. Some Chinese medicine scholars associate this feature with the

precise treatment of modern medicine. In fact, the diversity of prescriptions and the precise treatment of modern medicine based on omics data are completely different. The reasons for the diversity of prescriptions are various. For example, in ancient time, most herbs were difficult to transport for long distances. Therefore, in the face of herbs with similar functions, the doctors at that time were more willing to choose herbs with lower transportation costs. On the other hand, many doctors are not willing to share their own prescriptions with more people to study, so a certain prescription is often used, inherited and spread in a small scope, which led to the efficacy of different prescriptions is difficult to make a real comparison. In order to facilitate the scientific research, in some Chinese medicine hospitals, this diversified feature of formulas has been gradually abandoned. In these hospitals, the doctors are required to use the same prescription for the same diagnosis based on TCM theory.

The homology concept proposed by this algorithm is different from the homology concept proposed by some TCM researchers based on literary knowledge. The homology of this algorithm is more targeted and practical.

## 5. Conclusion

The standardization of TCM prescriptions is the basis for the modernization of TCM. This paper proposes a homology algorithm for TCM prescriptions, which can trace the source of any prescription based on fuzzy set theory. The goal of this algorithm is to cluster all existing recipes into the classic formulas. Because the modern research of classical formulas is very sufficient, the algorithm can let the owner of the prescription learn the latest and most comprehensive research results of his prescription in the shortest time. The homology algorithm was implemented on the MYSQL, PYTHON, and DJANGO architecture.

## Acknowledgment

This work is partially supported by *Shandong Natural Science Foundation Project (ZR2017LH062)*.

## References

- [1] Chen YC, Lin YH, Hu S, Chen HY. Characteristics of traditional Chinese medicine users and prescription analysis for pediatric atopic dermatitis: a population-based study. *BMC Complement Altern Med.* 16, 1(2016).
- [2] Huang TP, Liu PH, Lien AS, Yang SL, Chang HH, Yen HR. A nationwide population-based study of traditional Chinese medicine usage in children in Taiwan. *Complement Ther Med.* 22, 3(2014).
- [3] H.-J. Zimmermann. Fuzzy set theory. *Wiley Interdisciplinary Reviews.* 3, 2(2010).
- [4] Xu Xiaojuan, Wang Qinguo, Xu Haixu, et al. Study of Fuzzy Modularized Thoughts on Differentiation and Pattern in TCM. *Journal of TCM Univ. of Hunan.* 27, 2(2007).
- [5] Chen Lei. The application of fuzzy mathematics in the study of traditional Chinese medicine. *Journal of TCM of Liaoning.* 33, 6(2006).
- [6] Rongjiang Guo, Birong Ma, Min Xie, et al. TCM computer diagnosis and treatment system and fuzzy mathematics model. *Journal of Computer.* 7, 4(1981).
- [7] LIU Ming, GAO Yue, XIAO Rui, et al. Fuzzy mathematic quantitative law of composing principle in the study of traditional Chinese medicine. *Acta Pharmaceutica Sinica.* 44, 1(2009).
- [8] Li Guoxing, Yang Jianhong, Li Zhifang. Studies on ZiYinQingWeiTang treating recurrent ulcer of mouth. *Tianjin Pharmacy.* 19, 3(2007).