Contents lists available at ScienceDirect

Digital Chinese Medicine

journal homepage: http://www.keaipublishing.com/dcmed

# Research on knowledge reasoning of TCM based on knowledge graphs

GUO Zhiheng[a], LIU Qingping[a*], ZOU Beiji[a, b]

a. School of Informatics, Hunan University of Chinese Medicine, Changsha, Hunan 410208, China

b. School of Computer Science and Engineering, Central South University, Changsha, Hunan 410083, China

ARTICLE INFO

*Article history*
Received 19 October 2022
Accepted 20 November 2022
Available online 25 December 2022

*Keywords*
Traditional Chinese medicine (TCM)
Stroke
Knowledge graph
Knowledge reasoning
Assisted decision-making
Transloction Embedding (TransE) model

ABSTRACT

With the widespread use of Internet, the amount of data in the field of traditional Chinese medicine (TCM) is growing exponentially. Consequently, there is much attention on the collection of useful knowledge as well as its effective organization and expression. Knowledge graphs have thus emerged, and knowledge reasoning based on this tool has become one of the hot spots of research. This paper first presents a brief introduction to the development of knowledge graphs and knowledge reasoning, and explores the significance of knowledge reasoning. Secondly, the mainstream knowledge reasoning methods, including knowledge reasoning based on traditional rules, knowledge reasoning based on distributed feature representation, and knowledge reasoning based on neural networks are introduced. Then, using stroke as an example, the knowledge reasoning methods are expounded, the principles and characteristics of commonly used knowledge reasoning methods are summarized, and the research and applications of knowledge reasoning techniques in TCM in recent years are sorted out. Finally, we summarize the problems faced in the development of knowledge reasoning in TCM, and put forward the importance of constructing a knowledge reasoning model suitable for the field of TCM.

## 1 Introduction

Thousands of years have seen the development of traditional Chinese medicine (TCM) and the accumulation of relevant data in the forms of ancient books, research papers, medical cases, and clinical data [1]. How to make full and efficient mining and use of these valuable data has become a pressing issue. The advent of knowledge graphs provides the basis of ideas and methods for the interconnection and visualization of TCM knowledge. With the construction of TCM knowledge graphs and knowledge reasoning based on knowledge graphs, the potential connections between TCM entities are explored and the complex relationships within TCM knowledge are sorted out, which can facilitate assisted decision-making in TCM clinics and can effectively promote the rapid development of intelligent TCM.

Reasoning, as an important part of human cognition, has been defined by many scholars since ancient Greek as starting with existing knowledge and arriving at new knowledge [2], and reasoning is the process of using the appropriate rules to get from the starting point to the endpoint. In 2012, Google introduced the Knowledge Graph technology, which transforms independent knowledge into a triple of rules and builds an interconnected semantic network [3]. With the creation of many knowledge bases, such as YAGO, WordNet, and Freebase, the knowledge graph has become a focus of attention [4]. Studies

*Corresponding author: LIU Qingping, Associate Professor, E-mail: liuliu@hnucm.edu.cn.

Peer review under the responsibility of Hunan University of Chinese Medicine.

DOI: 10.1016/j.dcmed.2022.12.005
**Citation:** GUO ZH, LIU QP, ZOU BJ. Research on knowledge reasoning of TCM based on knowledge graphs. Digital Chinese Medicine, 2022, 5(4): 386-393.

show that the development of knowledge graphs and knowledge reasoning complement each other, and that knowledge reasoning plays a key role in the construction of knowledge graphs [5, 6]. Knowledge reasoning based on knowledge graphs is the process of reasoning about the knowledge of entities, attributes, and relationships within the graphs with the help of relevant technologies on the basis of the constructed graphs. The knowledge graph is composed in the form of entities, relationships, and attributes, which is in line with people's cognition of knowledge and can better reason and model the real, objective world. Therefore, the knowledge reasoning methods based on knowledge graphs have become typical in the field of knowledge reasoning. From Aristotle's syllogism to the knowledge reasoning algorithms based on knowledge graphs, reasoning has evolved from simple reasoning through continuous and iterative updates to knowledge graph-based knowledge reasoning models combined with computer technology, which are widely used in various industries and fields, including medical science and healthcare [7, 8]. The rapid development of Internet technology has made it easier to access knowledge of ancient TCM literature, facilitating the storage of TCM data and accelerating the adoption of knowledge graph technology in TCM [9]. In this paper, we take stroke as an example to explain knowledge reasoning methods, summarize the principles and characteristics of commonly used knowledge reasoning methods, and sort out the research and application of knowledge reasoning technology in the field of TCM in recent years.

## 2 Overview of knowledge reasoning

Knowledge reasoning based on knowledge graph has evolved from traditional knowledge reasoning methods, which has the advantages of traditional knowledge reasoning and also integrates modern computer technology. Traditional knowledge reasoning methods include deductive reasoning, inductive reasoning, and default reasoning, etc. These methods have been continuously improved by researchers, laying the foundation for the research of knowledge reasoning methods based on knowledge graphs [10]. With Google proposing the concept of knowledge graph, knowledge reasoning based on knowledge graphs has also become the focus of attention as an important technique for complementing and verifying knowledge graph triples [11]. The existing literature roughly classifies knowledge reasoning methods based on knowledge graphs into three categories, namely, those based on traditional rules, distributed feature representation, and neural networks [12, 13]. Among them, knowledge reasoning based on traditional rules is roughly similar to human thinking and can make use of people's existing experience for assistance, but it relies too much on expertise and has little potential for generalization and cross-discipline application; knowledge reasoning based on distributed feature representation maps entities and relationships into a low-dimensional vector space, making full use of the structural information between entities to carry out accurate computation; knowledge reasoning based on neural networks uses models to learn about entities and relationships, build models, and make predictions using structural and path information between entities in the graph.

Traditional knowledge reasoning methods have played an important role in the development of TCM, but it often falls short of understanding users' semantic queries when searching for TCM knowledge, and there are problems such as weak self-learning ability, low data utilization, and low search accuracy. For example, when a user needs to look up the symptoms of stroke, the traditional reasoning search engine will give hundreds of web pages, and most of them are textual, which to a certain extent affects the user's ability to find the target content accurately and quickly, resulting in low search efficiency and users' poor experience [14]. Knowledge reasoning search engine based on knowledge graphs can make extended searches for "stroke" and "symptoms" based on the user's semantics, which is more accurate and faster. The search engine can also list the causes, symptoms, diagnosis, treatment, and prevention measures for the disease.

## 3 Knowledge reasoning technology based on knowledge graphs

Knowledge reasoning technology based on knowledge graphs is to proceed to new factual conclusions from existing partial entity information basing on relevant rules, and feed the new factual conclusions back to the knowledge graph to improve its relevant contents. According to the analysis of existing literature, there are three main types of knowledge reasoning based on knowledge graphs, namely those based on traditional rules, distributed feature representation, and neural networks.

### 3.1 Knowledge reasoning based on traditional rules

Knowledge reasoning based on traditional rules can use knowledge graphs to learn simple rules and features [15] and obtain new facts by inference. The method takes full account of the symbolic meaning of knowledge, has a high degree of accuracy, and gives a clear explanation of the inference results. Depending on the method of rule representation relied on, knowledge reasoning based on traditional rules can be divided into knowledge reasoning based on first-order predicate logic and those based on descriptive logic.

**3.1.1 Knowledge reasoning based on first-order predicate logic** Knowledge reasoning based on first-order predicate logic uses propositions as the smallest unit of reasoning. Propositions consist of individuals and predicates, where individuals can be either concrete things in the physical world or abstract definitions in the mental space. A predicate is a property or fact used to describe an individual. Using propositions as the unit of reasoning can better simulate the human mind in reasoning, making the reasoning process more explanatory, and the results more precise. Taking "stroke" in the 10th edition of the 13th Five-Year Plan textbook on *Chinese Internal Medicine* as an example, the model is used to reason about the relationship between the symptoms of stroke and the medication used, e.g. (stroke, symptom, deviation of mouth and tongue, hemiplegia) $\wedge$ (Buyang Huanwu Decoction, treatment, deviation of mouth and tongue, hemiplegia) $\Rightarrow$ (stroke, medication, Buyang Huanwu Decoction). Given that a patient with stroke has symptoms such as deviation of mouth, tongue, and hemiplegia, and that Buyang Huanwu Decoction (补阳还五汤) is suitable for treating symptoms of deviation of mouth, tongue, and hemiplegia, it is evident that stroke can be treated with Buyang Huanwu Decoction.

**3.1.2 Knowledge reasoning based on descriptive logic** Propositional logic and first-order predicate logic have been developed and improved to form the prototype of description logic, which transforms the complex reasoning problems concerning entities and relationships into a set of term and assertion consistency checking problems through sets of terms and assertions. Specifically, descriptive logic maximizes both the complexity of entity reasoning and the expressive ability. Descriptive logic has logic-based semantic descriptions and strong expressive power, reasoning about more complex concepts based on simple concepts. It boasts an efficient representation mechanism that facilitates the handling of large amounts of data, thus improving the quality of service for reasoning and semantic query networks.

## 3.2 Knowledge reasoning based on distributed feature representation

Knowledge reasoning based on distributed feature representation represents entities and relationships in the knowledge graph as vectors, matrice or tensors, and performs the knowledge reasoning task through matrix transformation or inner product of vectors. Typical models include those based on tensor factorization, and translational distance model, etc.

**3.2.1 Knowledge reasoning based on tensor factorization** In the reasoning process, the graph is represented as a tensor, and a high-dimensional matrix is decomposed into a low-dimensional matrix by means of tensor decomposition to infer unknown relationships or entities. The RESCAL model is one of the classical models, transforming high-dimensional and multiple data into a third-order tensor, reducing the dimensionality while retaining its original data characteristics [16]. After decomposition, the composition structure of the entity or relationship vector is obtained, and whether the two are related is inferred by determining whether the vectors have similar composition structures. As shown in Figure 1, Patient A was diagnosed with stroke, showing symptoms such as deviation of mouth, tongue, and hemiplegia, poor skin, lack of verbal fluency, salivation at the corners of the mouth, thin white tongue, and fine pulse. Patient B had similar symptoms to Patient A. The RESCAL model showed that the symptoms of both patients were highly similar in structure, so it was inferred that Patient B also had a stroke.

**3.2.2 Knowledge reasoning based on translational distance model** Knowledge reasoning based on translational distance model maps vector entities to a low-dimensional vector space, regards the relationship in the map as a translation transformation between the head entity and the tail entity, and verifies whether the

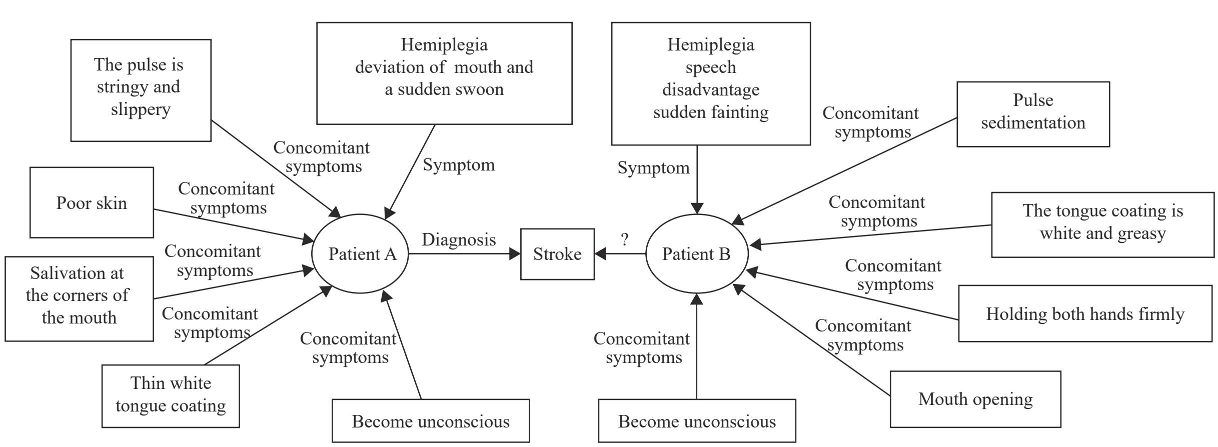


**Figure 1** Diagram of the RESCAL model

fact triple is reasonable by judging whether the translation vector meets the condition $h + r \approx t$. A typical model is Translating Embedding (TransE), which converts potential feature representations by specific relational offsets rather than just using simple vector matrix multiplication [17]. The performance of the TransE model is measured by the score function $f(h,r,t) = h + r - t_{l_1/l_2}$, where denotes the norm of the difference vector $l_1$ or $l_2$. The model allows inference to be made on fact triples. For example, the triple formed by the symptom, prescription, and the name of the prescription (weakness of the limbs and yellowing of the face, prescription, Buyang Huanwu Decoction) are shown in Figure 2. The symptoms of limb weakness and yellowing are represented by the vector h, the prescription of Buyang Huanwu Decoction is represented by *t*, and the relationship between them is represented by *r*. If $h + r \approx t$ is met in the vector space, it means that the symptoms of weak limbs and yellowing of the face can be treated with the prescription Buyang Huanwu Decoction. Due to the limitations of the TransE model in dealing with self-reflexive, one-to-many, many-to-one, and many-to-many relationships, researchers have enhanced the TransE model and designed the TransH model [18]. The TransH model solves the problem of using different representations of the same entity in different relationships by defining a hyperplane and a relationship vector for each relationship and projecting the entity onto the hyperplane of the specific relationship to achieve different representations of the entity in different relationships. For example, the disease names "stroke" "cerebral infarction" and "cerebral ischaemia" are expressed differently in different semantic contexts, but the projections in vector space all represent "stroke". This model can promote the research and application of relevant criteria of TCM.

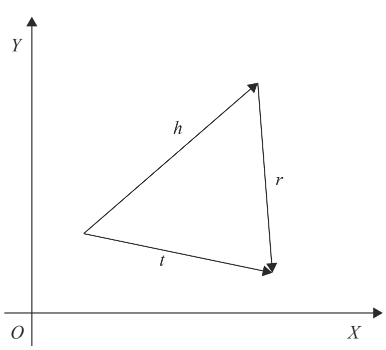


**Figure 2** TransE model diagram
Head + relation ≈ tail, $h + r \approx t$, head represents the head entity, tail represents the tail entity, and relationship represents the relationship between the two.

## 3.3 Knowledge reasoning based on neural networks

Neural networks learn to simulate human thought and perception. The input data of each neuron is multiplied by a certain weight and the product of the two is mapped to another feature vector space by a function mapping. The neurons are then bias adjusted, and the activation function normalized in the space to the output. Knowledge graphs combined with neural network technology learn the structure and semantic features of the knowledge graphs in greater depth, and can better infer the implicit relationships in knowledge graphs and verify the existing entities and relationships. Knowledge reasoning based on neural networks generally refers to the use of certain properties of neural networks for inference [19], such as predicting the missing part of a triple or reasoning about the existence of a relationship between two entities at the beginning and end of a multi-step path. Representative methods include knowledge reasoning based on convolutional neural networks and knowledge reasoning based on recurrent neural networks.

**3.3.1 Knowledge reasoning based on convolutional neural networks** Knowledge reasoning based on convolutional neural network refers to the extracting local features in entities through convolutional operations, embedding representations of entities and relationships in the knowledge graph to form an image-like two dimensional (2D) structure, and the extracting entity features through convolutional kernels in the convolutional network to predict the implicit relationships between entities. Among the typical methods are models such as Description-embodied Knowledge Represen-tation Learning (DKRL) and Convolutional 2D Knowledge Graph Embeddings (ConvE).

The DKRL model is a continuous bag of words model (CBOW) and a deep convolutional neural model (CNN), which are used for the semantic description of entities to learn the unordered features and word order features of text data in order to reduce the impact of data sparsity [20]. As shown in Figure 3, the CBOW model encoder selects the keywords of the entity and sums up the embeddings of all the keywords to obtain the embedding of the entity. Since the CBOW model has difficulty in capturing the word order information, which reduces the quality of the extracted keywords, the CNN encoder is added to obtain the entity encoding. As shown in Figure 4, the CNN model first preprocesses the original text and word representation, while the phrase vocabulary from the training set is initialized with embedding using word2vec, then convolutional operations are performed to extract some of the features, and then the pool is trained using the scoring

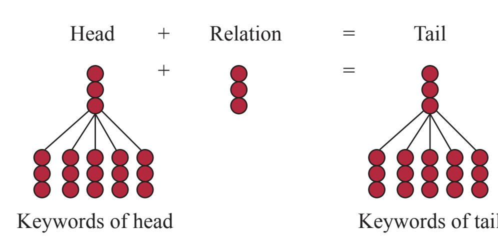


**Figure 3** CBOW encoder model diagram

function. The DKRL model not only improves the entity representation differentiation capability, but also achieves the embedding representation of new entities. When a new entity is generated, the DKRL model can generate a corresponding entity representation based on a short information description of the new entity for tasks such as knowledge reasoning.

Subsequently, scholars have proposed the ConvE model, which is a multilayer convolutional neural network model including convolutional, projection, and inner product layers [21]. The model is converted into a 2D tensor by embedding a representation entity and a relational vector stack, and undergoes operations such as convolution, pooling, and fully connected projection to extract entity and relational feature information, which is then computed with the matrix and tail entities to determine the credibility of the current triple. The model performs embedding matching with the candidate targets during the inner product calculation, and reasons about the new links in the knowledge graph therewith.

**3.3.2 Knowledge reasoning based on recurrent neural networks** Knowledge reasoning based on recurrent neural networks refers to the acquisition of entity content through recurrent structures to predict missing parts of the graph. NEELAKANTAN et al. [22] argue that knowledge reasoning that fuses relational paths and neural networks is more worthy of study and propose the Path-Recurrent Neural Network (Path-RNN) model, where Path Ranking Algorithm (PRA) is used to find a different path for each relation type and the embedding of the binary relation in that path is used as an input vector to output a relation vector through the relational semantic neighborhood of the entire path between the head entity and the tail entity. Similarity is caultulated for the whole path vector representation and the vector representation of the relationship to be predicted, and the predicted relationship vector corresponding to the highest value of the similarity result is the target path. For example, using the Path-RNN model to calculate the similarity between the overall path formed by stroke symptoms, syndrome types, prescriptions, and the relationship between stroke pointing to Buyang Huanwu Decoction, it can be reasoned that stroke characterized by Qi deficiency and blood stasis can be treated with Buyang Huanwu Decoction. The Path-RNNN model diagram is shown in Figure 5.

In addition to the typical reasoning methods mentioned above, many experts and scholars at home and abroad have continued to work on knowledge reasoning techniques and their application in recent years, proposing many new reasoning models and applying knowledge reasoning techniques to a wide range of areas in the medical field. For example, KUMAR et al. [23] proposed a hybrid approach that comines instance-based reasoning and rule-based reasoning, which was applied to a clinical decision support system in an intensive care unit to improve its monitoring and care capabilities. KAREGOWDA et al. [24] improved the accuracy of diabetes assisted diagnosis to a certain extent by using a hybrid reasoning model combining genetic algorithm and back propagation neural network. LIU et al. [25] proposed a knowledge-based reasoning model for semantic prediction based on cross-language vocabulary (CLSP), defining the language with annotated semantics as the source language and the language without annotated semantics as the target language, embedding the words of the original language and the target language into the same semantic space through joint learning, and then seeking the source language with similar meanings to the words to be recommended in the target language, and recommending semantics for the target words with the annotation information of the source language. In addition, path-based graph structure

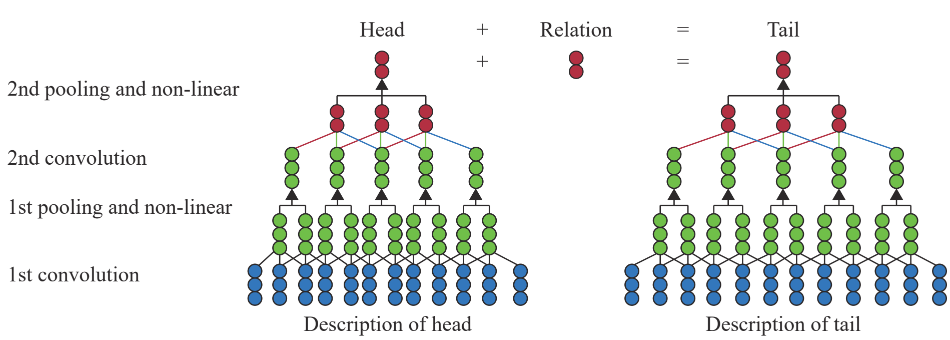


**Figure 4** CNN encoder model diagram

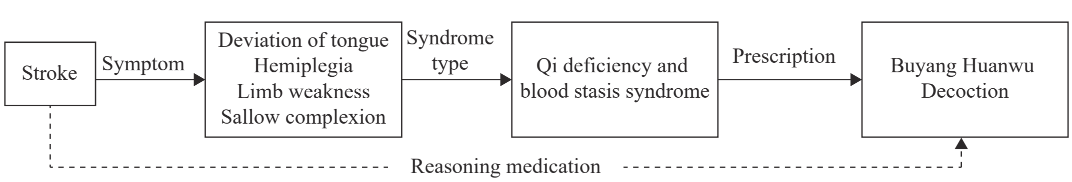


**Figure 5** Path-RNN model diagram

inference algorithms, such as the PRA [26] and its modified version Constant and Reversed Path Ranking Algorithm (Cor-PRA) [27], Coupled Path Ranking Algorithm (CPRA) [28], Enhanced Link Prediction (ELP) and others can also use the paths between entities as features to determine whether there is a target relationship between entities, and these methods can automatically mine path rules and are interpretable. In summary, research and results related to knowledge reasoning have contributed to the development of the healthcare field, and have extremely broad application prospects and social value [29].

## 4 Application of TCM knowledge reasoning

Domestic and foreign researchers have paid much attention to the study of knowledge reasoning techniques in medicine. In foreign studies based on western medicine, scholars have continued to try to apply knowledge reasoning techniques to assistive clinical treatment systems. For example, FECHOK et al. [30] used the biolink model as a high-level ontology for knowledge representation of relationships between medical entities with the help of the knowledge graph-based biomedical data translator, and combined it with the MACT question evaluation system, which led to improved performance of the Q&A system. The MYCIN treatment system developed by Stanford University is able to perform rule-based reasoning on the knowledge of medical experts with a heuristic backward reasoning strategy to help doctors with diagnosis assistance and antibiotic dosing recommendations for hospitalized patients with bloodstream infections, and gives reasonable antibiotic dosing recommendations in relation to the patient's weight [31]. Subsequently, Stanford University developed a PUFF pulmonary function testing system to assist medical specialists in performing pulmonary function tests, which has been validated in clinical trials with a success rate of 93% [32]. CASNET, a system developed by Rugers University, uses uncertainty-based reasoning to assist doctors in the diagnosis and treatment of glaucoma. In China, many scholars have applied knowledge reasoning techniques to the field of TCM, using traditional knowledge reasoning and knowledge graph-based knowledge reasoning techniques to carry out theoretical innovations and their applications. For example, LIU et al. [33] quantified symptom descriptions into numerical matrices with the help of BP neural networks, automatically reasoned out the evidence typing of hypertension by defining a rule base, and simulated prescriptions to realize the direct mapping from evidence to prescription, which improved the clinical decision-making ability of TCM practitioners. WEI et al. [34] discussed the knowledge representation and reasoning methods used in the development of digital dialectic in TCM, focusing on the idea of “treatment based on pattern differentiation”, which is of great significance in promoting the development of intelligentual assisted diagnosis in TCM. ZHU et al. [35] took an ontology-based approach to the semantic expression of asthma disease patterns, symptoms, and medication by extracting the knowledge related to asthma from TCM literature. WU et al. [36] used rule-based reasoning to build a decision aid system based on the Drools rule engine to predict specific symptoms and complications of diabetes mellitus, improving the self-measurement ability of the diabetic population, and the clinical diagnosis and decision-making ability of health care professionals. HAO [37] described the five steps of constructing the knowledge graph, and gave the definitions of four types of entities, namely, “western disease” “syndromes” “Chinese medicine” and “symptoms”. By obtaining entities, attributing relationships and semantics, the entities and relationships are fused to generate ontologies, and the rule-based knowledge reasoning method is used to realize the construction of the TCM health knowledge graph. In summary, experts and scholars at home and abroad are keen to apply knowledge reasoning technology in the research of assisted clinical diagnosis and treatment systems, and achieve corresponding results, which has to a certain extent promoted the development of information-based and intelligent TCM.

In recent years, the state has attached great importance to the inheritance, innovation, and development of TCM, and actively encouraged the application of new generation information technology such as artificial intelligence to assisted TCM diagnosis and treatment systems to promote the inheritance and innovation of the clinical experience of prestigious Chinese physicians [38]. TCM knowledge reasoning technology is a key component in the assisted TCM diagnosis and treatment system, which can be used to conduct research on assisted diagnosis and treatment with the clinical experience of prestigious Chinese physicians, and effectively improve the accuracy of diagnosis and treatment of the assisted TCM diagnosis and treatment system.

With the “healthy China strategy” in full swing, people are paying more and more attention to health issues, and the problem between the increasing demand for medical care and the insufficient supply of superior medical resources is becoming more and more pressing. The application of knowledge reasoning based on knowledge graph technology in assisted diagnosis and treatment [39], assisted decision making [40], knowledge recommendation [41], and medical knowledge Q&A [42, 43] has greatly improved the level of primary healthcare services and alleviated certain urgent problems in the medical field.

## 5 Summary

Through a typological investigation of knowledge reasoning methods and research on the application of knowledge reasoning in the field of TCM, it is found that there are still many problems to be solved in the research of knowledge reasoning in Chinese medicine. Among them, the development of knowledge reasoning based on knowledge graphs in the field of TCM is still limited by the complexity of TCM knowledge and the existing inference capability, and no significant breakthrough has yet occurred for three main reasons. Firstly, it is difficult to construct TCM knowledge graphs. The lack of unified standards for TCM knowledge, the high number of specific textual symbols and conventional symbols signifying professional concepts in the field, the strong professionalism of terms, as well as the frequent existence of phenomena, such as multiple meanings of one word and multiple words with one meaning, have increased the difficulty of constructing TCM knowledge graphs to a certain extent. Secondly, it is difficult to choose a method for reasoning about TCM knowledge. The knowledge system in the field of TCM is characterized by ambiguity, complexity, and diversity. The individualization and personalization of TCM treatment are prominent, and the resulting knowledge graph with few samples in the field of TCM has a greater probability of incorrect a priori knowledge during model training, which further affects the interpretability of the knowledge reasoning model and the accuracy of the results, making it difficult to select a suitable knowledge reasoning model for the TCM assisted diagnosis and treatment system. Thirdly, the interpretability of TCM knowledge inference models is relatively low. At present, various types of knowledge reasoning methods are being developed under different research paradigms, and it is difficult to strike a balance in the process and accuracy of TCM knowledge reasoning. Their degree of fitting to complex and diverse TCM knowledge is not high, and they cannot solve problems such as deviations between inference results and reality due to variable disease symptoms, which reduces the interpretability of the models.

In summary, the current problem of restricted development of TCM needs to be further addressed by exploration of more diverse reasoning structures as well as the theoretical construction and practical application of knowledge reasoning strategies suitable for the complex reasoning patterns of TCM, so as to assist TCM experts to carry out diversified decision-making and promote the development of TCM inheritance and innovation.

## Fundings

The National Key R&D Program of China (2018AAA0102100), Hunan Provincial Department of Education Outstanding Youth Project (22B0385), Open Fund of the Domestic First-class Discipline Construction Project of Chinese Medicine of Hunan University of Chinese Medicine (2018ZYX17), Electronic Science and Technology Discipline Open Fund Project of School of Information Science and Engineering, Hunan University of Chinese Medicine (2018-2), and Hunan University of Chinese Medicine Graduate Innovation Project (2022CX122).

## Competing interests

The authors declare no conflict of interest.

## References

[1] ZHANG Y, GUO WZ, LIN S, et al. Review of the combination of deep learning and knowledge reasoning. Computer Engineering and Applications, 2022, 58(1): 56–69.
[2] CHEN X, JIA S, XIANG Y. A review: knowledge reasoning over knowledge graph. Expert Systems with Applications, 2020, 141: 112948.
[3] MA RX, LI ZY, CHEN ZK, et al. A review of knowledge graph reasoning. Computer Science, 2022, 49(S1): 74–85.
[4] MA A, YU YH, YANG SL, et al. Review of knowledge graph based on reinforcement learning. Journal of Computer Research and Development, 2022, 59(8): 1694–1722.
[5] QI GL, GAO H, WU TX. The research advances of knowledge graph. Information Engineering, 2017, 3(1): 4-25.
[6] WANG WG. Knowledge graph reasoning: modern methods and applications. Big Data, 2021, 7(3): 42–59.
[7] CHEN RC, HUANG YH, BAU CT, et al. A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. Expert Systems with Applications, 2012, 39(4): 3995–4006.
[8] BOUSQUET C, HENEGAR C, LILLOLE LA, et al. Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach. International Journal of Medical Informatics, 2005, 74(7-8): 563–571.
[9] FU LJ, CAO Y, BAI Y, et al. Development status and prospect of knowledge graph in vertical field in China. Application Research of Computers, 2021, 38(11): 3201–3214.
[10] GUAN SP, JIN XL, JIA YT, et al. Research progress of knowledge reasoning oriented to knowledge map. Journal of Software, 2018, 29(10): 2966-2994.
[11] WANG S, DU ZJ, MENG XF. Research progress of large-scale knowledge graph completion technology. Science in China: Information Science, 2020, 50(4): 551–575.
[12] MA ZG, NI RY, YU KH. The latest progress, key technologies and challenges of knowledge graph. Chinese Journal of Engineering, 2020, 42(10): 1254–1266.
[13] TIAN L, ZHANG JC, ZHANG JH, et al. Survey of knowledge graph: representation, construction, reasoning and knowledge hypergraph theory. Computer Applications, 2021, 41(8): 2161–2186.
[14] ZHOU XY, LIAO SY, DONG ZH, et al. Research on data map visualization and knowledge question answering platform of proprietary Chinese medicines. Software Guide, 2022, 21(5): 158–162.
[15] XIA Y, LAN MJ, CHEN XH, et al. An overview of interpretable knowledge map reasoning methods. Journal of Network and Information Security, 2022, 8(5):1–25.
[16] NICKEL M, TRESP V, KRIEGEL HP. A three-way model for collective learning on multi-relational data. Proceedings of the 28th International Conference on Machine Learning, 2011.
[17] BORDES A, WESTON J, COLLOBERT R, et al. Learning structured embeddings of knowledge bases. Twenty-fifth AAAI Conference on Artificial Intelligence, 2011: 301–306.
[18] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes. Proceedings of the AAAI Conference on Artificial Intelligence. 2014. doi: 10.1609/aaai.v28il.8870.
[19] ZHANG JX, ZHANG XS, WU CX, et al. Review of knowledge graph construction technology. Computer Engineering, 2022, 48(3): 23–37.
[20] XIE R, LIU Z, JIA J, et al. Representation learning of knowledge

graphs with entity descriptions. Proceedings of the AAAI Conference on Artificial Intelligence, 2016, 2659–2665.
[21] ARVIND N, BENJAMIN R, ANDREW M. Compositional vector space models for knowledge base completion. Stroudsburg: Association for Computational Linguistics, 2015: 156–166.
[22] NEELAKANTAN A, ROTH B, MCCALLUM A. Compositional vector space models for knowledge base completion. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015: 156–166.
[23] KUMAR KA, SINGH Y, SANYAL S. Hybrid approach using case-based reasoning and rule-based reasoning for domain independent clinical decision support in ICU. Expert Systems with Applications, 2009, 36(1): 65–71.
[24] KAREGOWDA AG, MANJUNATH AS, JAYARAM MA. Application of genetic algorithm optimized neural network connection weights for medical diagnosis of pima Indians diabetes. International Journal on Soft Computing, 2011, 2(2): 15–23.
[25] LIU ZY, HAN XU, SUN MS. Knowledge graph and deep learning. Beijing: Tsinghua University Press, 2020.
[26] LAO N, COHEN W W. Relational retrieval using a combination of path-constrained random walks. Machine Learning, 2010, 81(1): 53–67.
[27] LAO N, MINKOV E, COHEN W. Learning relational features with backward random walks. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), 2015: 666-675.
[28] WANG Q, LIU J, LUO Y, et al. Knowledge base completion via coupled path ranking. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2016: 1308-1318.
[29] KANG L. Research and implementation of cardiovascular disease question answering system based on knowledge graph. Guangzhou: South China University of Technology, 2020.
[30] FECHO K, BALHOFF J, BIZON C, et al. Application of MCAT questions as a testing tool and evaluation metric for knowledge graph-based reasoning systems. Clinical and Translational Science, 2021, 14(5): 1719–1724.
[31] TU ZS. Research on the development of web-based TCM diagnostic expert system using JESS. Zhengzhou: Zhengzhou University, 2010.
[32] DONG D. Research on medical image intelligent diagnosis system based on case reasoning. Harbin: Harbin University of Science and Technology, 2009.
[33] LIU J, JIANG WM, SHEN GJ. Construction and application of TCM expert diagnosis and treatment system for hypertension oriented to big data. China Journal of Library and Information of Traditional Chinese Medicine, 2019, 43(5): 5–9.
[34] WEI CF, YAN JF. Discussion on the development of digital dialectics in traditional Chinese medicine from the perspective of knowledge representation and reasoning methods. Chinese Journal of Traditional Chinese Medicine, 2019, 34(10): 4471–4473.
[35] ZHU L, LI JH, YU Q, et al. Construction of ontology in asthma domain of traditional Chinese medicine. Chinese Journal of Experimental Traditional Medical Formulae, 2017, 23(15): 222-226.
[36] WU J, ZHU L, KANG L, et al. Research on diabetes syndrome and complication prediction based on Drools rule engine. China Digital Medicine, 2021, 16(6): 43–47.
[37] HAO WX. Research on the construction of TCM health knowledge graph. Beijing: Beijing Jiaotong University, 2017.
[38] LIU F, WANG MQ, LI LX, et al. Exploration on the construction method of clinical experience knowledge graph of famous and old Chinese medicine. Chinese Journal of Traditional Chinese Medicine, 2021, 36(4): 2281–2285.
[39] WANG HN, SUN YQ, ZHANG KX. Research on the application of knowledge graph in the field of traditional Chinese medicine. Journal of Liaoning University of Traditional Chinese Medicine, 2022, 24(8): 182–185.
[40] FAN YY, LI ZM. Research and application progress of Chinese medical knowledge graph. Computer Science and Exploration, 2022, 16(10): 2219–2233.
[41] QIN C, ZHU HS, ZHUANG FZ, et al. A review of recommendation system based on knowledge graph. Science in China: Information Science, 2020, 50(7): 937–956.
[42] HU HJ, ZHOU Y, KUANG ZM, et al. Research progress in application of medical knowledge map. Journal of Medical Informatics, 2022, 43(5): 30-33, 39.
[43] REN W. Research on intelligent question answering system of traditional Chinese medicine knowledge. Tangshan: North China University of Science and Technology, 2020.

# 基于知识图谱的中医药知识推理研究

郭志恒[a], 刘青萍[a*], 邹北骥[a, b]

*a. 湖南中医药大学信息科学与工程学院，湖南 长沙 410208，中国*
*b. 中南大学计算机学院，湖南 长沙 410083，中国*

**【摘要】**随着互联网技术的广泛应用，中医药行业领域数据规模呈指数型增长，如何从中筛选出有用的知识并有效组织和表达备受关注。知识图谱由此而生，基于知识图谱的知识推理成为研究的热点之一。本文首先简要介绍知识图谱和知识推理的发展及探讨知识推理的意义。其次，介绍主流的知识推理方法分类，包括基于传统规则的推理、基于分布式特征表示的推理、基于神经网络的推理。再以脑卒中疾病为实例，对知识推理方法进行阐述，总结常用知识推理方法的原理及特点，并梳理近些年知识推理技术在中医药领域的研究与应用。最后，总结中医药知识推理发展所面临的问题，提出构建适合中医药领域的知识推理模型的重要性。

**【关键词】**中医药；脑卒中；知识图谱；知识推理；辅助决策；TransE 模型