

Construction of Chinese Traditional Culture Knowledge Graph and Intelligent Question Answering System Based on Graphormer

Shengxue Lei

School of Marxism, Luohe Medical College, Luohe 462002, Henan, China

Corresponding author: Shengxue Lei (e-mail: leishengxue@126.com)

ABSTRACT Traditional graph neural networks have limitations in modeling long-range multi-hop relation paths and sparse semantic entity representations when constructing knowledge graphs for Chinese traditional culture. To address variant character forms, syntactic abbreviations, and sparse entity semantics in ancient texts, this paper proposes a knowledge graph construction and question-answering system based on Graphormer. The Confucian four-element inheritance structure of “teacher-student-writing-school-event” and the three-element mapping of “rites, textiles, and crafts” are formalized as node types and directed relations in a knowledge graph. An edge-type encoding and path-constraint mechanism is constructed accordingly. Based on this structure, Graphormer’s shortest-path distance bias and edge-feature encoding are guided to prioritize multi-hop paths consistent with Confucian transmission logic. By embedding the philosophical transmission structure into a graph attention mechanism, the method achieves collaborative modeling of sparse semantic relations and long-range dependencies. Experiments show that the proposed method achieves Hits@10 of 0.659 in the knowledge completion task and accuracies of 0.773, 0.715, and 0.667 in four-hop multi-hop question answering, verifying the effectiveness of the structural expression and reasoning paradigm.

INDEX TERMS graphormer, knowledge graph, traditional chinese culture, intelligent question answering, textile techniques

I. INTRODUCTION

CHINESE traditional culture encompasses a vast array of historical figures, schools of thought, classic texts, and significant historical events, all of which constitute the essential knowledge system of Chinese civilization [1,2]. Due to the diverse morphology of the corpus, the high number of ancient and modern variants in textile terminology, and ambiguous semantic expressions (e.g., in describing weaving patterns), existing digital Knowledge Graph (KG) construction for textile cultural heritage faces challenges such as incomplete association extraction and low disambiguation accuracy [3,4]. Existing methods struggle to effectively model semantic dependency paths exceeding two hops, resulting in low recall and confusing embedding representations in knowledge graphs for complex relational reasoning tasks. Improving the search efficiency of multi-hop relational paths and the inter-class separation of entity embeddings are core technical challenges currently facing the construction of cultural knowledge graphs.

Current research on the construction of Chinese traditional cultural KGs and intelligent question-answering systems focuses on solving the problems of low Named

Entity Recognition (NER) accuracy, insufficient relationship extraction, and insufficient understanding of question-answering systems in specific cultural fields [5-7]. Liu Shuang constructed a KG of ancient Chinese history and culture, using a BiLSTM (Bidirectional Long Short-Term Memory)-CNN (Convolutional Neural Network)-CRF (Conditional Random Field) NER method combined with DeepKE (Deep Knowledge Extraction) relationship extraction, and integrating information box triples to improve the structuring and dissemination efficiency of cultural knowledge [8]. Chansanam Wirapong studied the construction of KG by taking Thai culture as the research object. He collected Wikipedia data by crawling and used the BiLSTM model to complete NER and relationship extraction with an accuracy of 84.73%, verifying the effectiveness of this method in the structuring of cultural knowledge [9]. In order to reveal the profound cultural connotation of traditional Chinese silk weaving technology, Xu Liang took Yunjin knowledge as the object. It built a knowledge graph-based intelligent question-answering system for Nanjing Yunjin digital resources. The system used the intent recognition technology based on BERT (Bidirectional Encoder Rep-

representations from Transformers) to classify user queries according to intent. It used the BERT+BiGRU (Bidirectional Gated Recurrent Unit)+CRF model for NER. Experimental results showed that the proposed question-answering system can efficiently and accurately generate answers to users' natural language queries, greatly promoting the retrieval and utilization of Yunjin knowledge [10]. Yuan Hua proposed an intelligent question-answering system based on a domain-specific knowledge graph, which uses the RoBERTa (Robustly Optimized BERT Pretraining Approach)-BiLSTM-SDPA (Scaled Dot-Product Attention)-CRF model to integrate NER, question classification, and retrieval modules, improving the accuracy of cultural question answering [11]. From the perspective of knowledge completion, Zhang Min proposed the BERT model that combines entity type information, enriching entity semantics, and reducing relationship ambiguity through text serialization input [12]. Overall, these studies have effectively improved the accuracy and adaptability of cultural KG construction and question-answering systems. However, these methods mostly focus on modeling local entities and binary relationships, lacking the ability to model the global structure of knowledge graphs. In particular, in complex semantic scenarios such as the teacher-student relationship of Confucianism and the citation network between classics, existing methods struggle to model long-range dependency logic.

To address the challenge of jointly modeling the complex structure and deep semantics of traditional cultural knowledge, some research is gradually turning to the integration of GNNs and deep semantic models [13-15]. Wang Yi proposed a knowledge extraction and construction method based on the Graph Attention Network (GAT), and verified its improved knowledge extraction efficiency on Tang Dynasty gold and silver artifacts data. This method breaks through the limitations of traditional fixed-weight aggregation and enables key cultural entities to be more accurately represented in the KG [16]. As graph structure modeling capabilities gradually improve, researchers have begun to explore how to integrate semantically rich Transformer architectures into knowledge representation processes to enhance the contextual understanding of nodes. Hou Wenjun designed a graph representation learning model that integrates cultural attributes. By concatenating word vectors to construct a feature matrix, he combined TransR to embed domain-specific knowledge, improving entity representation capabilities [17]. Wang Dongsheng proposed a Transformer-based multimodal knowledge graph link prediction model, which uses a multimodal encoder to extract structural, visual, and textual features, and implements symmetric hybrid computing in the Transformer architecture, effectively solving the feature fusion problem [18]. These studies have shown that the global context modeling capability of the Transformer can effectively supplement the shortcomings of graph neural networks in depicting semantic details. However, existing methods often separate GNN and Transformer modules or shallowly stack them, failing to achieve deep fu-

sion of graph structure and semantic information. The latest research attempts to build a unified graph-sequence joint architecture to achieve end-to-end collaborative optimization. GraphFormers proposed by Yang Junhan fuses hierarchical GNNs with Transformer blocks, enabling node encoding to have both global semantics and local structural information, significantly enhancing information integration capabilities [19]. Current research has demonstrated the feasibility of the Graph and Transformer architecture in the field of cultural KGs, but existing research still lacks systematic exploration of applying GraphFormers to the construction of Chinese traditional cultural KGs and question-answering. In particular, there is limited research on forming a unified coding framework for the entire process of entity and relationship extraction, knowledge completion, and question-answering reasoning, resulting in challenges in the reasoning ability of complex cultural logic.

The transmission of Confucian thought is essentially a multi-hop semantic network with clear directionality, intergenerational connections, and textual dependencies. This structure requires knowledge representation models to distinguish relation types; model long-range dependencies across multiple generations; and maintain semantic consistency of paths under sparse connections. Traditional GNNs, due to their local aggregation mechanisms, struggle to meet these requirements. To improve the authenticity and interpretability of answers and promote the intelligent organization and dissemination of traditional cultural knowledge, this paper constructs a Chinese traditional culture KG construction and an intelligent question-answering system based on Graphormer. This paper uses the inheritance of Confucian thought as its logical framework and the etiquette and clothing system, weaving ethics, and skill transmission in traditional Chinese textile culture as its final application scenario. It does not merely use Confucian thought as a source of data, but rather its inheritance logic itself as a structural knowledge modeling paradigm. The intergenerational transmission of Confucian thought has clear topological characteristics: it uses core figures as hubs, forms a genealogy through teacher-student relationships, solidifies the thought through classical writings, and evolves into schools of thought through historical events. A graph structure is constructed based on a four-layer progressive structure of "figures, writings, schools, and events." In the knowledge graph construction, this philosophical model is formalized into a set of node types and directed relations in a heterogeneous graph. Shortest path distance bias strengthens reasoning paths that conform to the length of the Confucian inheritance chain, while edge type embedding distinguishes between institutional and ideological relationships, allowing attention weights to prioritize culturally rational paths in multi-hop reasoning, achieving a layer-by-layer mapping from the philosophical model to the graph structure and then to the attention mechanism. Its innovations are: (1) Graphormer is applied to the Confucian thought inheritance modeling scenario, and node degree and path encoding

are combined to achieve long-range semantic dependency modeling of thought inheritance; (2) Cultural relationships are adapted through edge type encoding to achieve a unified representation of multi-type edges of cultural knowledge; (3) A multi-hop reasoning mechanism guided by global attention weight is designed, and candidate reasoning chains are dynamically expanded according to path importance to cope with the complex reasoning paths of Confucian thought context; (4) Explanatory, complex comparative and factual cultural knowledge questions and answers are applied as scenario tests to ensure that the reasoning path conforms to the logic of traditional cultural thought. This paper not only addresses the application gap of Graphormer in the domain of Chinese traditional textile culture but also offers a novel solution for the digital modeling, educational dissemination, and intelligent inheritance of this specialized cultural knowledge.

II. GRAPHORMER-ENHANCED CHINESE TRADITIONAL CULTURE KNOWLEDGE GRAPH MODELING AND QUESTION-ANSWERING SYSTEM CONSTRUCTION

A. CONSTRUCTION OF THE CHINESE TRADITIONAL CULTURE KNOWLEDGE GRAPH

1) Corpus Preprocessing and Normalization

This paper uses the inheritance of Confucian thought as the subject of knowledge modeling. The data sources include individuals, classics, schools of thought, and related historical events. Data collection adheres to the principles of "authority and comprehensiveness," selecting two sources: digital classics and semi-structured historical and cultural materials. The two are used for entity and relation extraction, mainly including: (1) Confucian core classics and their commentaries, including classics such as the *Analects*, *Mencius*, *Book of Rites*, *Xunzi*, *Zhu Xi's Collected Sayings*, and the *Four Books with Commentary*; (2) textile-related historical documents, including monographs such as the *Book of Sericulture*, the *Record of Weaving and Dyeing*, and the *Record of Brocade*, to establish a three-dimensional mapping of ritual, textiles, and craftsmanship. Table 1 shows the scale of the corpus data collection:

In Table 1, the digitized classics library primarily relies on the "Chinese Ancient Books Resource Library" of the National Library of China as its core text source. This library contains high-definition images and some OCR (Optical Character Recognition)-recognized text for over 100,000 ancient books, covering the four categories of Classics, History, Philosophy, and Anthology, as well as Confucian classics such as the *Analects of Confucius* and their historical annotations. This paper selects core classics related to the transmission of Confucian thought from the pre-Qin to the Ming and Qing dynasties and extracts their OCR-recognized text. Semi-structured historical and cultural data, sourced from Chinese Wikipedia entries and selected academic article abstracts from CNKI and the Wanfang Data Platform, possess a well-defined entity-attribute structure. During the collection process, the fields of the semi-structured data

are directly mapped to attribute values, while the original text retains contextual information. Data is stored in the standardized UTF-8 (Unicode Transformation Format-8) encoding format, and format checks are performed during the import phase to eliminate samples containing garbled characters, missing segments, or excessively high OCR error rates.

Preprocessing is performed to address issues with format heterogeneity, text noise, and differences in word usage between ancient and modern times in the original corpus. Since most Confucian classics are written in classical Chinese, their sentence boundary identifiers are irregular, with sentences sometimes separated by ".", ";", or even commas and spaces. This paper uses a CRF-based classical Chinese sentence segmentation model, whose conditional probability formula for determining sentence boundaries is:

$$P(y|x) = \frac{\exp(\sum_{i=1}^n \sum_k \lambda_k f_k(y_{i-1}, y_i, x, i))}{\sum_{y'} \exp(\sum_{i=1}^n \sum_k \lambda_k f_k(y'_{i-1}, y'_i, x, i))} \quad (1)$$

x is the input character sequence, y is the sentence boundary label sequence, f_k is the feature function, and λ_k is the corresponding weight. Among them, the feature function f_k contains three types of information: (1) character 1–3 grammar; (2) local punctuation context window; (3) the frequency features of sentence-end function words based on the statistics of the Chinese Dictionary.

To address missing and misplaced punctuation marks that occur during the digitization of ancient books, a BiLSTM-CRF-based automatic punctuation recovery method uses a combined training approach based on syntactic features and contextual embeddings to enable the model to predict the location and type of missing punctuation marks, restoring the standard punctuation system of modern Chinese. Based on the sentence segmentation results, modern standard punctuation marks are inserted at the predicted beginning and end of the sentence. The BiLSTM encoder integrates character embeddings and part-of-speech context features, and the CRF layer models the punctuation transfer constraints. For the numerous ancient and modern variant forms and forms of characters found in Confucian classics, a normalized mapping table M is constructed based on the General Standard Chinese Character Table and the Chinese Dictionary. It covers mapping relationships such as phonetic loan characters, taboo characters, and colloquial characters. The mapping function is defined as:

$$Norm(c) = \begin{cases} M[c], & \text{if } c \in M, \\ c, & \text{otherwise} \end{cases} \quad (2)$$

In the preprocessing stage, the character stream is transformed in sequence through the normalization function $Norm(\cdot)$. Then, an entity mapping table is established on this basis to unify the multiple expressions of the same cultural entity in different texts. Through manual annotation

TABLE 1. Corpus Collection Scale

Data type	Data Source	Number of texts	Average number of words per article	Number of structured fields
Chinese Ancient Books Resource Library	Digital Classics Library	12571	1050	0
Semi-structured historical and cultural materials	Wikipedia Chinese articles	2582	386	7
Semi-structured historical and cultural materials	CNKI (China National Knowledge Infrastructure)	1331	285	6
Semi-structured historical and cultural materials	Wanfang Data Platform	1049	261	6

and vector similarity judgment, a mapping relationship set is constructed:

$$S = \{(e_i^{variant}, e_i^{canonical}) \mid e_i^{variant} \in Variants, e_i^{canonical} \in Canonical\ Entities\} \quad (3)$$

$e_i^{variant}$ is an alias, $e_i^{canonical}$ is a standard entity, and the similarity calculation uses the cosine similarity of the word vector generated by the Chinese pre-training model based on contrastive learning sentence vector representation:

$$sim(v_a, v_b) = \frac{v_a \cdot v_b}{\|v_a\| \|v_b\|} \quad (4)$$

Threshold $\tau = 0.85$ is set. When the similarity exceeds the threshold and is manually verified, it is included in the mapping table. The final output of the normalized corpus maintains a high degree of consistency in word form, sentence structure, encoding, and entity reference, providing reliable input for subsequent knowledge extraction and graph construction.

2) Domain-Adaptive Entity and Relation Extraction

This paper constructs a domain-adaptive extraction framework based on a pre-trained language model for a joint corpus consisting of Confucian classics and textile works. This framework aims to address domain transfer barriers, scarce labeled samples, and ambiguity in knowledge-intensive contexts during entity and relation extraction. It jointly trains with a small amount of manually labeled data and weakly labeled data generated by remote supervision, and achieve entity recognition, relation extraction, entity disambiguation, and relation standardization through multi-stage optimization, thereby generating a high-confidence initial triple set (h, r, t) , where h is the head entity, r is the relation type, and t is the tail entity.

RoBERTa is used as the basic language model. The RoBERTa model completes mask modeling pre-training on a large-scale Chinese corpus and can capture Chinese lexical and syntactic features [20,21]. To adapt to the mixed ancient and modern literary characteristics of Confucianism, Domain-Adaptive Pretraining (DAP) is performed on the pre-processed corpus, and the optimization objective function is:

$$L_{MLM} = -\frac{1}{N} \sum_{i=1}^N \log P(\omega_i | \hat{\omega}_i; \theta) \quad (5)$$

$\hat{\omega}_i$ represents a randomly masked word fragment, ω_i represents the original word, θ represents the model parameters, and N represents the number of samples. The manual annotations are constructed by domain experts based on 1,200 high-value sentences. These annotations include four entity types: person, work, school, and event, as well as seven relationships: "disciple," "authored by," and "influenced by thought." Weakly supervised data utilizes the known triple "Confucius-Q3130-founded-Confucianism" from Wikipedia to automatically annotate sentences containing "Confucius" and "Confucianism" in the corpus as instances of the "founded" relationship. By setting a context window for co-occurring entity pairs, weakly annotated samples are generated and used in training alongside the high-quality manually annotated samples.

Entity and relation extraction uses the Span-based Joint Extraction model, which unifies entity recognition and relation prediction into a span-based classification problem. Given a RoBERTa-encoded sentence vector representation $H \in R^{n \times d}$, where n is the sequence length and d is the hidden layer dimension, the entity category probability is calculated for all possible text spans (i, j) [22-24]:

$$P_{ent}(i, j) = Softmax(W_{ent}[H_i; H_j] + b_{ent}) \quad (6)$$

Relation prediction is done by computing relation vectors on pairs of identified entities:

$$P_{rel}(s, o) = Softmax(W_{rel}[h_s; h_o] + b_{rel}) \quad (7)$$

h_s and h_o are the average pooling results of the entity's first and last token vectors, and W_{ent} and W_{rel} are task-specific weight matrices.

The joint training loss function is:

$$L = L_{ent} + \zeta L_{rel} \quad (8)$$

L_{ent} and L_{rel} are the cross-entropy losses for entity recognition and relationship prediction, respectively, and ζ is the balance coefficient. For synonymous or near-synonymous relationship labels, a sentence vector clustering method based on the relationship description text is used

to project the relationship vectors into a unified semantic space. K-means clustering and manual verification are used to form a standard relationship set to ensure cross-corpus consistency. After standardization, the entity pairs and their relationships are output as an initial set of triples:

$$G_0 = \{(h_i, r_i, t_i, c_i) | i = 1, \dots, N\} \quad (9)$$

$c_i \in [0, 1]$ represents the confidence of the model prediction. To improve the quality of subsequent Graphormer knowledge completion, only the triples of $c_i \geq 0.7$ are retained. The initial triple size statistics are shown in Table 2:

TABLE 2. Initial Triple Size

Triple type	Average Confidence	Number of covered entities	Number of coverage relationships
Person-Relationship-event	0.82	321	7
Person-Relationship-Work	0.85	293	3
School-Relationship-Person	0.80	90	2

3) Graphormer-Driven Knowledge Completion and Global Structure Encoding

The initial KG is input into the Graphormer model, and global structure modeling is achieved through node distance encoding and edge feature encoding. To substantially guide the Graphormer encoding mechanism through the Confucian inheritance model, a three-layer mapping mechanism—structural, semantic, and attentional—is established. In structural mapping, the Confucian figures, works, factions, event transmission chains, and the "ritual system, textiles, and crafts" ternary mapping are transformed into nodes and edge constraints of the graph. In the semantic mapping, during edge feature encoding, learnable embedding vectors are assigned to different relation types, and cultural rationality weights are introduced to adjust path scores. In the attention mapping, within the Graphormer spatial bias, the shortest path distance not only reflects graph distance but also implicitly contains the number of generations of inheritance. Attention response is enhanced based on Shortest-Path Distance (SPD) bias values. In the link prediction task, missing entity relationships are completed, high-dimensional node and subgraph vector representations are generated, and a multi-hop reasoning mechanism guided by graph attention is constructed. As shown in Figure 1:

G_0 is considered a heterogeneous graph with relationship types and directions. To facilitate unified encoding, the relationship type $r \in R$ is embedded in the vector $e_r \in R^{d_r}$, and the direction ($h \rightarrow t/t \rightarrow h$) is encoded as $e_{\rightarrow}, e_{\leftarrow}$. The initial node feature x_v is obtained by concatenating the output entity type embedding, the name subword average pooling vector, and statistical attributes (frequency, time distribution), and then unified into $h_v^{(0)} \in R^d$ after linear transformation. Graphormer's attention applies a learnable

spatial-structural bias b_{ij} , which is composed of three parts [25,26]:

$$b_{ij} = b_{ij}^{SPD} + b_{ij}^{EDGE} + b_{ij}^{CENT} \quad (10)$$

The SPD is mapped via discrete embedding table E_{spd} :

$$b_{ij}^{SPD} = a^\top E_{spd}[\min(\text{dist}(i, j), K)] \quad (11)$$

Edge types and directions are mapped via Table E_{edge} :

$$b_{ij}^{EDGE} = a_e^\top E_{edge}[r(i, j), \text{dir}] \quad (12)$$

The node centrality is mapped to Table E_{cent} :

$$b_{ij}^{CENT} = a_c^\top (E_{cent}[\text{cent}(i)] + E_{cent}[\text{cent}(j)]) \quad (13)$$

Designing a multi-hop reasoning mechanism guided by graph attention. For intra-layer attention and update, the l -th layer self-attention adopts biased scoring [27,28]:

$$e_{ij}^{(l)} = \frac{q_i^{(l)}(k_j^{(l)})^\top}{\sqrt{d}} + b_{ij} \quad (14)$$

$$\alpha_{ij}^{(l)} = \text{softmax}_j(e_{ij}^{(l)}) \quad (15)$$

$$h_i^{(l+1)} = W_o \sum_j \alpha_{ij}^{(l)} v_j^{(l)} \quad (16)$$

$[q_i^{(l)}, k_i^{(l)}, v_i^{(l)}]$ is obtained by linearly mapping $h_i^{(l)}$, with residuals and normalization between layers, resulting in a stack of L layers to obtain $z_v = h_v^{(L)}$. The Graphormer input and bias encoding components are shown in Table 3:

In knowledge completion, the goal is to estimate the probability $s(h, r, t)$ of triples existing and add high-confidence candidates to the graph under the verification threshold τ^* . Relation-specific bilinear scoring is used [29,30]:

$$\phi(h, r, t) = z_h^\top R_r z_t \quad (17)$$

$$s(h, r, t) = \sigma(\phi(h, r, t)) \quad (18)$$

$R_r \in R^{d \times d}$ is a relation-specific parameter, and σ is a sigmoid. Minimizing the binary cross entropy for positive and negative samples:

$$L_{LP} = -\frac{1}{|B|} \sum_{(h,r,t,y) \in B} [y \log s(h, r, t) + (1 - y) \log(1 - s(h, r, t))] \quad (19)$$

At the same time, mask modeling is applied to enhance the discrimination of relationship types. The edge set ϵ_m is randomly masked, and its relationship type $\hat{r}$ is predicted using z_h and z_t on Graphormer, and cross entropy L_{mem} is used; the total loss is [31,32]:

$$L = L_{LP} + \lambda_{mem} L_{mem} + \lambda_{L2} \|\Theta\|_2^2 \quad (20)$$

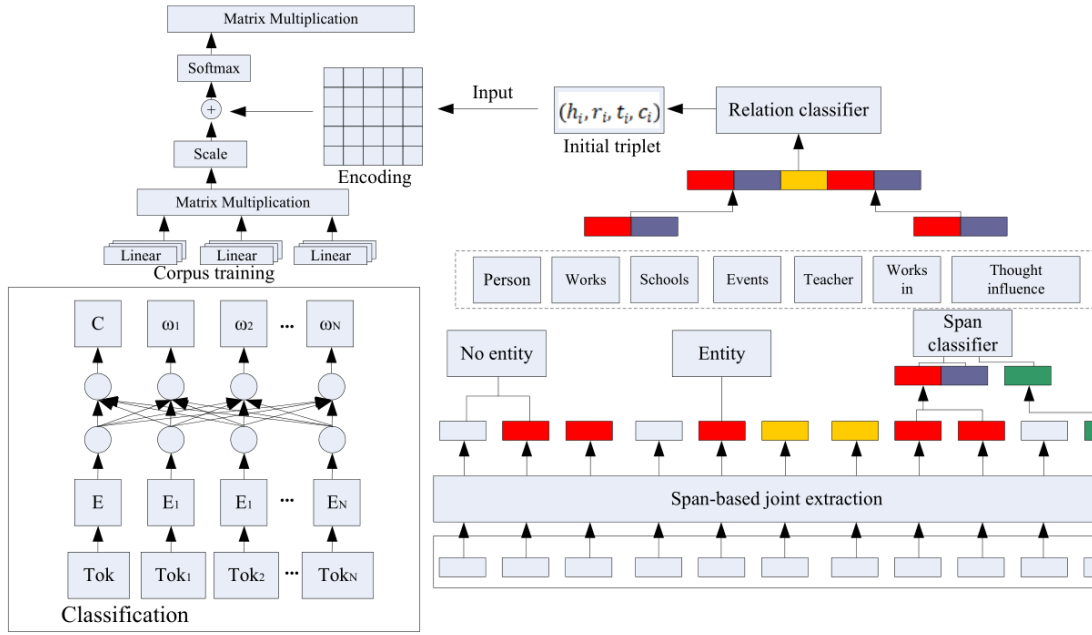


FIGURE 1. Graphormer Model

TABLE 3. Graphormer Input and Bias Encoding Components

Coding components	Symbol	Calculation method	Dimension/Description
Initial characteristics of nodes	$h_v^{(0)}$	Text embedding + type + statistical feature linear transformation	d
SPD	b_{ij}^{SPD}	$E_{spd}[dist(i, j)]$	SPD bias d_b
EDGE	b_{ij}^{EDGE}	$E_{edge}[r, dir]$	Edge bias d_b
CENT	b_{ij}^{CENT}	$E_{cent}[cent]$	Neutral horizontal bias d_b
Multi head attention	–	Scale dot product with bias	Head count H

To reduce gradient degradation caused by easy samples, it uses relational distribution adaptive sampling $p(t|r)$ and reweighting of hard negative samples. It increases the weight $w_- = softplus(\phi)$ of negative samples that meet $\phi(h, r, t)$ higher, stabilizes convergence, and improves recall. It selects the threshold by maximizing the Youden index:

$$\tau^* = \arg \max_{\tau} (TPR(\tau) - FPR(\tau)) \quad (21)$$

And imposing the minimum precision constraint of the relationship dependency. Finally, the completed graph is obtained:

$$G^+ = G_0 \cup \{(h, r, t) | s(h, r, t) \geq \tau^*\} \quad (22)$$

In global structure encoding and subgraph representation, the z_v after multi-layer GraphOrmer is used as the final node vector for entity retrieval and upstream and downstream question-answering reasoning. To serve the multi-hop retrieval in the question-answering stage, a general subgraph readout independent of the question is defined: the candidate subgraph $S = (V_s, \epsilon_s)$ is aggregated with the relational attention weight β_r :

$$g_S = \sum_{(i,j,r) \in \epsilon_s} \beta_r(z_i \| z_j \| e_r) \quad (23)$$

$$\beta_r = \frac{\exp(\gamma \cdot \pi_r)}{\sum_{r'} \exp(\gamma \cdot \pi_{r'})} \quad (24)$$

π_r is set to the inverse of the relation frequency or obtained by learnable gating, and γ is the temperature coefficient. This readout forms a structured semantic summary without applying the question text, which facilitates subsequent alignment with the question vector.

B. DESIGN OF INTELLIGENT QUESTION ANSWERING SYSTEM

To achieve efficient, accurate, and interpretable access to traditional Chinese cultural knowledge, this paper constructs an intelligent question-answering system based on Graphormer-enhanced KGs. This system uses the Graphormer-enhanced KG representation directly as the core knowledge source for question-answering reasoning. Using a four-stage process consisting of "question semantic encoding—graph structure alignment—path reasoning based on Graphormer embeddings—and answer generation," the system explicitly integrates the high-dimensional

node embeddings and global attention weights generated by Graphormer into the entire question-answering process, enabling collaborative reasoning in semantic and structural spaces. The overall system architecture consists of four core components: a front-end interaction module, a question understanding module, a knowledge retrieval and reasoning module, and an answer generation module, as shown in Figure 2.

The front-end interaction module serves as the direct interface between the system and the user, responsible for receiving user queries and presenting the final answer. This is implemented using a separate front-end and back-end architecture. The front-end is a responsive web interface developed using the Vue.js framework, supporting natural language input, historical session logging, keyword highlighting, and reasoning path visualization. After the user enters a question, the front-end sends the text to the back-end Flask API (Application Programming Interface) service via Hypertext Transfer Protocol Secure. All front-end resources are deployed on an Nginx (engine x) server, supporting cross-platform access.

The core goal of the question understanding module is to transform the user's natural language question into a semantic representation aligned with the KG structure. This is achieved by encoding the input question into a vector representation based on entity type embedding and name subword average pooling. The system first performs text normalization on the input question Q . Then, it uses a domain-tuned pre-trained language model to perform contextual semantic encoding on the question, resulting in a vector representation of the question $q = f_{enc}(Q) \in R^d$. f_{enc} represents the encoder. Using a domain-adaptive NER model and entity linking module, it identifies explicitly mentioned entities E_q from Q and maps them to a graph node set V . At the same time, it uses a relation extraction model to identify a set of potential relation candidates R_q in the question, and retains the confidence scores for subsequent retrieval weighting. After the initial identification of entities and relations, it calculates the semantic similarity between the node embedding matrix and the question vector q :

$$sim(q, z_v) = \frac{q \cdot z_v}{\|q\| \|z_v\|} \quad (25)$$

A candidate node set V_q is retrieved from high to low similarity and then combined with E_q to perform a union fusion, ensuring that the recalled entities cover both explicit mentions and semantic neighbors. Finally, using V_q as the source, a multi-hop restricted expansion is performed on the completed G^+ . The system prioritizes expansion of high-weight paths within the multi-hop limit, thereby obtaining the multi-hop knowledge subgraph most relevant to the question semantics.

The knowledge retrieval and reasoning module, leveraging Graphormer's enhanced multi-hop reasoning capabilities, performs question-aware path search and optimization on retrieved candidate subgraphs. The question vector rep-

resentation q and the subgraph node embedding z_v are fed into the Graphormer encoder, allowing attention weights to be explicitly influenced by question semantics, thereby strengthening question-relevant path features. During reasoning, the system calculates the composite score of all possible paths $Path = \{path_1, path_2, \dots, path_n\}$:

$$Score(path) = \zeta_1 \cdot Rel(path, q) + \zeta_2 \cdot c(path) \quad (26)$$

$Rel(path, q)$ represents the degree of match between the path and the question semantics, $c(path)$ is the path confidence, which is provided by the scoring function $s(h, r, t)$ in the knowledge completion phase, and ζ_1 and ζ_2 are balance coefficients. After path rescoring and denoising, only high-confidence paths are retained as the final reasoning basis, achieving the construction of explainable multi-hop reasoning links.

The answer generation module generates answers that conform to natural language conventions based on optimized reasoning paths and their terminal nodes. The entities, relationships, and attributes within the reasoning chain are organized into a set of structured answer candidates. For explanatory questions, a template-filling mechanism is used, with several pre-set sentence templates. The retrieved triples and paths are then filled in to generate coherent sentences. For complex comparative questions, contrast generation logic is applied to extract the common adjacent nodes and difference relationships between two entities to generate structured comparative answers. For factual questions, a list of entities is directly returned. During the generation process, the system retains traceable information about the reasoning path, allowing users to view the reasoning behind the answer and the relevant knowledge triples.

III. EXPERIMENTAL ANALYSIS OF KNOWLEDGE GRAPHS AND INTELLIGENT QUESTION ANSWERING SYSTEMS

A. EXPERIMENTAL SETUP

To verify the effectiveness and reliability of this Graphormer-enhanced KG modeling and intelligent question-answering system for traditional Chinese culture, it conducts an experimental analysis. The dataset used in this experiment primarily comes from the digitized classics library and semi-structured historical and cultural materials preprocessed in Section "Corpus Preprocessing and Normalization". The digitized classics library includes open ancient book data provided by the "Chinese Ancient Books Resource Library" of the National Library of China, covering core classics such as classics, history, philosophy, and anthologies, as well as annotations from various dynasties. The semi-structured historical and cultural materials are derived from Chinese Wikipedia entries and abstracts of selected academic entries on CNKI and the Wanfang Data Platform. After cleaning and entity alignment, the experimental corpus is finally formed.

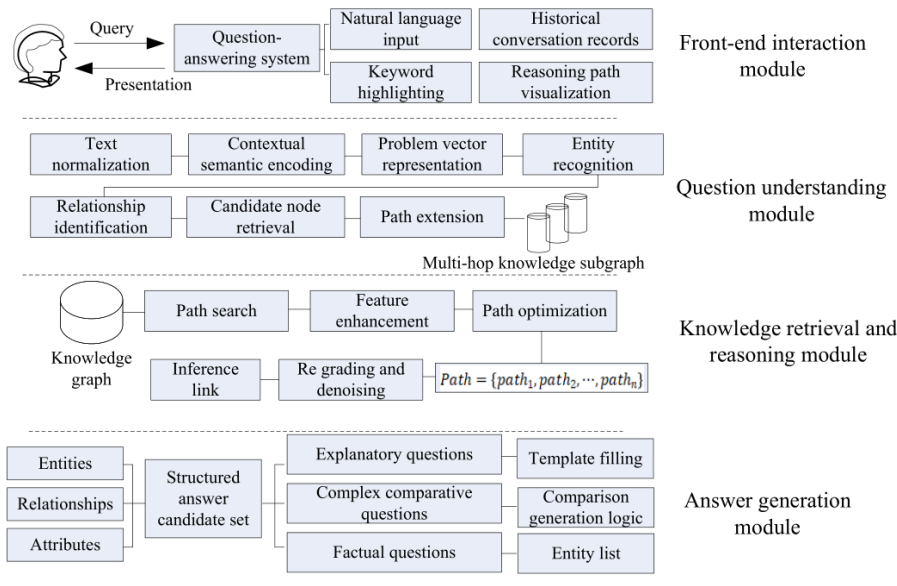


FIGURE 2. Intelligent Question-Answering System Architecture

To evaluate the model's multi-hop reasoning ability, this paper constructs a hierarchical multi-hop question-answering test set, divided into four categories from Hop=1 to Hop=4 based on the number of hops required for the reasoning path. Hop represents the number of edges required to reach the answer entity from the starting entity explicitly or implicitly mentioned in the question via the shortest effective reasoning path in the knowledge graph. Three domain experts manually constructed the question set based on the knowledge ontology of Confucian heritage and textile technology, covering three semantic scenarios: explanatory, comparative, and factual. All questions were verified for their shortest path hop count through subgraph traversal, and ambiguous samples with shorter alternative paths were excluded. The question-answering accuracy (Answer Accuracy@Hop=k) at each hop count is used as the core indicator, representing the proportion of times the model's predicted answer completely matches the standard answer.

All experiments are implemented using the PyTorch framework. The Graphormer model uses a multi-layer Transformer encoder, combining global structure encoding with a distance embedding mechanism. The main parameter settings are shown in Table 4:

TABLE 4. Model Parameter Settings

Classification	Parameter	Specifications
Model structure parameters	Hidden size	256
Model structure parameters	Number of Transformer layers	12
Model structure parameters	Number of attention heads	8
Figure Structure Modeling Parameters	Node position encoding method	SPD
Figure Structure Modeling Parameters	Edge feature embedding dimension	64
Training and optimizing parameters	Dropout ratio	0.1
Training and optimizing parameters	Learning rate	1.00E-04
Training and optimizing parameters	Batch size	64
Training and optimizing parameters	Optimizer	Adam with weight decay (weight decay 0.01)
Training and optimizing parameters	Epochs	50

The corpora used in this study were all sourced from the Ancient Books Database of the National Library of China, Chinese Wikipedia, CNKI, and Wanfang. Preprocessing tools included an ancient text segmentation model, a punctuation restoration model, a variant character mapping table, and entity normalization rules. Given the unique nature of the ancient texts involved in this study, direct sharing of the complete original corpus is practically limited. However, the configuration parameters for the preprocessing rules, word segmentation, and punctuation recovery model can be obtained by contacting the corresponding author, subject to relevant data usage policies.

To verify the advantages of Graphormer in knowledge completion and question-answering reasoning, it is evalu-

ated from two dimensions: KG completion and intelligent question-answering. In this experiment, a unified KG construction process is used and compared with mainstream methods:

(1) TransE (Translating Embeddings): a classic knowledge representation method based on vector translation, used for link prediction and knowledge completion [33].

(2) R-GCN (Relational Graph Convolutional Networks): an extension of GCN on heterogeneous relational graphs, used for relation-aware graph structure learning [34].

(3) K-BERT (Knowledge-Based BERT): a knowledge representation method that directly encodes triples using a pre-trained language model [35].

B. EXPERIMENTAL RESULTS

1) Knowledge Graph Completion Task

In the KG completion task, to verify the effectiveness of the proposed method in the KG construction phase, it evaluates entity recognition and relationship extraction. Entity recognition is a prerequisite for KG construction and directly determines the completeness and accuracy of subsequent knowledge units. It uses entity recognition accuracy and F1 score as initial evaluation metrics and compares them with existing methods to verify the effectiveness of our proposed method in semantic extraction, as shown in Figure 3.

The results in Figure 3 show that the proposed Graphormer model significantly outperforms the comparison methods in entity recognition. In Figure 3A, the average recognition accuracy of our method on the four categories of cultural entities reaches 0.929, while the average recognition accuracy of the TransE, R-GCN, and K-BERT methods on the four categories of entities is 0.794, 0.833, and 0.875, respectively. In Figure 3B, the average recognition F1 value of our method on the four categories of cultural entities reaches 0.908, while the average recognition F1 values of the TransE, R-GCN, and K-BERT methods on the four categories of entities are 0.740, 0.790, and 0.850, respectively. Judging from the specific comparison results, the method in this paper performs more ideally. Graphormer effectively models the long-range semantic dependencies and complex topological structures in the inheritance of Confucian thought by applying node degree, shortest path and edge type encoding; its global attention mechanism can integrate the contextual information of the inheritance of Confucian thought and enhance the ability to recognize entities; compared with R-GCN that only relies on local aggregation and TransE that relies on static embedding, Graphormer has stronger semantic resolution and generalization performance in heterogeneous cultural graphs. Although K-BERT combines a text pre-training model, it lacks explicit global graph structure modeling and, therefore, is still inferior to Graphormer in complex relationship extraction tasks.

In the KG completion experiment, in addition to the local performance of entity recognition and relation extraction, to evaluate the model's performance in overall reasoning and

link prediction tasks, it further uses the Hits@k ($k=1, 3, 10$) metric to compare the effectiveness of different models in knowledge completion tasks by measuring the coverage of the correct answer in the top k candidate results. This is shown in Figure 4:

Figure 4 shows the overall performance differences among the four models under different evaluation metrics: Graphormer consistently performs best, achieving 0.487, 0.602, and 0.659 on Hits@1, Hits@3, and Hits@10, respectively. K-BERT follows, achieving 0.423, 0.539, and 0.628 on Hits@1, Hits@3, and Hits@10, respectively. R-GCN outperforms traditional TransE overall, but still has shortcomings in handling complex relationship combinations, with specific results of 0.358, 0.487, and 0.583, respectively. TransE, relying on a simple vector translation assumption, performs worst in complex semantics and high-order path reasoning, with results of 0.312, 0.428, and 0.522, respectively. Graphormer uses graph attention and global position encoding to effectively alleviate the performance degradation of traditional methods under multi-hop reasoning and sparse relations in the context of Confucian thought inheritance, enabling it to maintain higher recall coverage and stability under different k thresholds.

To evaluate the model's ability to discriminate fuzzy relationships, it applies the AUC (Area Under Curve) metric for link prediction for comparative analysis. The AUC value calculates the model's ability to distinguish positive and negative samples at different thresholds, further validating the stability advantage of our method in complex relationship reasoning. This is shown in Figure 5.

The results in Figure 5 show that Graphormer performs best in terms of the overall discriminative ability of link prediction, followed by K-, R-GCN, and TransE. In Figure 5A, TransE has an AUC value of 0.753. Its translation hypothesis has the weakest ability to model complex semantic patterns, so its AUC is the lowest; in Figure 5B, the AUC value of R-GCN is 0.806. Although it can model local structure and improve the capture of neighborhood consistency, it has limited ability to distinguish long-range dependencies and heterogeneous relationships; in Figure 5C, the AUC value of K-BERT is 0.834. It enhances the semantic understanding of the inheritance of Confucianism through language models and improves text-based discrimination, but the lack of explicit global graph bias results in its performance in complex path judgment still being inferior to the model in this paper; in Figure 5D, the AUC value of Graphormer in this paper is 0.854. Graphormer combines the global self-attention of Transformer with graph structure bias, which can simultaneously capture long-range dependencies and heterogeneous relationship signals, making it possible to effectively identify real links in Confucianism KG with sparse multi-hop semantic associations and complex paths; the link prediction target and mask modeling applied during training improve the model's robustness in distinguishing relationship types, thereby reducing the false positive rate of false positive samples.

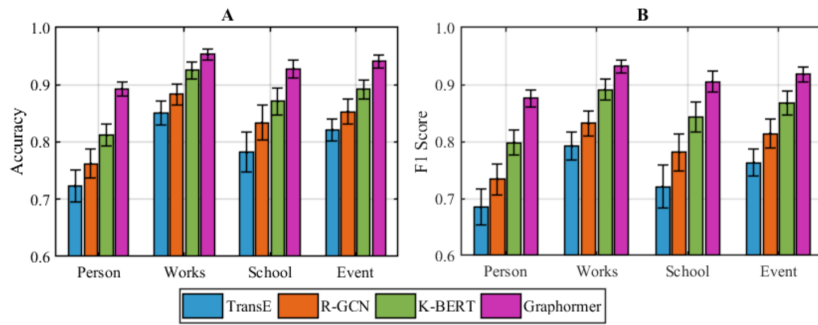


FIGURE 3. Comparison of entity recognition. Figure 3A shows the accuracy; Figure 3B shows the F1 value

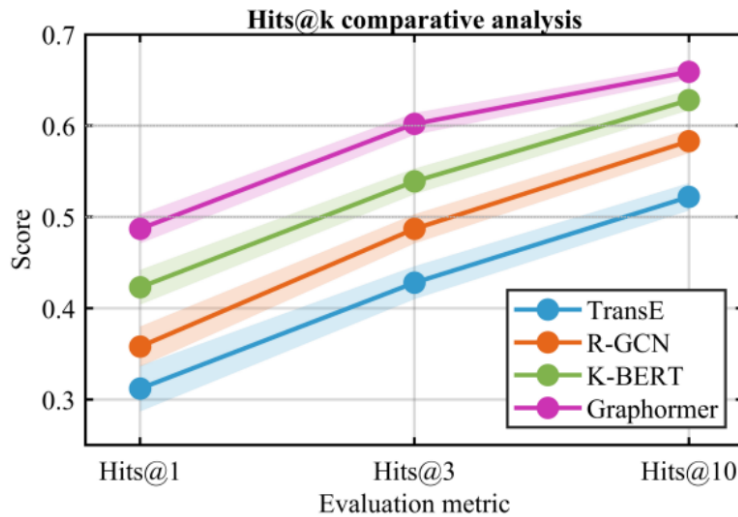


FIGURE 4. Hits@k Comparison

2) Intelligent Question Answering Task

The Mean Rank metric is used to evaluate the performance of different methods in interpretative, complex comparative, and factual questions, as well as in different text lengths. Mean Rank is the average ranking of correct answers, with smaller values being better. By calculating the average ranking and fluctuation of each model in each length interval, the retrieval performance of this method in intelligent question-answering tasks under complex texts is verified. This is shown in Figure 6.

As shown in Figure 6, the Mean Rank increases with question length. Graphormer performs best across different text lengths and question types. In Figures 6A, 6B, and 6C, when the text length of explanatory, complex comparative, and factual questions exceeds 25 characters, their Mean Ranks are 58.6, 61.0, and 62.7, respectively. This demonstrates their superior accuracy in processing local semantics and entity relationships, making them particularly suitable for multi-hop reasoning in Chinese culture. K-BERT is slightly inferior to our proposed method. Its knowledge-enhanced semantic fusion effectively mitigates information loss to a certain extent, but when the text length exceeds 25

characters, its Mean Ranks are 64.7, 66.7, and 68.6. R-GCN and TransE achieve Mean Ranks of 70.1, 71.2, and 74.7, and 77.3, 79.2, and 79.1, respectively. This demonstrates that simple vector embeddings or traditional graph convolutions struggle to handle long-chain reasoning or complex semantic dependencies. Graphormer leverages global self-attention to capture long-range dependencies, enabling it to maintain high semantic understanding capabilities even in long-text questions.

To verify the system's effectiveness in long-range semantic dependency modeling, the model's performance was evaluated on a hierarchical multi-hop question-answering test set. As shown in Figure 7, accuracy was calculated separately for Hop=1 to Hop=4 to ensure that the evaluation focused on multi-hop reasoning ability rather than single-hop fact retrieval.

Figure 7 shows significant differences in the performance of different models in the multi-hop question answering task, with accuracy decreasing as the number of hops increases. In Figures 7A, 7B, and 7C, the Graphormer model maintains accuracy of 0.773, 0.715, and 0.667 at Hop=4, demonstrating its stable performance in multi-hop

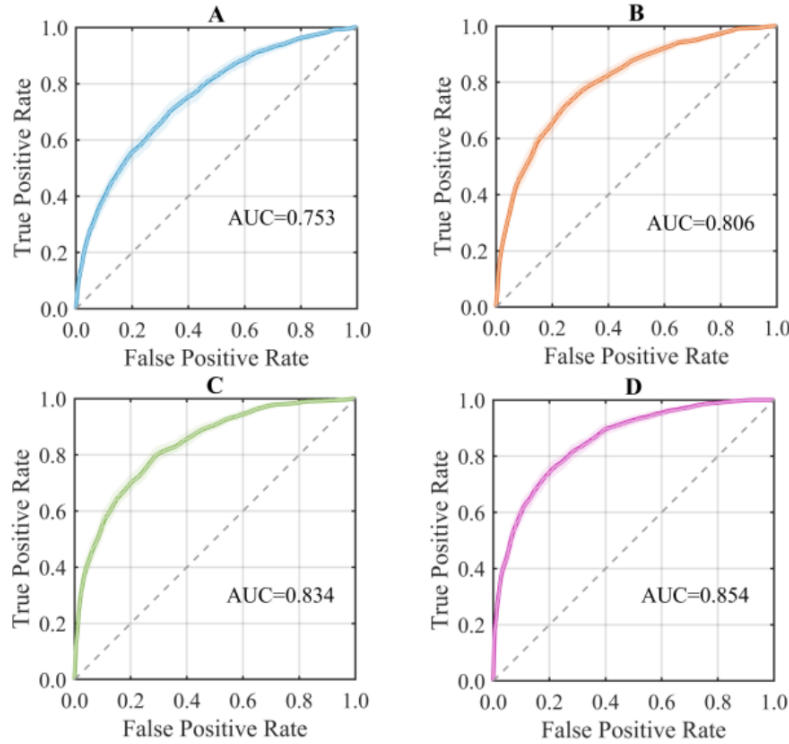


FIGURE 5. AUC comparison of Link Prediction. Figure 5A Shows the AUC of TransE; Figure 5B Shows the AUC of R-GCN; Figure 5C Shows the AUC of K-BERT; Figure 5D Shows the AUC of Graphormer

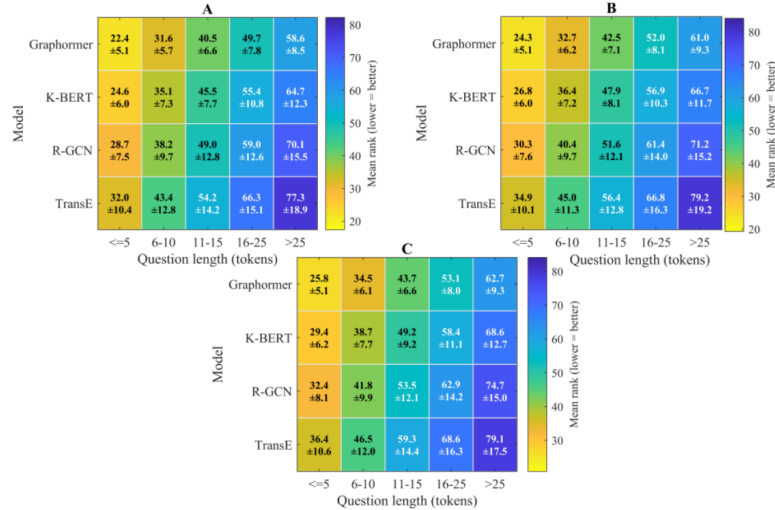


FIGURE 6. Mean Rank Comparison. Figure 6A Shows the Mean Rank for Explanatory Questions; Figure 6B Shows the Mean Rank for Complex Comparative Questions; Figure 6C Shows the Mean Rank for Factual Questions

reasoning. K-BERT achieves accuracy of 0.742, 0.684, and 0.629, respectively, at Hop=4; R-GCN achieves accuracy of 0.682, 0.631, and 0.583, respectively, at Hop=4; and TransE performs the weakest, achieving accuracy of 0.635, 0.589, and 0.547, respectively, at Hop=4, demonstrating the limitations of traditional embedding models when handling direct factual questions. Overall, Graphormer demonstrates superior semantic understanding and comprehensive reasoning capabilities. By enhancing structured representations

and long-term dependency modeling, it can better cope with the complexity introduced by the increasing number of hops, thereby maintaining accurate question answering performance.

To comprehensively evaluate the performance of intelligent question-answering systems in generative answering tasks, it applies the BLEU-4 (Bilingual Evaluation Understudy-4) and ROUGE-L (Recall-Oriented Understudy for Gisting Evaluation-Longest Common Subsequence) met-

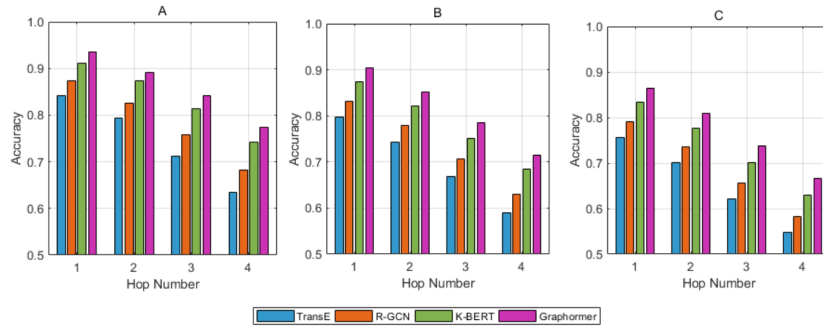


FIGURE 7. Comparison of Multi-hop Question Answering Accuracy. Figure 7A Shows the Accuracy of Explanatory Questions; Figure 7B Shows the Accuracy of Complex Comparative Questions; Figure 7C Shows the Accuracy of Factual Questions

rics. BLEU-4 primarily measures the accuracy of answers in terms of fluency and word order matching, while ROUGE-L captures the semantic coverage between generated content and reference answers through the longest common subsequence. The results are shown in Figure 8:

In Figure 8, the Graphormer model performs better in both BLEU-4 and ROUGE-L indicators. In Figure 8A, Figure 8B, and Figure 8C, its BLEU-4 reaches 0.578, 0.521, and 0.607, respectively; and its ROUGE-L reaches 0.656, 0.589, and 0.682, respectively. This shows that the fluency and semantic coverage of the answers generated by Graphormer are higher, and it can organize and express information more accurately. The TransE model has the lowest BLEU-4, 0.421, 0.372, and 0.453, respectively; and ROUGE-L is 0.502, 0.439, and 0.532, respectively. R-GCN achieves BLEU-4 results of 0.468, 0.415, and 0.498, respectively, and ROUGE-L results of 0.548, 0.482, and 0.578, respectively. K-BERT achieves BLEU-4 results of 0.523, 0.468, and 0.554, respectively, and ROUGE-L results of 0.602, 0.536, and 0.634, respectively. Overall, Graphormer captures the multi-hop semantic paths in Confucian thought through a global attention mechanism and, in combination with the KG structure, generates logically coherent and rich natural language responses. It demonstrates stable generation capabilities across a wide range of question types.

3) Ablation Experiment

To evaluate the effectiveness of each structural encoding component in Graphormer, ablation experiments were conducted. Node degree centrality bias, SPD bias, and Edge Encoding were removed sequentially; and the results were compared with the complete system presented in this paper. The results are shown in Table 5.

TABLE 5. Ablation Experiment Results

Model	Hits@10	AUC
Graphormer (Full model)	0.659	0.854
Remove node degree centrality bias	0.582	0.812
Remove SPD bias	0.601	0.823
Remove edge encoding	0.627	0.839

Table 5 shows that the SPD bias has the greatest im-

pact on performance, with its absence causing Hits@10 to drop to 0.601. Edge feature encoding is the second most important factor, as it carries the semantic distinguishing ability of Confucian relationship types. Although degree centrality contributes less, it improves the stability of node representations when identifying core figures in the lineage.

4) Baseline Contrast Experiment

Two classic graph neural networks, GCN and GAT, are introduced as lightweight references. Both use the same input node features and training strategies as Graphormer to eliminate feature engineering interference. All models are fairly compared on a uniformly partitioned training/validation/test set. Evaluation metrics include Hits@10, multi-hop question answering accuracy, and inference parameter count. The results are shown in Table 6.

Table 6 shows that Graphormer has a larger number of parameters and higher inference latency, mainly due to its global attention computation and structural bias caching. However, it exhibits significant performance advantages in long-range dependency modeling, proving that its computational complexity is reasonable in the context of high semantic complexity. Especially when dealing with sparse, nonlocal relations, standard graph neural networks struggle to capture global path signals due to their limited receptive field, while Graphormer's global attention mechanism and structural bias mechanism jointly overcome this deficiency.

IV. CONCLUSION

This paper utilizes Graphormer to study the construction of a specialized Knowledge Graph (KG) and intelligent question-answering system. By combining global attention with graph structure encoding, it models the long-range dependencies and complex topological structures in the inheritance of Confucianism. In the KG completion task, Graphormer achieves entity recognition accuracy and F1-score of 0.929 and 0.908, respectively. In the Hits@10 metric, Graphormer achieves 0.659, and in the link prediction task, its AUC reaches 0.854, demonstrating its superior discriminative power for complex semantics and sparse paths. In intelligent question answering, the Graphormer-based intelligent question answering system maintains su-

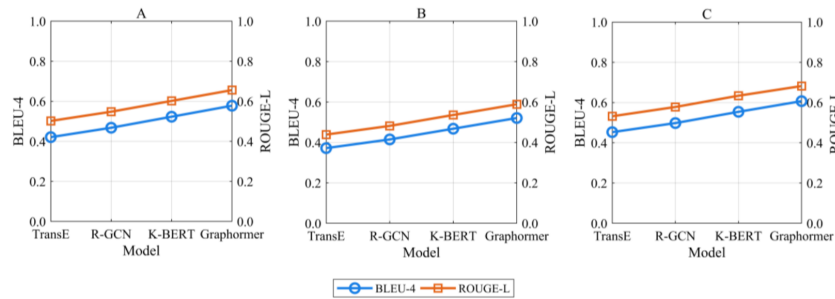


FIGURE 8. Comparison of BLEU-4 and ROUGE-L. Figure 8A Shows BLEU-4 and ROUGE-L on Explanatory Questions; Figure 8B Shows BLEU-4 and ROUGE-L on Complex Comparative Questions; Figure 8C Shows BLEU-4 and ROUGE-L on Factual Questions

TABLE 6. Performance Comparison with Baseline

Model	Hits@10	Explanatory Questions (Hop=4)	Complex Comparative Questions (Hop=4)	Factual Questions (Hop=4)	Inference latency (ms/query)	Parameter Count (M)
GCN	0.498	0.582	0.521	0.503	2.1	24.6
GAT	0.537	0.624	0.573	0.551	2.8	31.8
Graphormer	0.659	0.773	0.715	0.667	18.6	89.3

prior performance under multi-hop reasoning. In complex comparative question answering, BLEU-4 and ROUGE-L scores reach 0.521 and 0.589, respectively, demonstrating their combined fluency and semantic coverage. This paper integrates Graphormer’s traditional cultural KG with an intelligent question-answering system to effectively address the long-range dependencies and complex relational reasoning between Confucian thought and textile heritage. This provides reliable support for the intelligent transmission of specialized Chinese cultural textile knowledge.

AUTHOR CONTRIBUTIONS

Shengxue Lei designed, collected and analyzed the data, and drafted the manuscript. Shengxue Lei conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Shengxue Lei participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research was supported by:

1. The 2023 Philosophy and Social Science Planning Research Project in Luohe City, titled "Agricultural Culture in Shuowen Jiezi" (No. SKGH-2023-109).
2. The 2023 Science and Technology Innovation Project of Luohe Medical College, titled "Food Culture in Shuowen Jiezi" (No. 2023SK03).

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] D. Sun, Y. Zheng, and C. Huo, "A Study on the Dissemination Paths of China's Excellent Traditional Culture in the Context of Intelligence," *Chinese Traditional Culture*, vol. 12, no. 3, pp. 474-479, 2024, doi: 10.12677/cnc.2024.123075.

[2] J. Yang and M. Zhu, "An Analysis of the Paths of Artificial Intelligence Empowering the Education of China's Excellent Traditional Culture," *College and Job*, vol. 14, no. 4, pp. 45-51, 2025, doi: 10.12677/ve.2025.144152.

[3] Y. Cui, J. Shen, J. Wu, and M. Lv, "Linking past insights with contemporary understanding: An ontological and knowledge graph approach to the transmission of ancient Chinese classics," *Heritage Science*, vol. 12, no. 1, pp. 382-399, 2024, doi: 10.1186/s40494-024-01504-x.

[4] B. Bai and W. Hou, "The application of knowledge graphs in the Chinese cultural field: The ancient capital culture of Beijing," *Heritage Science*, vol. 11, no. 1, pp. 77-93, 2023, doi: 10.1186/s40494-023-00922-7.

[5] L. Wang, F. N. Mohamed, and Z. Liu, "Lanjiefu in Wenzhou City under the Perspective of Digital Humanities: Constructing a Knowledge Graph to Empower the Cultural Resource Base," *International Journal of Education and Humanities*, vol. 5, no. 3, pp. 404-414, 2025, doi: 10.58557/ijeh.v5i3.322.

[6] Z. Qian, T. Chen, Y. Xu, Q. Li, and D. Zhang, "Research on the Construction and Application of Knowledge Graphs for Ancient Classical Dictionaries," *Library Journal*, vol. 42, no. 388, pp. 82-88, 2023, [Online]. Available: <https://www.libraryjournal.com.cn/CN/>.

[7] L. Liu, Y. Qi, C. Liu, W. Li, and D. Wang, "Research on Ancient Book Citations and the Construction of a Full-Text Knowledge Base Based on Computational Humanities," *Journal of the China Society for Scientific and Technical Information*, vol. 42, no. 12, pp. 1498-1512, 2023.

[8] Liu, S, H. Yang, J. Li, and S. Kolmanic, "Preliminary study on the knowledge graph construction of Chinese ancient history and culture," *Information*, vol. 11, no. 4, pp. 186-206, 2020, doi: 10.3390/info11040186.

[9] W. Chansanam, Y. Jaroenruen, N. Kaewboonma, and K. Tuamsuk, "Culture knowledge graph construction techniques," *Education for Information*, vol. 38, no. 3, pp. 233-264, 2022, doi: 10.3233/EFI-220028.

[10] L. Xu, L. Lu, and M. Liu, "Construction and application of a knowledge graph-based question answering system for Nanjing Yunjin digital resources," *Heritage Science*, vol. 11, no. 1, pp. 222-238, 2023, doi: 10.1186/s40494-023-01068-2.

[11] H. Yuan, Y. Li, B. Wang, K. Liu, and J. Zhang, "Knowledge graph-based intelligent question answering system for ancient Chinese costume heritage," *Npj Heritage Science*, vol. 13, no. 1, pp. 198-213, 2025, doi: 10.1038/s40494-025-01776-x.

[12] M. Zhang, G. Geng, S. Zeng, and H. Jia, "Knowledge graph completion for the Chinese text of cultural relics based on bidirectional encoder representations from transformers with entity-type information," *Entropy*, vol. 22, no. 10, pp. 1168-1182, 2020, doi: 10.3390/e22101168.

- [13] J. Liu, X. Ma, L. Wang, and L. Pei, "How Can Generative Artificial Intelligence Techniques Facilitate Intelligent Research into Ancient Books?," *ACM Journal on Computing and Cultural Heritage*, vol. 17, no. 4, pp. 1-20, 2024, doi: 10.1145/3690391.
- [14] Q. Wu, W. Zhao, C. Yang, H. Zhang, F. Nie, H. Jiang, et al., "Sgformer: Simplifying and empowering transformers for large-graph representations," *Advances in Neural Information Processing Systems*, vol. 36, no. 2826, pp. 64753-64773, 2023, doi: 10.48550/arXiv.2306.10759.
- [15] C. Zhou, Q. Yu, L. Zhang, Z. Hu, and M. Zhao, "Overview of Research Progress in Graph Transformers," *Journal of Computer Engineering & Applications*, vol. 60, no. 14, pp. 1-37, 2024, doi: 10.3778/j.issn.1002-8331.2308-0216.
- [16] Y. Wang, J. Liu, W. Wang, J. Chen, X. Yang, L. Sang, et al., "Construction of cultural heritage knowledge graph based on graph attention neural network," *Applied Sciences*, vol. 14, no. 18, pp. 8231-8258, 2024, doi: 10.3390/app14188231.
- [17] W. Hou, B. Bai, and C. Cai, "CR-TransR: A knowledge graph embedding model for cultural domain," *ACM Journal on Computing and Cultural Heritage*, vol. 17, no. 1, pp. 1-11, 2024, doi: 10.1145/3625299.
- [18] D. Wang, K. Tang, J. Zeng, Y. Pan, Y. Dai, H. Li, et al., "MM-transformer: A transformer-based knowledge graph link prediction model that fuses multimodal features," *Symmetry*, vol. 16, no. 8, pp. 961-973, 2024, doi: 10.3390/sym16080961.
- [19] J. Yang, Z. Liu, S. Xiao, C. Li, D. Lian, S. Agrawal, et al., "Graphformers: Gnn-nested transformers for representation learning on textual graph," *Advances in Neural Information Processing Systems*, vol. 34, no. 2206, pp. 28798-28810, 2021, doi: 10.48550/arXiv.2105.02605.
- [20] Z. Shou, Y. Liu, D. Li, J. Mo, and H. Zhang, "A Bloom Cognitive Hierarchical Classification Model for Chinese Exercises Based on Improved Chinese-RoBERTa-wwm and BiLSTM," *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 3, pp. 479-489, 2024, doi: 10.14569/ijacsa.2024.0150349.
- [21] P. Feng, B. Zhang, L. Yang, and S. Feng, "Entity-Alignment Interaction Model Based on Chinese RoBERTa," *Applied Sciences*, vol. 14, no. 14, pp. 6162-6176, 2024, doi: 10.3390/app14146162.
- [22] C. Ozkurt, "Comparative Analysis of State-of-the-Art Q&A Models: BERT, RoBERTa, DistilBERT, and ALBERT on SQuAD v2 Dataset," *Chaos and Fractals*, vol. 1, no. 1, pp. 19-30, 2024, doi: 10.69882/adba.chf.2024073.
- [23] F. Cai, Q. Hu, R. Zhou, and N. Xiong, "REEGAT: RoBERTa Entity Embedding and Graph Attention Networks Enhanced Sentence Representation for Relation Extraction," *Electronics*, vol. 12, no. 11, pp. 2429-2444, 2023, doi: 10.3390/electronics12112429.
- [24] E. Hassan, A. S. Talaat, and M. A. Elsabagh, "Intelligent text similarity assessment using Roberta with integrated chaotic perturbation optimization techniques," *Journal of Big Data*, vol. 12, no. 1, pp. 164-196, 2025, doi: 10.1186/s40537-025-01233-3.
- [25] Y. Luo, H. Li, L. Shi, and X. Wu, "Enhancing graph transformers with hierarchical distance structural encoding," *Advances in Neural Information Processing Systems*, vol. 37, no. 1821, pp. 57150-57182, 2024, doi: 10.48550/arXiv.2308.11129.
- [26] X. Jiang and J. Chen, "Heterogeneous graphormer for extractive multimodal summarization," *Journal of Intelligent Information Systems*, vol. 63, no. 2, pp. 355-373, 2025, doi: 10.1007/s10844-024-00886-5.
- [27] J. Liao, X. Zhao, J. Tang, W. Zeng, and Z. Tan, "To hop or not, that is the question: Towards effective multi-hop reasoning over knowledge graphs," *World Wide Web*, vol. 24, no. 5, pp. 1837-1856, 2021, doi: 10.1007/s11280-021-00911-5.
- [28] L. Muller, D. Kusuma, B. Bonet, and C. Morris, "Towards principled graph transformers," *Advances in Neural Information Processing Systems*, vol. 37, no. 4026, pp. 126767-126801, 2024, doi: 10.48550/arXiv.2401.10119.
- [29] Z. Hu, J. You, Z. Yang, X. Li, H. Xie, Q. Li, et al., "A Multi-Hop Graph Reasoning Network for Knowledge-Based VQA," *ACM Transactions on Intelligent Systems and Technology*, vol. 16, no. 3, pp. 1-23, 2025, doi: 10.1145/3724125.
- [30] J. Yu, Q. Su, X. Quan, and J. Yin, "Multi-hop reasoning question generation and its application," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 1, pp. 725-740, 2021, doi: 10.1109/TKDE.2021.3073227.
- [31] S. Sahoo, M. Arriola, Y. Schiff, A. Gokaslan, E. Marroquin, J. Chiu, et al., "Simple and effective masked diffusion language models," *Advances in Neural Information Processing Systems*, vol. 37, no. 4135, pp. 130136-130184, 2024, doi: 10.48550/arXiv.2406.07524.
- [32] D. Mizrahi, R. Bachmann, O. Kar, T. Yeo, M. Gao, A. Dehghan, et al., "4M: Massively multimodal masked modeling," *Advances in Neural Information Processing Systems*, vol. 36, no. 2544, pp. 58363-58408, 2023, doi: 10.48550/arXiv.2312.06647.
- [33] D. Song, F. Zhang, M. Lu, S. Yang, and H. Huang, "DTransE: Distributed translating embedding for knowledge graph," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 10, pp. 2509-2523, 2021, doi: 10.1109/TPDS.2021.3066442.
- [34] S. Yao and D. Pi, "A new relational reflection graph convolutional network for the knowledge representation," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 4, pp. 4191-4200, 2023, doi: 10.1007/s12652-023-04516-w.
- [35] W. Hu, L. He, H. Ma, K. Wang, and J. Xiao, "Kgner: Improving chinese named entity recognition by bert infused with the knowledge graph," *Applied Sciences*, vol. 12, no. 15, pp. 7702-7717, 2022, doi: 10.3390/app12157702.