



OPEN

Construction of knowledge constraints: a case study of 3D structural modeling

Cai Lu¹, Xinran Xu^{2✉} & Bingbin Zhang¹

The uncertainty of structural interpretation complicates the practical production and application of data-driven complex geological structure modeling technology. Intelligent structural modeling excavates and extracts structural knowledge from structural interpretation through human-machine collaboration and combines structural interpretation to form a new model of complex structural modeling guided by knowledge. Specifically, we focus on utilizing knowledge rule reasoning technology to extract topological semantic knowledge from interpretive data and employ knowledge inference to derive structural constraint information from complex geological structure models, thus effectively constraining the 3D geological structure modeling process. To achieve this, we develop a rule-based knowledge inference system that derives theoretical models consistent with expert cognition from interpretive data and prior knowledge. Additionally, we represent the extracted knowledge as a topological semantic knowledge graph, which facilitates computer recognition and allows estimation of intersection lines during 3D geological modeling, resulting in the creation of accurate models. The applicability of our proposed method to various complex geological structures is validated through application tests using real-world data. Furthermore, our method effectively supports the realization of intelligent structure modeling in real working area.

Keywords 3D model, Knowledge reasoning, Interpretation, Complex geological structures

Structural modeling is usually not the ultimate goal but supports the numerical and physical simulation of complex phenomena (such as seismic propagation and fluid migration), depth domain imaging, lithology interpretation, and reservoir modeling. The three-dimensional model of the underground structure visually shows the geometric shape and spatial relationship between underground geological interfaces and geological bodies, such as strata and faults^{1–3}. Based on structural modeling, sequence and attribute modeling can be performed to directly support reserve calculations, well location deployment, and oil and gas development plan formulation. This is one of the most important tasks in underground resource exploration and development^{4,5}. In these areas, the quality of seismic data is often poor, and the relationship between key stratum reflection and stratum contact in the seismic profiles is unclear⁶. The transformation of the multistage tectonic movement led to strong deformation of the rock mass, forming a very complex underground structure. Such cases often complicate the acquisition of high-quality structural interpretations, causing considerable uncertainty in the traditional data-driven structural modeling methods^{7–9}. The presence of uncertainties can make it challenging to establish a direct link between the geometry of a 3D structural model and the corresponding geological interface^{10,11}. Because of the high cost of obtaining interpretation data in actual structural modeling, only limited data can be obtained in a certain research area, which requires more expert experience and interpretation to construct a relatively accurate stratigraphic model^{1,12,13}.

Based on the above questions, Zhan et al. constructed the geometric constraints of a structural model through a knowledge graph, which they used to characterize the constraint relationship among knowledge. When the experts failed to comprehend the structural model, a quality assessment was conducted by modifying the knowledge graph to avoid repeated modeling operations¹⁴. From the graph perspective, knowledge graph is conceptual network and symbolic expression of the physical world. Its nodes represent entities in the real world, and the edges connected by entities represent the semantic links between entities^{15,16}. Knowledge reasoning often involves overcoming two challenges: the difficulty of obtaining data, which results in sparse and uneven distribution of data samples, and the complexity of spatial relationships between structural elements¹⁷. These challenges

¹School of Information and Communication Engineering, University of Electronic and Science Technology of China, Chengdu, China. ²School of Resources and Environment, University of Electronic and Science Technology of China, Chengdu, China. ✉email: xdsxxr@163.com

complicate the accurate expression of geological structures using only text data, and the problem of missing structures is sometimes encountered. To address the above issues, this study proposes a process of constructing constraint information of complex geological structure modeling based on knowledge reasoning. The study aims to establish a large-scale structural modeling knowledge base, which is used to fuse the topological position relationship of geological surface space elements and provide technical support for 3D geological structure modeling technology, as shown in Fig. 1. The construction of a knowledge graph is divided into three stages: (1) conversion of geological data into constraints of a topology knowledge graph, (2) mining of entity and relationship information in geological data through knowledge reasoning, and (3) expert determination of the reliability of the knowledge graph using the wire-frame model.

This study makes the following contributions:

- (1) We propose a framework and process to obtain modeling constraint information by knowledge inference of complex geological structure models (Fig. 1), transform the knowledge of geological experts into knowledge graph data structures that can be recognized by computers, and represent them into wire frame models that can be recognized by experts.
- (2) We construct a common knowledge reasoning rule base in the field of structural modeling and introduce the semantic information of the geological structure into the topological network of the structural model.
- (3) We have demonstrated that our approach can effectively deal with real world work and avoid modifying the original data, but change the knowledge in the knowledge base, and improve the accuracy and robustness of the modeling by introducing expert knowledge.

The remainder of this article follows the following organizational structure. “Materials and methods” section briefly describes the relevant research methods of this study. In “Complex geological structure knowledge reasoning” section, the research methods and processes of constructing constraint knowledge of topological geological modeling through knowledge inference are introduced in detail. “Result” section shows the application of the proposed framework to a field case. “Discussion” section discusses the feasibility of the proposed framework in comparison with existing approaches. The last section will summarize the work of this study.

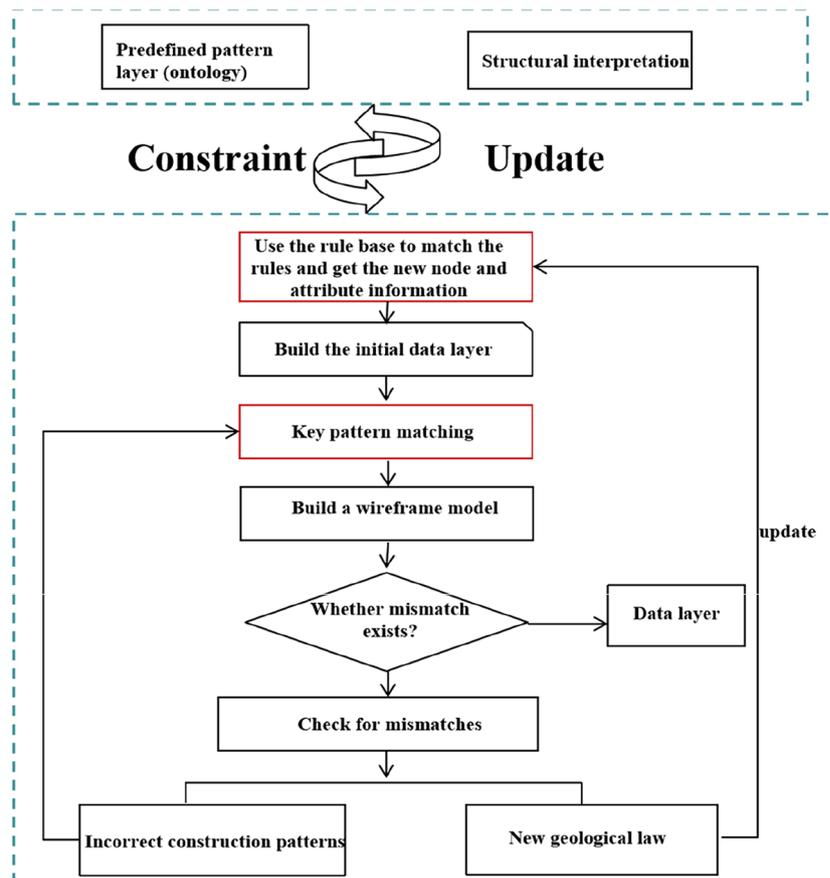


Figure 1. Knowledge constraint construction process.

Materials and methods

As far as we know, only Zhan et al.¹⁴ have introduced knowledge into the study of structural modeling at present, and there is no relevant study on the constraint information of knowledge inference to construct modeling. In this section we will discuss the following two types of research work that are relevant to this paper: knowledge reasoning and explicit modeling¹⁸.

Explicit structural modeling

Explicit interactive modeling is a classical 3D modeling method to reconstruct 3D geological structures using sparse data. When modeling complex regions containing various types of geological structures, the limitations of using a single method show up. In display modeling, people introduce the workflow of building closed geological models by introducing multi-source data information to constrain 3D geological models. First, all geological surfaces are reconstructed, and then the intersections between them are found by cutting each other under constraints. Most of the subsequent explicit modeling approaches use this workflow. Display modeling allows a large number of interactive modifications, adding appropriate underground control constraints according to the experience of the modeler, but the amount of interaction is very large and prone to topological contradictions¹⁹. The square deterministic modeling of structural modeling mainly includes the following methods, namely reservoir seismology methods²⁰, reservoir sedimentological methods²¹ and kriging interpolation methods²². It includes some discrete techniques commonly used in software^{23–26}. Yan et al. shown that spatial explicit models produce better results than non-spatial models, thus showing that space is indeed special in terms of summary²⁷. In order to build a more reasonable 3D geological structure model, but now lack of understanding of the specific structure. In display modeling, where it is difficult to choose a way to integrate all types of information, geological histories are used to combine multiple scalar fields, merely showing the geological interfaces cutting each other. We focus on introducing knowledge reasoning into structural modeling to construct structural models, and introducing earth scientists' cognition into display modeling to provide more comprehensive and intuitive structural models. We provide a workflow to construct modeling constraints by knowledge reasoning, and improve the stability and efficiency of display modeling by introducing knowledge reasoning. Figure 2 provides constraint information for the pattern layer.

Knowledge reasoning of model constraints

The model constraint can be represented by knowledge graph in data form the construction of the knowledge constraints of the structural model starts from the application field and determines the scope of knowledge constraints²⁸. The key to constructing the data layer of the knowledge graph is knowledge reasoning²⁹. New insight is obtained through knowledge reasoning, and the given knowledge graph is inferred based on expert knowledge to determine whether it conforms to cognition for updating the rule base of the pattern layer of the knowledge graph.

The target world can be described based on the relationships between entities. Based on this information, the data are not sufficiently meaningful. Relevant data were combined to form the information. Semantics comprise two components: data and relevance. When describing a structural model, the semantic description encompasses both low-level features, such as geometric properties, and high-level features, such as logical relationships between structural elements. By utilizing basic semantic entities, the structural geological model is divided into geometric units, and their spatial topological relationships are described through the relationships between geometric elements. The semantic reasoning between spatial topological geometries yields the construction process of a knowledge graph³⁰.

The semantic entities in the modeling knowledge graph refer to geometric objects. Their basic semantic entities can be divided into four types: point (0-cell), line (1-cell), surface (2-cell), and body (3-cell). Relation refers to the topological geometric, positional, and compositional relationships between two entities (including the relationships between target entities). An attribute refers to the position and closure of geometric objects in structural modeling¹⁴.

Definition 1 (*Topological semantic knowledge graph*) The creation and update of topological semantic knowledge graphs can be recorded using meta-knowledge, thereby enabling evolution analysis and traceability of complex geological structural knowledge.

$$GeoHfKG = \langle E, V, \{[P_T, Ti]|Ti_O, [P_L, L]|L_O\}|R_{Ti,L} \rangle \leftrightarrow GeoMetaK, \quad (1)$$

where E, V represent the basic elements of the topological semantic knowledge graph. Usually, they are expressed in the form of "head entity, relationship between entities, tail entity". $GeoMetaK$ represents the meta-knowledge of the topological semantic knowledge graph. It is usually used to indicate the updating of knowledge, P_T and P_L are used to represent the temporal and spatial relationships of topological semantic knowledge graphs, such as the deposition time order of the interpretation data and the topological spatial location association.

Introducing a knowledge-reasoning algorithm into a knowledge graph to constrain its construction and obtain accurate data samples is necessary. In constructing a topological structure knowledge graph, it is very important to determine the relationship between structural elements. Burns et al. introduced a technique for representing the geological topological relations using a network diagram, in which node entities denote spatial elements and edges indicate the topological connections between them³¹. Based on the above ideas, we used the hierarchical network to obtain the computer representation of the topological structure knowledge graph.

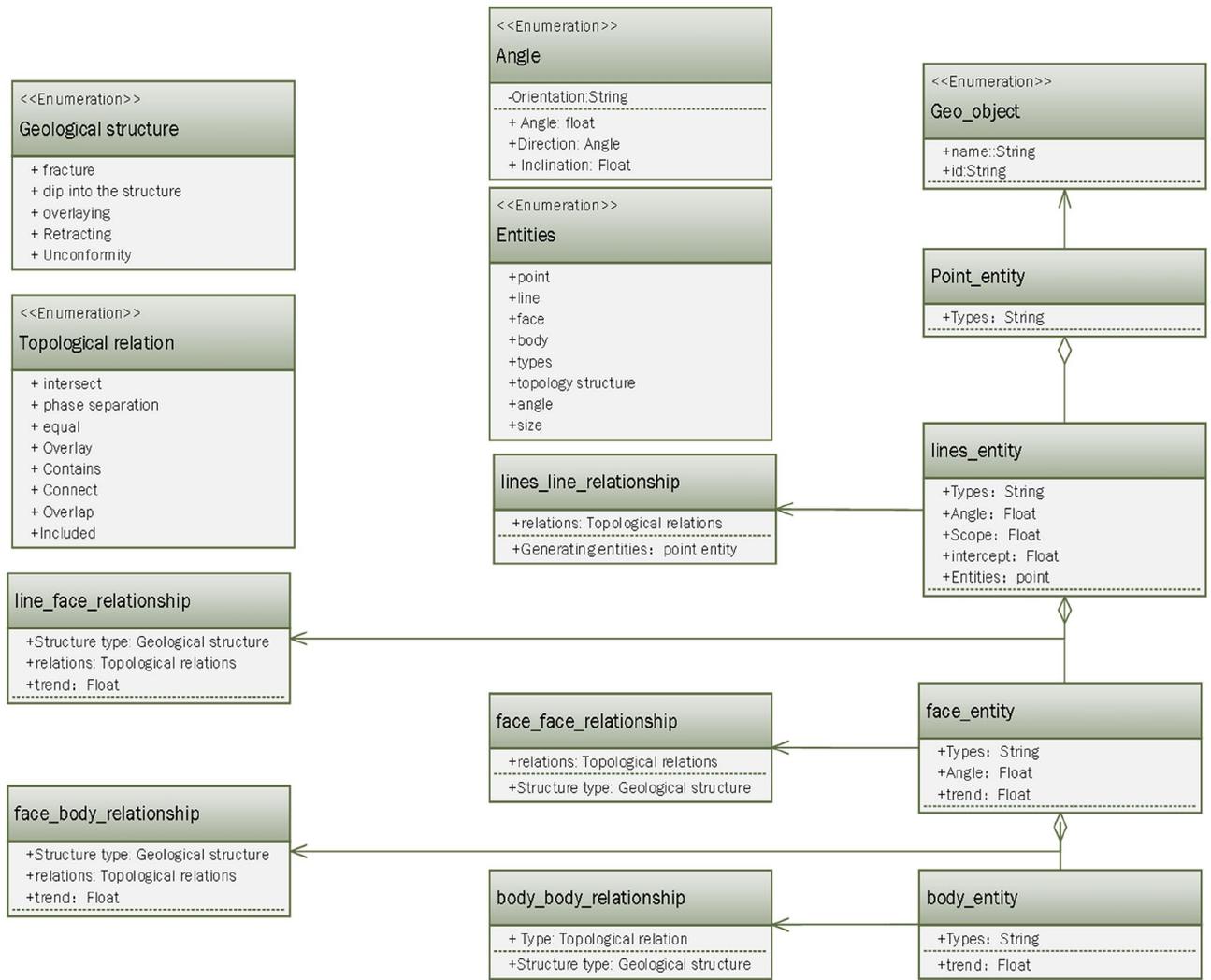


Figure 2. Knowledge constraint construction process.

The topological structure information fully represented the geometric and structural features of the structural geological model.

Complex geological structure knowledge reasoning Knowledge reasoning rule base construction

The pattern layer of a knowledge graph consists of defining entities, relationships, and attributes, as well as building the knowledge graph's rule base. Structural geology has a long history. Numerous rules and a lot of information regarding the genesis and regularity of movement of geological entities have been formed. Some of these rules can be summarized to form knowledge.

Therefore, according to the definition of a knowledge graph in the construction of a complex-structure knowledge graph, and the schema layer and rule base summary of knowledge constraints are shown in Figs. 3 and 4.

The basic form of the production rule base in this article is

$$R_k := \bigvee_{j=1}^n \left(\bigwedge_{j=1}^m E_{ijk} \right) \rightarrow C_k \quad m \geq 1, n \geq 1. \tag{2}$$

Among them, R it represents the number of rules, k is the k rule, which represents the obtained rule conclusion, and the extraction of the knowledge graph data layer is realized through logical rules. Where E and C represent the conditions and conclusions of rule reasoning. The pattern layer of a knowledge graph consists of defining entities, relationships, and attributes, as well as building the knowledge graph's rule base. Structural geology has a long history. Numerous rules and a lot of information regarding the genesis and regularity of movement of geological entities have been formed. Some of these rules can be summarized to form knowledge. Therefore, according to the definition of a knowledge graph in the construction of a complex-structure knowledge graph, the rule base of knowledge reasoning is summarized as shown in Fig. 4.

In the initial rule base, we first added nine separate geological structure rule patterns as the initial state and then used the geological constraint rule base (prior knowledge) to match the common structural patterns. For

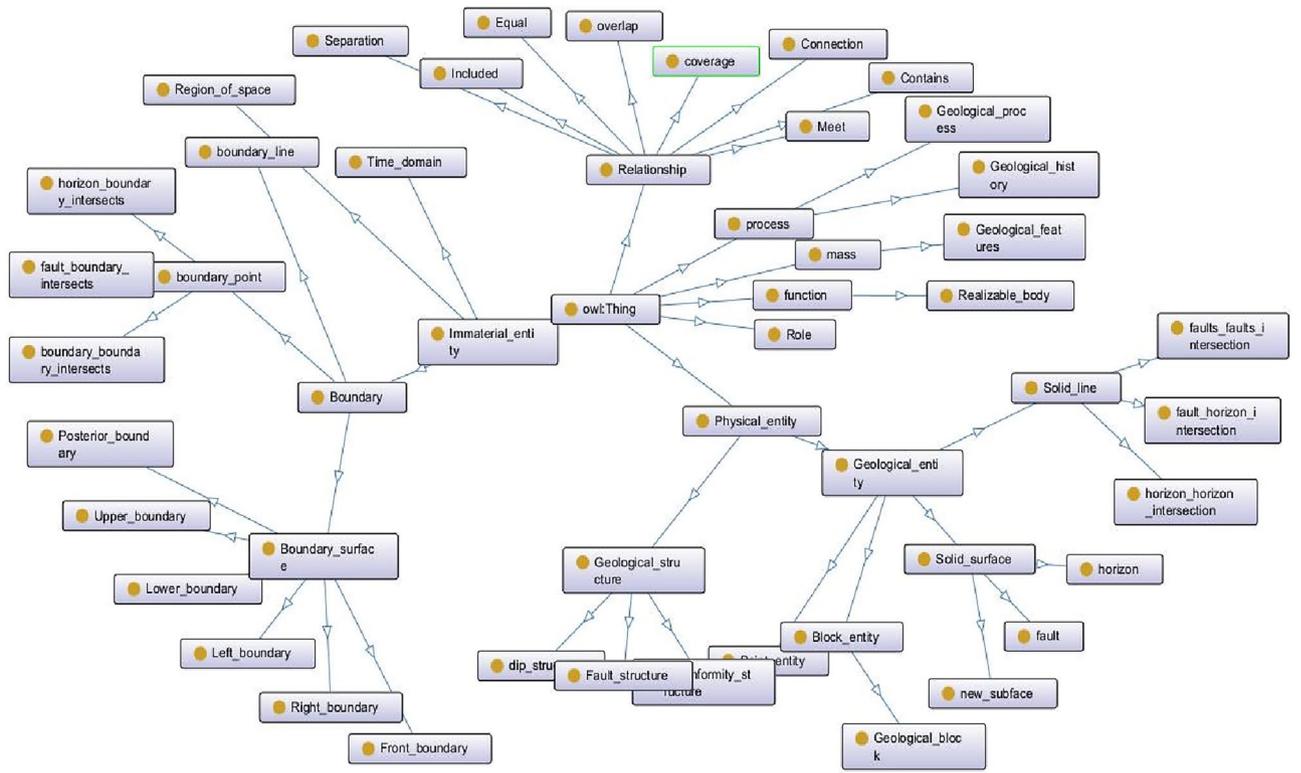


Figure 3. Structural geology knowledge ontology graph.

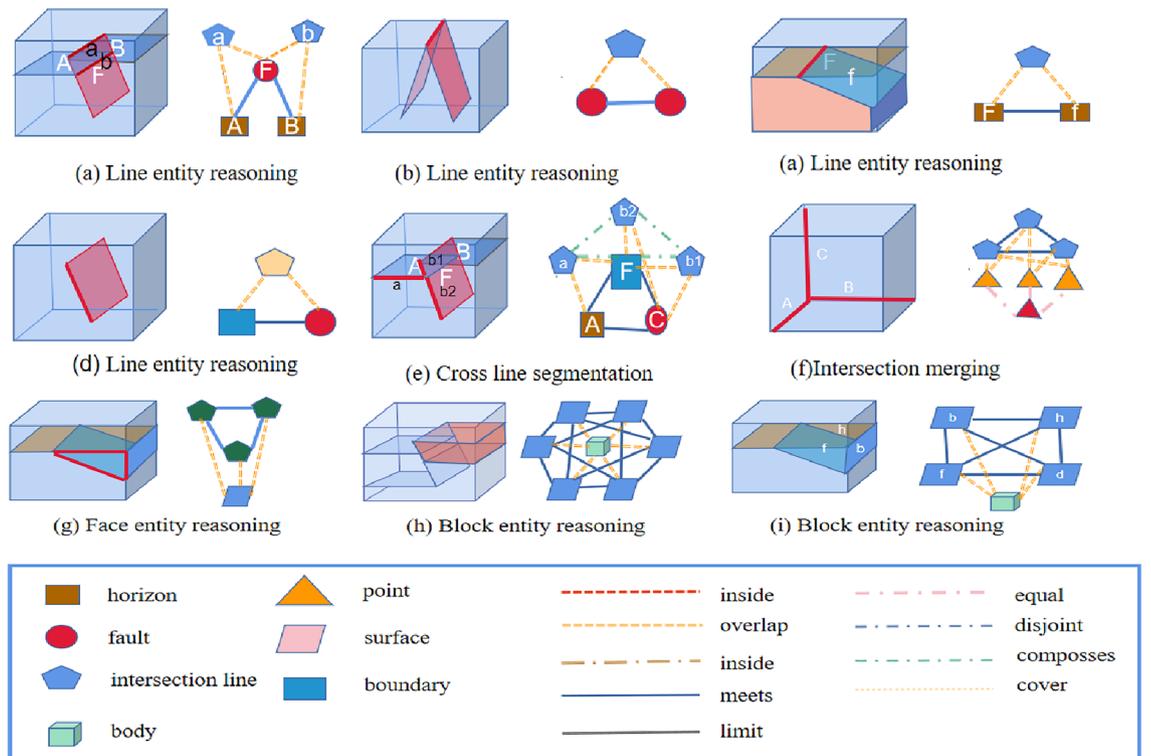


Figure 4. Topological rule base of knowledge reasoning, initial predefined geological structure rule base (geological structure diagram), including six common reasoning modes and subgraph reasoning query diagram; the geometric topology rule base (geometric structure diagram) contains common geometric reasoning patterns.

parts that match, we inferred the interpretation of the corresponding geological structure to be geologically reasonable. In the construction of model knowledge reasoning, the focus was on geometric topology knowledge graphs from the perspective of computer geometry and knowledge graph reasoning. In the new reasoning process, we focused on two possibilities: (1) different structural superposition and superposition methods since exhausting all geological structure rules is difficult, and (2) predefined rule errors (prior knowledge errors), as shown in Fig. 4. Experts must determine whether new rules need to be added or whether existing rules are suitable. In the process, the inference rule base is gradually improved and optimized.

The topological structure knowledge graph can be used to represent and describe various geological structural models, including fractures, intrusions, and unconformities. By representing these construction patterns as nodes and using edges to describe the relationships between them, a topology knowledge graph can be constructed. In the knowledge graph, the relationship between different nodes can be expressed as a topological relationship, which can more accurately represent the topological relationship in the geological structure and automatically perform topological inspection and constraints during the modeling process to ensure the correctness of the model.

The geological processes and structural types covered in the rule base are extensive, covering common types of geological structures such as faults, folds, and intrusions and geological structures of different scales, such as large-scale topography and small-scale rock structures. We described the universality of the rule base from three perspectives. (1) The first perspective involves the types of covering tectonic patterns; we listed the types of tectonic patterns contained in the rule base, such as faults, intrusions, and unconformity structures, as shown in Fig. 5. This rule base can cover the most common construction patterns in graphs. (2) The second perspective is the spatial distribution of cover tectonic models. In addition to the types of cover tectonic models, we considered the distribution of these tectonic models in geological space, which can describe various tectonic models in different geological periods and regions. (3) The third perspective includes the complexity of the cover structure model; the complexities of the geological structure models differ. Certain simple models may be relatively easy to describe and identify, whereas others may require more rules to describe. In the subsequent simulation, the given rules were used to achieve a more complex model (Fig. 5). Therefore, our rule base can cover and handle construction patterns with different complexities.

Knowledge reasoning for constraints of geological structure modeling

Knowledge reasoning, which can accurately connect to structural modeling, is the main technology used to build a knowledge graph in the field of structural modeling. The aim of constructing a knowledge graph is to obtain expert knowledge from seismic interpretation data and structured and semi-structured data from seismic interpretation data, such as coordinate data of the horizon and fault plane, intersection information of the horizon and fault plane, and the extraction of geological entities, spatial relations, semantic relations, and sedimentary relations. Furthermore, we determine whether the number of rules is finite or countable. Finally, a complete knowledge graph of the structural model was obtained by combining the knowledge reasoning of the spatial relations. In structural modeling, knowledge reasoning can combine spatial relations to achieve a more

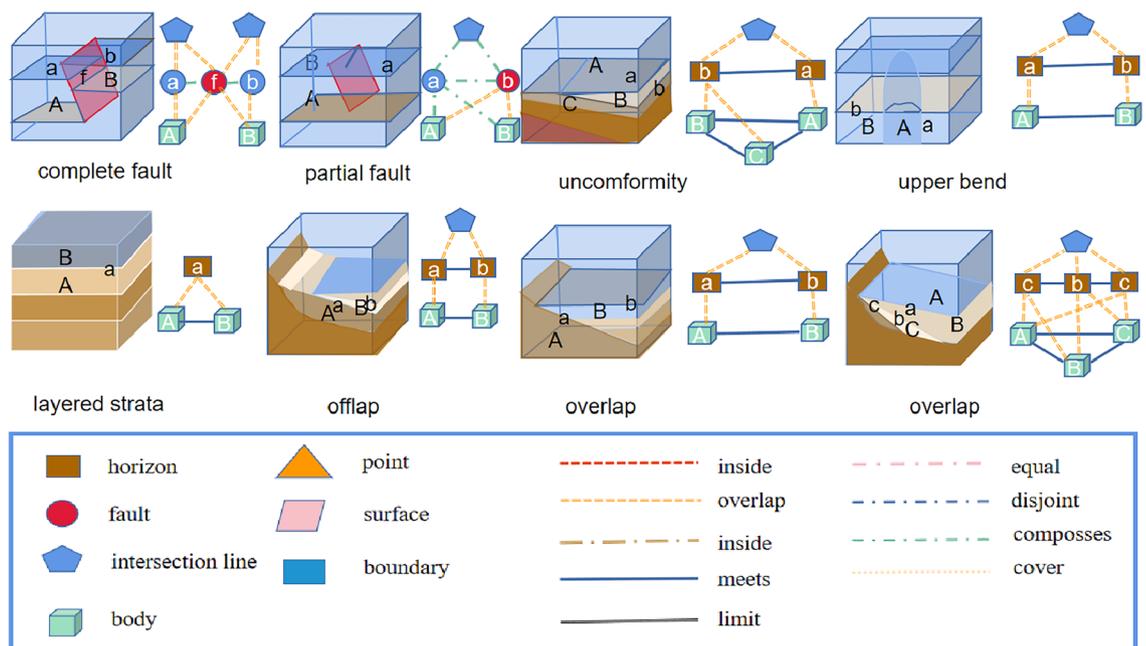


Figure 5. Lists common geological structure models, such as faults, intrusions, and unconformity structures. Most of the common geological structures exist in fault structures and unconformity structures. In the proposed rules, the above models can be expressed in the form of a knowledge graph.

accurate reasoning process. For example, the relationship between geological entities can be inferred based on their relative positions and intersections in space.

Reasoning constraints in the knowledge graph are realized through the rules of knowledge reasoning. The process of knowledge reasoning is as follows. The specific reasoning process is shown in Fig. 6. The aim of knowledge reasoning is to make human knowledge comprehensible to computers, construct a geometric topology knowledge graph, and guide 3D geological structure modeling. In our input, the effective horizon data is $H_i = \{h_{i1}, h_{im}, \dots, h_{ia}\}$, where h_{im} represents the horizon interpretation data at the m point of the i section, $M \leq a$, a is the number of horizon. And the effective fault data is $F_j = \{f_{j1}, f_{jn}, \dots, f_{jb}\}$ where the f_{jn} represents the fault interpretation data at the n th point of the j section, $1 \leq n \leq b$, b is the number of fault. the quantity of a and b is determined according to the specific work area.

To facilitate processing, we modeled the input prior knowledge as an adjacency matrix. According to the prior knowledge of geological structure (the intersection information of plane and fault plane) obtained from the perspective of geological experts, the prior information is represented by $G_1 = (E, V)$, where $E = \{GeoMetaK, E_{e_g}\}$,

$$GeoMetaK = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1(m+n)} \\ e_{21} & e_{22} & \dots & e_{2(m+n)} \\ \vdots & \vdots & \ddots & \vdots \\ e_{(m+n)1} & e_{(m+n)2} & \dots & e_{(m+n)(m+n)} \end{bmatrix}, \quad (3)$$

where $e_{ij} = 1$ indicates that the i -th face and the j th face have an intersection relationship, $GeoMetaK = \{E_F, E_H, V\}$ E_F and E_H represent the fault plane and horizon plane entity respectively.

And $(m+n)(m+n)$ represents the number of intersection relations between horizon planes and fault planes calculated by the bounding box method as our meta-knowledge. $e_{ij} = 0$ indicates that there is no intersection relationship. And the E_{e_g} represents the point, line, surface, block semantic entity that interprets the data generated, and its specific form is $E_{e_g} = \{e_{p1}, e_{p2}, \dots, e_{pn_p}, e_{l1}, e_{l2}, \dots, e_{ln_l}, e_{s1}, e_{s2}, \dots, e_{sn_s}, e_{b1}, e_{b2}, \dots, e_{bn_b}\}$ where n_p, n_l, n_s and n_b represent the number of the topological semantic entities of the generated point, line, surface and block entities respectively. What's more $V = \{V_{meet}, V_{overlap}, V_{inside}, V_{disjoin}, V_{cover}, V_{compose}, V_{equal}, V_{lim\ of}\}$. The following tables of the above edges represent the topological position relationships between entities respectively, represents the intersection information of prior knowledge (1 represents intersection, and 0 represents non-intersection) to construct the adjacency matrix of prior knowledge. Through the inference of rule base, the transformation of geological structure information to the topological relationship is realized, and the language (i.e., the form of adjacency matrix) that can be read by a computer is combined to create conditions for the subsequent sub-graph inference. The pseudo-code of knowledge reasoning related to algorithm 1 based on the rule base is as Table 1.

In the first stage of reasoning, prior human knowledge is combined with data that the computer can recognize, and computer cognition is realized by matching the corresponding rule base.

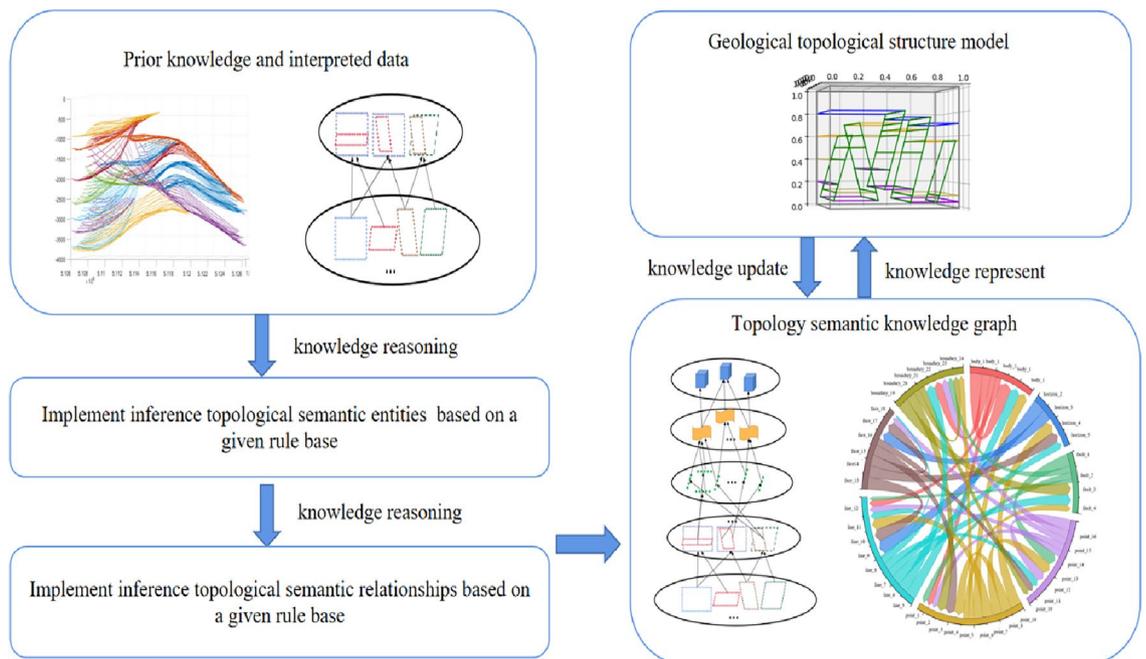


Figure 6. Complete process of knowledge reasoning can be divided into three parts: transforming geological data into geological constraints that can be read by computer; based on the constructed knowledge base, the knowledge graph is constructed via the knowledge reasoning algorithm. The knowledge graph is represented as a wireframe model to allow expert knowledge to participate in the knowledge graph update.

Algorithm 1: rule matching

Input: for each E_F, E_H and $V \in \{E_F, E_H\}$ interpretation data and calculate according to spatial bounding box to obtain prior knowledge.

for i in $1, \dots, E_H$ do

 for j in $1, \dots, E_F$ do

 if $v[i][j] = 1$ then

$e_{\{new_edge\}} = 1, new_edge += 1;$

$e_{\{new_edge\}} = 1, new_point += 2;$

 end if

 end for

end for

for i in $E_{e_g} = \{e_{i_1}, e_{i_2}, \dots, e_{i_{n-1}}\}$

 if $V[i] = 1$ then

$e_{\{new_surface\}} += 1, new_surface += 1;$

 end if

end for

Algorithm end

Output: knowledge graph that computers can understand $G_1 = \{E_H, E_F, E_{e_g}, V\}$

Table 1. The algorithm of rule matching.

After obtaining the basic topological entities and relationship information of the knowledge graph, pattern matching in the knowledge base is carried out through sub-graph matching research. The planar entities and

Algorithm 2 : face entity and block entity semantic mining algorithm

Input: Topology knowledge graph $G_d = \{E_{data}, V_{data}, L_{data}\}$ and query graph $G_q = \{E_{query}, V_{query}, L_{query}\}$

for i in $1, \dots, E_{data}$ do

 If $M(s)$ covers all the node of E_{query} then output $M(s)$;

 else compute the set $P(s)$ of the pairs candidate for inclusion in $M(s)$;

 initial value for node e

 for each $(e_{query}, e_{data}) \in P(s)$

 for each (e_{query}, e_{data}) in E_{query}

 if $F(s, e_{query}, e_{data})$ then

 compute the state $s1$ obtained by adding (e_{query}, e_{data}) to $M(s)$;

 call match ($s1$);

 end if

 end for each

 restore data structures;

end for

end procedure

Output: mapping between two graphs

Table 2. The algorithm of mining algorithm.

block entities existing in the knowledge graph are mined by graph isomorphism matching, and the entities and relationship information conforming to the geological structure are obtained by approximate subgraph matching. We studied the problem of sub-graph matching for knowledge graph. Specifically, given a query graph $G_q = \{E_{query}, V_{query}\}$ and data graph $G_d = \{E_{data}, V_{data}\}$ the sub-graph matching problem refers to obtaining all the data sub-graphs in G_2 that is isomorphic to G_3 to determine the new entity information in the knowledge graph. Through the study of each algorithm, we selected the VF2 algorithm with block efficiency and speed. The corresponding pseudo-code is as Table 2.

Generally, mapping M is expressed as the mapping of the node pair (e_{query}, e_{data}) ($e_{query} \in G_q$ and $e_{data} \in G_d$) each of which represents the mapping of node e_{query} of G_q and e_{data} of G_d , $M \in G_d \times G_q$ the function $F(s, e_{query}, e_{data})$ is a feasibility function that simultaneously increases the comparison of node and edge labels.

The return value is a Boolean value used to prune the search tree. At the same time, it can eliminate the situation where the two graphs can be isomorphic but the final matching result cannot be obtained, which is used to reduce the number of state spaces. $P(s)$ represents the set of all pairs of nodes to be matched. The VF2 algorithm is used to match the isomorphism graph, and a knowledge graph (face and block entities) that matches the given rule pattern is obtained. The VF2 algorithm cannot directly query the number of isomorphic sub-graphs; however, this function can be realized by modifying the VF2 algorithm. Specifically, based on the VF2 algorithm, each matched node pair can be marked, and unmatched node pairs can then be searched. Each time a new match is found, the marked node pairs are removed from the current search to query the number of isomorphic sub-graphs.

The specific process uses Algorithm 1 to realize rule matching and converts prior human knowledge into structured data on the computer with the help of prior rules to guide subsequent subgraph isomorphism matching. Through the modified sub-graph matching model, all sub-graphs that are isomorphic to the graph are traversed through sub-graph isomorphism to infer all surface entities and block entity knowledge and construct a complete knowledge graph. The quality of a knowledge graph obtained through knowledge reasoning is typically not guaranteed; therefore, it is added to the knowledge base. Before the process of quality evaluation is required, the wire-frame model is used in construction modeling to evaluate the quality of the knowledge graph. As shown in Fig. 7, when the constructed knowledge graph does not conform to expert cognition, the error information in the knowledge graph is modified and queried using subgraph matching to identify incorrect subgraph information in the knowledge graph. Quality assessment is the process of measuring and evaluating the credibility of new knowledge before it is added to a knowledge base to eliminate errors or conflicts.

Result

Figure 7 shows the workflow of knowledge reasoning construction modeling constraints participating in construction modeling. Firstly, the knowledge graph is represented as a wire-frame model to determine whether there are errors in the knowledge graph and realize the editing of the knowledge graph data. Secondly, the knowledge graph obtained by knowledge reasoning is a data format that can be recognized by the computer to constrain the boundary conditions of the construction model. Finally, the wire-frame model and the final construction model are consistent with expert knowledge.

In the construction process of knowledge reasoning in the knowledge graph, the entities and relationships in Fig. 8b,c are obtained through the first stage of knowledge reasoning using prior knowledge (Fig. 8a) (where prior knowledge is used to judge the intersection relationship between the horizon surface and fault surface by

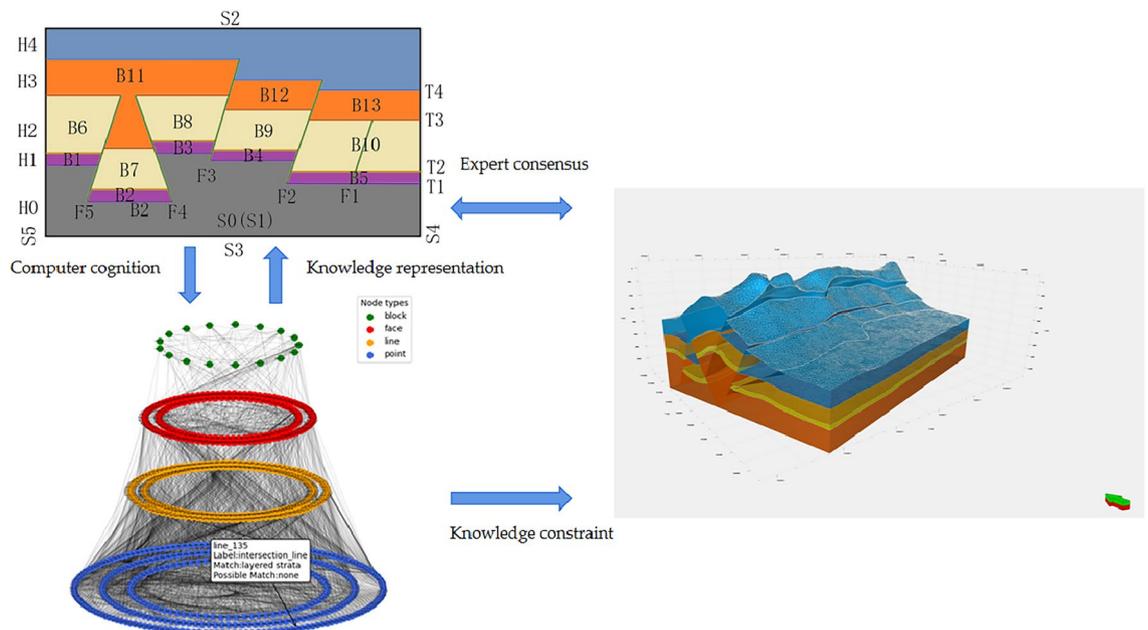


Figure 7. The whole knowledge reasoning ultimately participates in the workflow of the construction modeling.

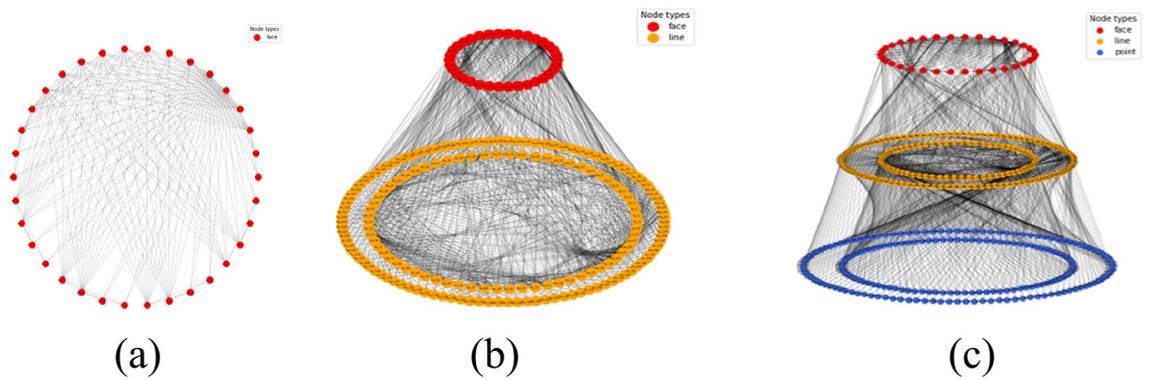


Figure 8. The process of point entity and line entity in the topological geometric knowledge graph obtained by rule base matching in the first stage of knowledge reasoning, which is divided into two steps. (a) Is prior knowledge, and geological experts input the intersection relationship between horizon surface and fault surface; (b) the topological position relationship between the line entity and the entity is obtained by matching the prior knowledge through the rule base; and (c) is the point entity obtained by the rule base matching for the intersection of line entities, and the entities with the same attributes are merged through knowledge fusion to complete the first stage of knowledge reasoning task.

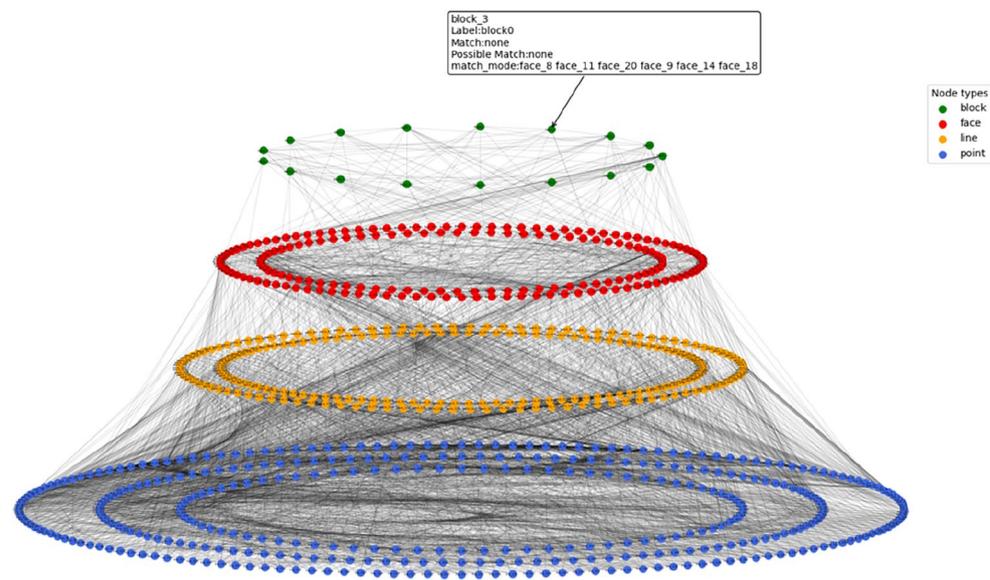


Figure 9. Second stage of knowledge reasoning, and the face entity and block entity obtained in the previous stage are used to obtain the subgraph isomorphic to the given graph model by subgraph isomorphism matching, and the matching entity and topological position relationship are recorded, to obtain the updated pure topological space knowledge graph.

interpreting the data). Here, entities refer to the intersection line and intersection point entities obtained from prior knowledge, as well as the topological semantic relationships between entities. Rule matching in the library was used to realize the reasoning of geological knowledge data to topological structure data, and a preliminary knowledge graph was constructed. In the second stage of knowledge reasoning, the hidden entity and relationship information in the knowledge graph are continuously enriched, and all sub-graphs that are isomorphic to the query sub-graph are obtained through the sub-graph isomorphism query. As shown in Fig. 9, the hidden surface entity and geological block entity knowledge (third and fourth layers) were obtained by reasoning. The knowledge graph obtained in this study does not contain prior knowledge of the initial definition. It is a purely topological semantic knowledge graph. The intersection point, intersection line, closed surface composed of an intersection line, and closed block entity composed of a surface were mined from prior interpretation data. The knowledge graph obtained by knowledge inference is used as the modeling constraint to estimate the boundary conditions of the constructed model[5] and get the final structural modeling, as show in Fig. 10.

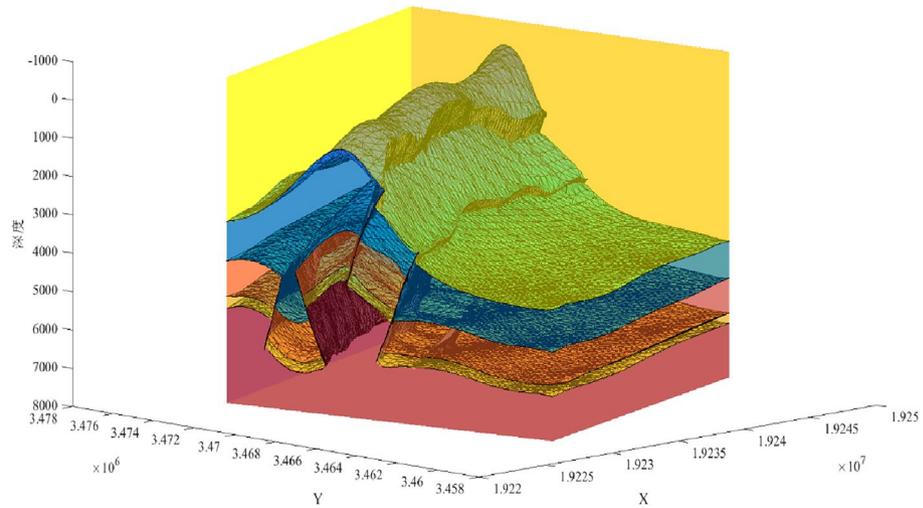


Figure 10. Closed geological body in the actual work area of three-dimensional geological structure under the constraint of knowledge graph.

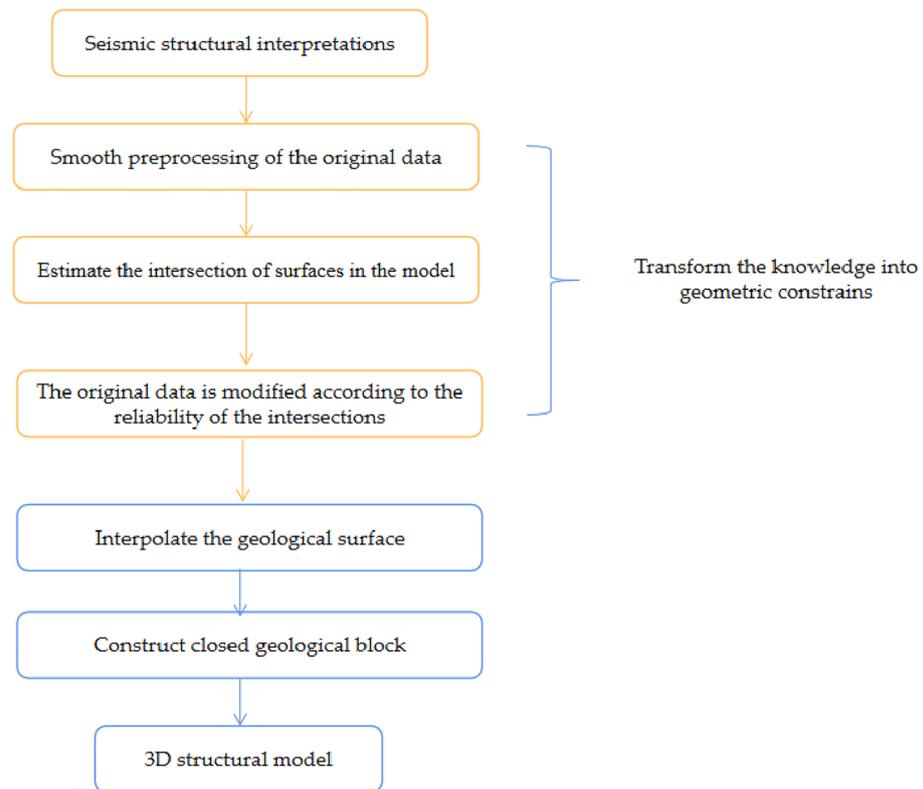


Figure 11. Complete process of knowledge reasoning constrained 3D construction modeling. It can be divided into two parts: knowledge reasoning participates in intersection estimation and three-dimensional geological body surface interpolation reconstruction closed three-dimensional geological body.

Discussion

Practical application of knowledge reasoning

The constructed geometric constraint knowledge graph can be used to constrain 3D geological structure modeling⁵. The specific construction process is shown in Fig. 11. Knowledge graph technology is used to construct knowledge graphs in the field of geology so that geologists can query and analyze geological data and identify geological laws and evolution trends. According to the geological knowledge in the knowledge graph, the model was constrained and optimized. The constraint relationship in the knowledge graph is used to specify

the topological relationship and geometric properties of the intersection line to improve the accuracy and controllability of the intersection line, provide the geometric constraints required for surface reconstruction, and ensure that the surface is as smooth as possible, while satisfying the geological structural characteristics. The optimized geological model was more consistent with the actual geological conditions.

In the subsequent modeling process, the topological structure knowledge graph is integrated into the boundary feature extraction, and then the boundary features are used as surface topology semantic constraints, and the morphological features are used as surface geometry semantic constraints, combined with multi-source information abstraction such as stratigraphic interpretation point clouds and well layered data. It is a multi-constraint surface regression model. By building a rectangular grid, based on the spatial autoregressive neural network, the fitting error and smoothing error are used as loss functions to solve the model, and then construct a model with high fitting degree, good smoothness and accurate morphological characteristics.

$$\begin{aligned} \arg \min_{w_i^{xTL}, w_i^{zTL}, w_i^{xBL}, w_i^{zBL}} LF = & \sum_{i=1}^m \frac{1}{2(\sigma_i^{hsTL})^2} \left(\mu_i^{hsTL} (x_i^{hsTL} - \varphi(y_i)w_i^{xTL} - b_i^{xTL}) - (z_i^{hsTL} - \varphi(y_i)w_i^{zTL} - b_i^{zTL}) \right)^2 \\ & + \sum_{i=1}^m \frac{1}{2(\sigma_i^{fsTL})^2} \left(\mu_i^{fsTL} (z_i^{fsTL} - \varphi(y_i)w_i^{zTL} - b_i^{zTL}) - (x_i^{fsTL} - \varphi(y_i)w_i^{xTL} - b_i^{xTL}) \right)^2 \\ & + \sum_{i=1}^m \frac{1}{2(\sigma_i^{hsBL})^2} \left(\mu_i^{hsBL} (x_i^{hsBL} - \varphi(y_i)w_i^{xBL} - b_i^{xBL}) - (z_i^{hsBL} - \varphi(y_i)w_i^{zBL} - b_i^{zBL}) \right)^2 \\ & + \sum_{i=1}^m \frac{1}{2(\sigma_i^{fsBL})^2} \left(\mu_i^{fsBL} (z_i^{fsBL} - \varphi(y_i)w_i^{zBL} - b_i^{zBL}) - (x_i^{fsBL} - \varphi(y_i)w_i^{xBL} - b_i^{xBL}) \right)^2 \\ \text{s.t.} \quad & \begin{cases} \varphi(y_i)w_{i+1}^{zTL} + b_{i+1}^{zTL} \geq \varphi(y_i)w_i^{zTL} + b_i^{zTL}, i = 1, \dots, k-1 \\ \varphi(y_i)w_{i+1}^{zBL} + b_{i+1}^{zBL} \geq \varphi(y_i)w_i^{zBL} + b_i^{zBL}, i = 1, \dots, k-1 \\ \varphi(y_i)w_i^{zTL} + b_i^{zTL} \geq \varphi(y_i)w_i^{zBL} + b_i^{zBL}, i = 1, \dots, k \\ \varphi(y_1)w_1^{zBL} + b_1^{zBL} \geq z_{\min}^{fs_1} \\ \varphi(y_m)w_m^{zTL} + b_m^{zTL} \leq z_{\max}^{fs_j} \end{cases} \end{aligned}$$

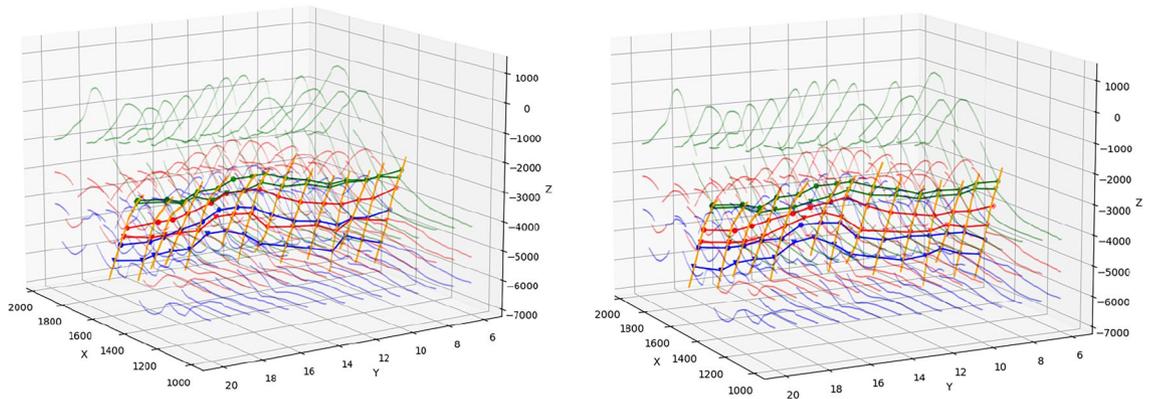


Figure 12. Comparison of boundary feature extraction methods. (a) Boundary characteristics guided by knowledge graph; (b) boundary characteristics of extrapolation interpolation.

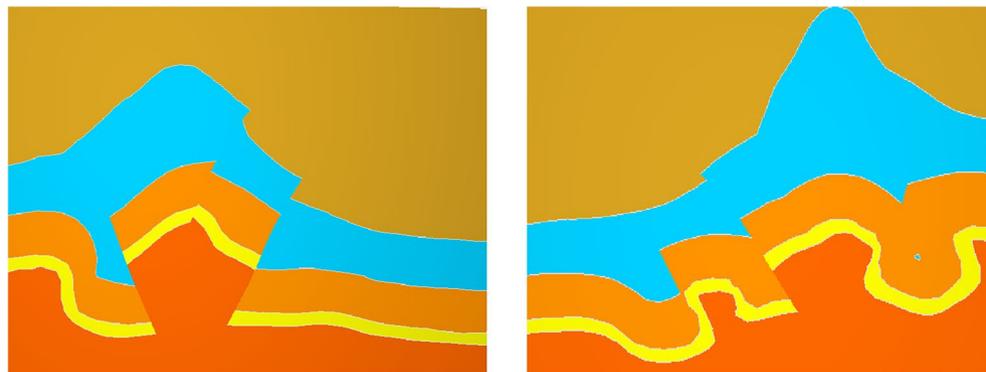


Figure 13. Three-dimensional geological structure (error).

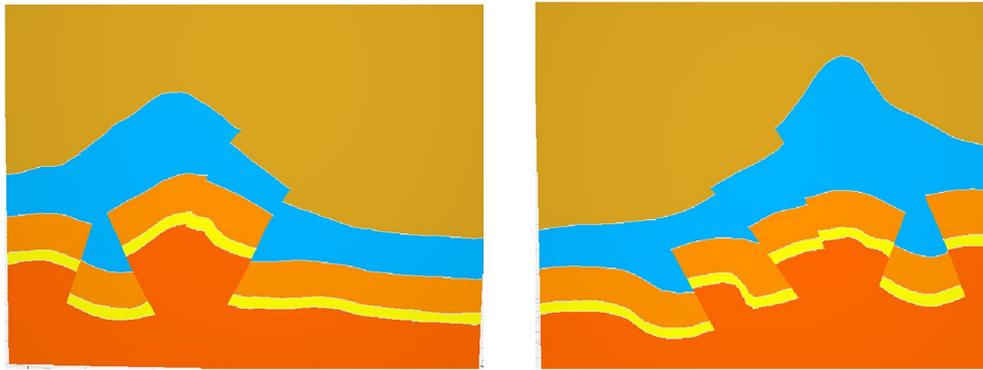


Figure 14. Three-dimensional geological structure.

where $(x_i^{hsTL}, y_i, z_i^{hsTL})$ and $(x_i^{hsBL}, y_i, z_i^{hsBL})$ represents the coordinate set of the end points of the middle level line of the i level plane in the upper and lower wall of the fault respectively. And $(x_i^{fsTL}, y_i, z_i^{fsTL})$ represents the set of coordinates corresponding to $(x_i^{hsTL}, y_i, z_i^{hsTL})$ in the fault line of the fault plane. In the same way, the $(x_i^{fsBL}, y_i, z_i^{fsBL})$ represents the set of coordinates corresponding to $(x_i^{hsBL}, y_i, z_i^{hsBL})$ in the fault line of the fault plane. μ_i^{hsTL} and σ_i^{hsTL} represent the mean and standard deviation of the dip angle near the end point $(x_i^{hsTL}, y_i, z_i^{hsTL})$ of the hanging wall horizon line. μ_i^{hsBL} and σ_i^{hsBL} represent the mean and standard deviation of the dip angle near the end point $(x_i^{hsBL}, y_i, z_i^{hsBL})$ of the hanging wall horizon line, μ_i^{fsTL} and σ_i^{fsTL} and μ_i^{fsBL} and σ_i^{fsBL} represent the mean and standard deviation of the reciprocal fault dip angle near fault points $(x_i^{fsTL}, y_i, z_i^{fsTL})$ and $(x_i^{fsBL}, y_i, z_i^{fsBL})$.

By comparison in Fig. 12, it is found that in the traditional extrapolation interpolation method, the intersection lines of the second horizon and the first and third horizons are staggered. At the same time, considering that the fault is a reverse fault, the hanging-footwall intersection line is required to be located above the hanging-footwall intersection line. In the traditional extrapolation interpolation method, the intersecting lines of the upper and lower disks of the first horizon also appear interleaved. Under the guidance of the principle of knowledge graph, the method in this paper can expertly extract reliable boundary feature lines that are consistent with geological laws and expert cognition. Ensure that the intersection line of the deposited lower horizon always remains below the intersection line of the deposited upper horizon. Moreover, the hanging-wall (thrust fault) or hanging-wall (normal fault) intersection line always remains above the hanging-wall (thrust fault) or hanging-wall (normal fault) intersection line. This method effectively prevents the occurrence of intersecting lines of adjacent layers or intersecting lines of upper and lower layers, and eliminates any unreasonable phenomenon.

To test the effectiveness of the complex geological knowledge graph based on knowledge reasoning, we constructed a three-dimensional geological model of the study area in Sichuan. When the traditional structural modeling method (Fig. 13) encounters an unreasonable fault or horizon, it obtains an accurate geological structural model by modifying the original data (Fig. 14). This method requires modification of a large amount of raw data, which is obviously not suitable for more complex work areas. But the method based on the knowledge reasoning (Fig. 10) can accurately constrain the three-dimensional geological modeling and reduce the uncertainty of structural modeling, what's more, this method modifies the knowledge to model the constraints and avoids the uncertainty caused by modifying the original data.

Other knowledge reasoning approaches

In addition, certain neural network-based methods, such as the graph neural network (GNN)^{32,33}, can be used for subgraph isomorphism reasoning. These methods use the learning ability of neural networks to query whether a corresponding subgraph exists by learning the feature vectors of nodes and edges and describing the reasoning problem of the graph in detail.

Owing to the high requirements of domain knowledge graphs, ensuring the accuracy of results using a deep learning method is difficult if applied to practical engineering. The existing method constructs a model rule base by human definition to realize the construction of a knowledge graph; however, following the development of deep learning, the rules existing in the construction of knowledge graphs can be automatically learned, efficient knowledge reasoning can be realized, and the rule base can be avoided. For example, the hidden subspace and closed block entities in a knowledge graph can be obtained by reasoning based on a random walk, and a topological semantic knowledge graph can be obtained. The AMIE algorithm proposed by Galárraga et al. is an association rule mining algorithm based on an incomplete knowledge base³⁴. Each rule is predicted by learning: for each relationship, starting from the rule whose body is empty, the body part of the rule is expanded by three operations, and the candidate rules whose knowledge degree exceeds the threshold are retained to realize association rule mining research of knowledge graphs.

Limitations

The method in this article aims to mine the topological structure of existing models from interpretation data. There is no better method for constructing topological semantic knowledge graphs for other data (such as seismic data); the limitation of this method is that it relies on Expert experience realizes the improvement of the knowledge reasoning rule base. For the limited work area data, our geological rule base can only build a knowledge graph and constrain the three-dimensional structural modeling for the existing work area. For other work areas, our method is theoretically applicable. However, for areas that are not accurate enough, we need to continuously add rules for reasoning to further improve our rule base. Based on a large number of experiments, our method is theoretically effective. Scalable. The disadvantage is that it relies on the participation of a large number of expert experience in the early stage, but in the actual work area, the accuracy of this method is relatively high and can meet the needs of actual production.

Conclusions

This study introduces a technique for knowledge reasoning in the field of structural modeling, which can be applied to create modeling constraints for 3D geological structure modeling in the context of oil exploration. The knowledge graph we constructed enables experts and users to access the semantic information contained in the model at any time and maintain expert knowledge throughout the modeling process. Geological experts can visualize the knowledge graph more easily and find connections between knowledge items. The knowledge graph is used to constrain the boundary information of complex structural models, improving the efficiency of intelligent structural modeling.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.

Received: 18 December 2023; Accepted: 20 February 2024

Published online: 27 February 2024

References

- Perrin, M. & Rainaud, J.-F. Shared earth modeling: knowledge driven solutions for building and managing subsurface 3D geological models. *Ed. Technip* **283**, 76–112 (2013).
- Caumon, G., Pellerin, J. & Laurent, G. Current bottlenecks in geomodeling workflows and ways forward. *Can. Soc. Petrol. Geol.* **37**, 32–64 (2013).
- Bentler, P. M. & Chou, C.-P. Practical issues in structural modeling. *Sociol. Methods Res.* **16**, 78–117 (1987).
- Guo, J., Wang, X. & Wang, J. Three-dimensional geological modeling and spatial analysis from geotechnical borehole data using an implicit surface and marching tetrahedra algorithm. *Eng. Geol.* **284**, 106047 (2021).
- Zhang, B., Tong, Y. & Du, J. Three-dimensional structural modeling (3D SM) and joint geophysical characterization (JGC) of hydrocarbon reservoir. *Minerals* **12**, 363 (2022).
- Radwan, A. E. Three-dimensional gas property geological modeling and simulation. In *Sustainable Geoscience for Natural Gas Subsurface Systems* (ed. Radwan, A. E.) 29–49 (Elsevier, 2022).
- Puzyrev, V., Salles, T. & Surma, G. Geophysical model generation with generative adversarial networks. *Geosci. Lett.* **9**, 1–9 (2022).
- Yu, P., Dempsey, D. & Rinaldi, A. P. Association between injection and microseismicity in geothermal fields with multiple wells: Data-driven modeling of Rotokawa, New Zealand, and Húsmúli, Iceland. *J. Geophys. Res. Solid Earth* **128**, e2022JB025952 (2023).
- Peng, H., Dukalski, M. & Elison, P. Data-driven suppression of short-period multiples from laterally varying thin-layered overburden structures. *Geophysics* **88**, V59–V73 (2023).
- Madsen, R. B., Høyer, A.-S. & Andersen, L. T. Geology-driven modeling: A new probabilistic approach for incorporating uncertain geological interpretations in 3D geological modeling. *Eng. Geol.* **309**, 106833 (2022).
- Hasan, M. & Shang, Y. Geophysical evaluation of geological model uncertainty for infrastructure design and groundwater assessments. *Eng. Geol.* **299**, 106560 (2022).
- Calcagno, P., Chilès, J.-P. & Courrioux, G. Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Phys. Earth Planet. Inter.* **171**, 147–157 (2008).
- Arantes, A. & Ferreira, L. M. D. Development of delay mitigation measures in construction projects: A combined interpretative structural modeling and MICMAC analysis approach. *Prod. Plann. Control* **1**, 1–16 (2023).
- Zhan, X., Lu, C. & Hu, G. 3D structural modeling for seismic exploration based on knowledge graphs. *Geophysics* **87**, 81–100 (2022).
- Steiner, T., Verborgh, R. & Troncy, R. Adding realtime coverage to the google knowledge graph. In *11th International Semantic Web Conference (ISWC 2012)* 65–68 (Citeseer, 2012).
- Alberti, B. Archaeologies of ontology. *Annu. Rev. Anthropol.* **45**, 163–179 (2016).
- Lee, J. & Zlatanova, S. A 3D data model and topological analyses for emergency response in urban areas. *Geospat. Inf. Technol. Emerg. Response* **1**, 159–184 (2008).
- Gray, J. & Rumpe, B. Explicit versus implicit models: What are good languages for modeling? *Softw. Syst. Model.* **21**, 839–841 (2022).
- Guo, J., Wang, J. & Wu, L. Explicit-implicit-integrated 3-D geological modelling approach: A case study of the Xianyan Demolition Volcano (Fujian, China). *Tectonophysics* **795**, 228648 (2020).
- Bourne, S., Oates, S. J. & Elk, J. F. V. A seismological model for earthquakes induced by fluid extraction from a subsurface reservoir. *J. Geophys. Res. Solid Earth* **119**, 8991–9015 (2014).
- Manzocchi, T., Carter, J. N. & Skorstad, A. Sensitivity of the impact of geological uncertainty on production from faulted and unfaulted shallow-marine oil reservoirs: Objectives and methods. *Petrol. Geosci.* **14**, 3–15 (2008).
- Oliver, M. A. & Webster, R. Kriging: A method of interpolation for geographical information systems. *Int. J. Geogr. Inf. Sci.* **4**, 313–332 (1990).
- Ming, J., Pan, M. & Qu, H. GIS: A 3D geological multi-body modeling system from netty cross-sections with topology. *Comput. Geosci.* **36**, 756–767 (2010).
- Guo, J., Wu, L. & Zhou, W. Section-constrained local geological interface dynamic updating method based on the HRBF surface. *J. Struct. Geol.* **107**, 64–72 (2018).

25. González-García, J. & Jessell, M. A 3D geological model for the Ruiz-Tolima Volcanic Massif (Colombia): Assessment of geological uncertainty using a stochastic approach based on Bézier curve design. *Tectonophysics* **687**, 139–157 (2016).
26. Mallet, J. L. Discrete smooth interpolation in geometric modelling. *Comput. Aided Des.* **24**, 178–191 (1992).
27. Yan, B., Janowicz, K. & Mai, G. A spatially explicit reinforcement learning model for geographic knowledge graph summarization. *Trans. GIS* **23**, 620–640 (2019).
28. Hogan, A., Blomqvist, E. & Cochez, M. Knowledge graphs. *ACM Comput. Surv.* **54**, 1–37 (2021).
29. Hao, X., Ji, Z. & Li, X. Construction and application of a knowledge graph. *Remote Sens.* **13**, 2511 (2021).
30. Zhan, X., Lu, C. & Hu, G. Event sequence interpretation of structural geological models: A knowledge-based approach. *Earth Sci. Inform.* **14**, 99–118 (2021).
31. Burns, K. L. *Retrieval of Tectonic Process Models from Geologic Maps and Diagrams* (1981).
32. Krleža, D. & Fertalj, K. Graph matching using hierarchical fuzzy graph neural networks. *IEEE Trans. Fuzzy Syst.* **25**, 892–904 (2016).
33. Liu, X., Pan, H. & He, M. Neural subgraph isomorphism counting. In *Proc. 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining 1959–1969* (2020).
34. Galárraga, L. A., Teflioudi, C. & Hose, K. AMIE: Association rule mining under incomplete evidence in ontological knowledge bases. In *Proc. 22nd International Conference on World Wide Web* 413–422 (2013).

Acknowledgements

In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Author contributions

The following statements should be used “Conceptualization, X.X. and C.L.; methodology, X.X.; software, C.L.; validation, C.L., X.X. and B.Z.; formal analysis, X.X.; investigation, C.L.; resources, X.X.; data curation, X.X.; writing—original draft preparation, C.L.; writing—review and editing, C.L.; visualization, C.L.; supervision, C.L.; project administration, C.L.; funding acquisition, X.X. All authors have read and agreed to the published version of the manuscript”.

Funding

This research was funded by National Natural Science Foundation of China, Grant Number 41974147.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to X.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024