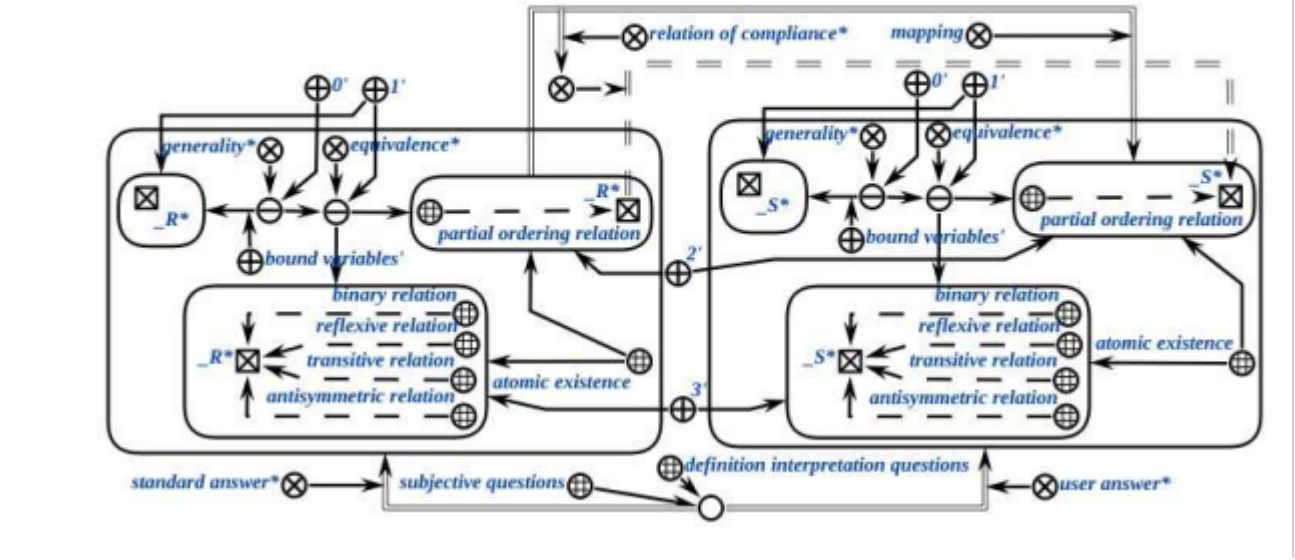

Figure 2. An example of establishing the mapping relationship between semantic graphs.

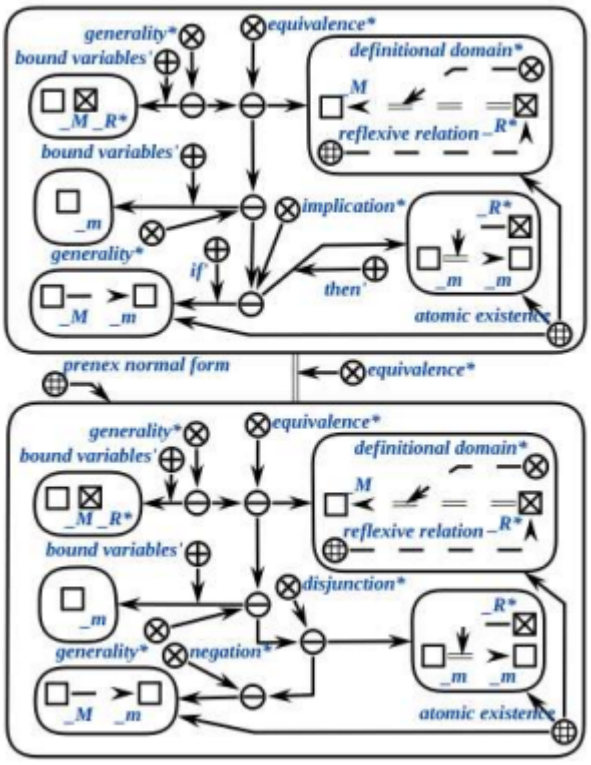

Figure 3. An example of converting a semantic graph into PNF representation.

code in strict accordance with the standard solution steps corresponding to the test question. The search template is used to search in the knowledge base for all semantic fragments corresponding to it [7], [28].

Since the user answers in natural language are converted into semantic graphs they are already integrated with the knowledge already available in the knowledge base. Therefore, when calculating the similarity between the semantic graphs, it is not necessary to consider the differences of the concepts at the natural language level. For example, Segment AB and Segment BA are represented by the same sc-node, they are just two identifiers of the sc-node [6], [7]. An approach to calculate the similarity between the semantic graphs of answers to proof questions and problem-solving task according to the reasoning tree of standard answer (semantic graph of standard answer) is proposed in this article, and the specific calculation process is shown below:

1) numbering each semantic sub-graph in the semantic graph of user answer (the numbering order started from 1);
2) each node in the reasoning tree (search template) is traversed in turn according to the DFS strategy. At the same time, the corresponding semantic sub-graph that is included in the semantic graph of the user answer are searched in the knowledge base using the search template currently being traversed. If such a semantic sub-graph exists, then determine whether the searched semantic sub-graph number is smaller than the semantic sub-graph number corresponding to the search template of the current search template parent node (except for the root node of the reasoning tree), and if so, the searched semantic sub-graph is considered correct;
3) repeat step 2) until all search templates in the reasoning tree have been traversed and the number of correct semantic sub-graphs is counted at the same time;
4) using formulas (1), (2) and (3) to calculate the precision, recall and similarity between answers.

Since this article focused on the entire process from test