Chinese/English Journal of Educational Measurement and Evaluation | 教育测量与评估双语季刊

Volume 2 | Issue 2 Article 1

2021

# The Augment Algorithm and its Role in Cognitive Diagnosis

Shuliang Ding
*Jiangxi Normal University*

Fen Luo
*Jiangxi Normal University*

Wenyi Wang
*Jiangxi Normal University*

Xiaofeng Yu
*Jiangxi Normal University*

Jianhua Xiong
*Jiangxi Normal University*

# The Augment Algorithm and its Role in Cognitive Diagnosis

Shuliang Ding [a], Fen Luo [a], Wenyi Wang [a], Xiaofeng Yu [a], and Jianhua Xiong [a]

[a] Jiangxi Normal University

**Abstract**
Both the augment algorithm and the reduction algorithm can be used to obtain non-zero knowledge states vector in testing. In particular, the augment algorithm based on the reachability matrix can imply the structure of the **Q**-matrix and its non-zero columns, thus proving the set of **Q**-matrix columns forms an algebraic structure (Lattice). Applying the augment algorithm based on the reachability matrix and the test **Q**-matrix, respectively, we can obtain the theoretical construct validity of the test **Q**-matrix (i.e., the degree of the test **Q**-matrix fitting the cognitive model) and use the results to evaluate test quality. We can also use the algorithm when constructing and evaluating cognitive models, as well as when developing cognitive diagnostic models. Moreover, the augment algorithm and its reverse algorithm (reduction algorithm) are suitable for analyzing and evaluating retrofitting data.

**Keywords**
Augment algorithm;
Reachability matrix;
Q-matrix;
Lattice;
Theoretical construct validity;
Retrofit

## 1 Problem Statement

To facilitate our following discussions, we first introduce the relevant terms. Let **A** denote the adjacency matrix, **E** denote the identity matrix of the same order as **A**, and **R** denotes the reachability matrix. In this paper, we regard the reachability matrix not as a fixed matrix, but as class of matrices that can become identical after column permutation. We also let **Q** denote either the Boolean matrix with element 0 (or 1) or the polytomous **Q**-matrix with non-negative integer elements. Obviously, the Boolean matrix is a particular case of the polytomous **Q**-matrix. Generally, the **Q**-matrix is considered as the incidence matrix of attributes and items (Tatsuoka, 1995, 2009). In fact, it is also the incidence matrix between attributes and the respondents (corresponding to their knowledge states). Each row of the **Q**-matrix corresponds to an attribute. Each column corresponds to the attribute vector of an item or the knowledge states of a respondent.

Given the hierarchical relationships among $K$ attributes, we can obtain the adjacency matrix **A** (and ultimately the reachability matrix **R**). Conversely, knowing the adjacency matrix, we can also obtain the attributes and their hierarchical relationships. Therefore, there is a one-to-one correspondence between attributes' hierarchical relationships and the adjacency matrix. Then, given the attributes and their hierarchical relationships represent the associated cognitive model, the adjacency matrix also has a one-to-one correspondence with the cognitive model. The adjacency matrix reflects the direct relationships or immediately prerequisite relationships between attributes (Tatsuoka, 1995, 2009), where as the reachability matrix **R** represents the direct, indirect, and reflexive relationships among the attributes. However, because the adjacency matrix only denotes the immediate prerequisite between the attributes, it cannot be a **Q**-matrix. In comparison, because the reachability matrix can denote the incidence relations between the attributes and the items, it can be regarded as a **Q**-matrix.

The Boolean matrix composed of all the different $K$-dimensional 0-1 column vectors is called the full matrix, denoted as $\mathbf{Q}_a$. After deleting columns that violate the limits imposed on **R** in the full matrix, we obtain the reduced **Q**-matrix ($\mathbf{Q}_r$; Tatsuoka, 1995, 2009). This method of obtaining $\mathbf{Q}_r$ from the full matrix by deleting some columns is called *the reduction method*. The test **Q**-matrix consists of the subset of the non-zero columns of $\mathbf{Q}_r$.

Tatsuoka (1995, 2009) pointed out that cognitive diagnostic classification is essentially statistical pattern recognition. There are two approaches (Duda et al.,

**CONTACT**: Shuliang Ding. ding06026@163.com. Jiangxi Normal University.

2003) to pattern recognition—supervised learning and unsupervised learning—that correspond to discriminant analysis (supervised learning) and clustering analysis (unsupervised learning), respectively. The purpose of cognitive diagnosis is to map the observed response patterns to some knowledge states. Therefore, the test **Q**-matrix is equivalent to the sensor in pattern recognition; the ideal response pattern is feature extraction; the knowledge state is the result of classification. If there are suitable sensors, feature extraction and classification results can be matched correspondingly. The diagnostic analysis becomes supervised learning if we are given the state of knowledge. The columns of the $\mathbf{Q}_r$ are the set of all possible knowledge states, and cognitive diagnosis is equivalent to supervised learning based on the set. In general, supervised learning is better than unsupervised learning in terms of interpretability. Because diagnosis aims to classify each observed response pattern as a certain state of knowledge by predefining classification rules, it is vital to find all the $\mathbf{Q}_r$ columns. The reduction method can generate all the knowledge state or item attribute vectors that satisfy the attribute hierarchy.

However, it is difficult to analyze the column-construction rules of $\mathbf{Q}_r$ via the reduction method, and we cannot infer what algebraic systems might be formed after we define some algebraic operations on the set of these columns. Additionally, it is difficult to conduct qualitative analyses on the retrofitting data obtained from any test that was not developed for diagnostic purposes. Examples of such a test include TOEFL, National Computer Rank Examination, and Public English Test System. We might ask the following questions. Does the attribute vector of the item in the retrofitting data conform to the attribute hierarchy? Can the reachability matrix be obtained from the retrofitting data? Even if the reachability matrix can be obtained, does the associated attribute hierarchical relationships match what the experts have specified? This means the incidence matrix corresponding to the reachability matrix must be explicit and can be used as a reference by field experts. Because the structure of the $\mathbf{Q}_r$'s columns cannot be obtained via the reduction method, it is challenging to apply algebraic methods and determine whether the **Q**-matrix obtained by domain experts based on the retrofitting data contains the reachability matrix, and whether it represents attribute hierarchical relationships.

It is possible to extract attributes and their hierarchical relationships through data mining techniques (Wang & Lu, 2021), but this requires response data with a large sample size. In general, proving the correctness of data mining results is also a challenging theoretical problem. Furthermore, it would be meaningful if we can determine whether the retrofitting data can be used for diagnostic analyses by only examining the test **Q**-matrix, without using examinees' test responses.

The augment algorithm can address the aforementioned issues (Ding et al., 2008, 2009; Yang et al., 2008) under certain conditions. To explicate this idea, the rest of this paper is organized as follows. First, we introduce the augment algorithm and its properties based on the reachability matrix or a general **Q**-matrix. Second, we present the role and application of the augment algorithm in cognitive diagnosis. Lastly, we provide a summary and discussion.

## 2 The Augment Algorithm Based on the Q-Matrix

The reduction algorithm introduced by Tatsuoka (1995, 2009) regards the reachability matrix as a "sieve," while the augment algorithm regards it as a "seed." What is the relationship between these two algorithms? Why is the augment algorithm still needed, given the reduction algorithm exists already? If the derivation of respondents' all possible knowledge states can be obtained using popular R packages (i.e., CDM, GDINA), what is the special role of the augment algorithm in this case? To answer these questions, we first discuss what kind of **Q**-matrix is augmented because the results may differ when different **Q**-matrices are used. In the following sections, if the augment algorithm is used based on a matrix **H**, then it is specifically called **H**-based augment, and the matrix **H** is referred to as the basis matrix.

### 2.1 The Basis Matrix is a Boolean Matrix

#### 2.1.1 The Augment Algorithm

Suppose **Q** is the basis matrix,which is a Boolean matrix of $K$ rows and $m$ columns. We then divide **Q** into multiple column vectors, starting from the column $j = 1$ (called the operational column), and perform Boolean union operations on the column right next to column $j$. Specifically, we perform element-wise Boolean union operations between the two vectors. If the resulting column is different from the existing columns (called the new column), add the new column to the far right of the **Q**-matrix. Next, we move the operational column one position to the right and repeat the above process until all the operational columns ($j = 1$, 2, ..., $K$) are completed. Based on the reachability matrix, the potential **Q**-matrix (Ding, Wang, & Luo, 2012) can be obtained by the augment algorithm (Ding et al., 2008, 2009; Yang et al., 2008).

---

**Algorithm** The Augment Algorithm

---

**Input:** The reachability matrix **R**. Let **R** be a matrix of size $K \times K$.
**Output:** The potential **Q**-matrix $\mathbf{Q}_p$.
  $\mathbf{Q} = \mathbf{R}$
  $\mathbf{Q} = (r_1, r_2, \ldots, r_K)$ # Divide **Q** into column vectors
  **for** $j = 1$ to $K$ **do**
    **for** $t = j + 1$ to$m$ **do**
      **if** $r_j \oplus r_t$ is not equal to any existing column **then**
        $m = m + 1$
        $r_m = r_j \oplus r_t$
          # $\oplus$ is the Boolean addition operation

---

### 2.1.2 Augmenting Based on the Reachability Matrix

**Proposition 1. The results of the reduction algorithm and the augment algorithm based on the reachability matrix are the same; that is, the same set of non-zero columns can be obtained based on these two algorithms.**

**Proof.** Obviously, the augment algorithm cannot generate zero columns. Firstly, all the columns derived from the augment algorithm, which is based on the reachability matrix, must satisfy the hierarchical relationships and hence will not be filtered out. Also, the column set ($S_2$) obtained by the augment algorithm is a subset of the set ($S_1$) derived from the reduction algorithm.

Conversely, we shall prove that $S_1$ is a subset of $S_2$—that is, if **x** belongs to $S_1$, then **x** also belongs to $S_2$.

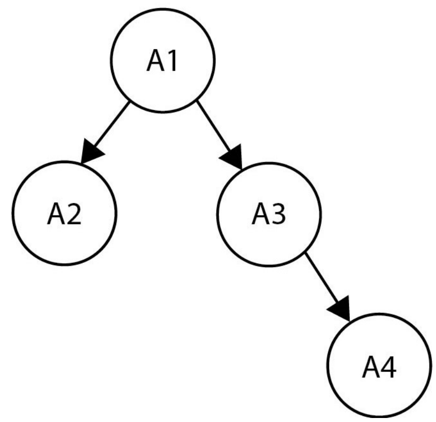
If **x** does not belong to $S_2$, then it cannot be expressed by the Boolean union of the columns of the reachability matrix, so **x** must contain elements that do not satisfy the requirements of the reachability matrix. In this case, **x** can be filtered out by the reduction algorithm, meaning **x** must not belong to $S_1$. Consequently, $S_1$ is a subset of $S_2$. In conclusion, $S_1 = S_2$. It can be seen that the potential **Q**-matrix is the same as all of the non-zero columns in the reduced **Q**-matrix. The proof is complete.

Then, let us consider a class of special **Q**-matrices.

**Definition 1. The Q-matrix is called a necessary Q-matrix if the reachability matrix is its sub-matrix (Ding et al., 2011). The columns of the reachability matrix are called the essential columns, while the columns augmented from the reachability matrix are called non-essential columns.**

**Example 1.** Based on the hierarchical attributes structure shown in Figure 1, we introduce the implementation processes of the augment algorithm and the reduction algorithm.

Figure 1
*An Attribute Hierarchy*




The reachability matrix based on the above figure is defined as follows:

| | A1 | A2 | A3 | A4 |
|---|---|---|---|---|
| A1 | 1 | 1 | 1 | 1 |
| A2 | 0 | 1 | 0 | 0 |
| A3 | 0 | 0 | 1 | 1 |
| A4 | 0 | 0 | 0 | 1 |

The augment process are defined as follows:

$$\left|\begin{matrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{matrix}\right| \xrightarrow{\text{First cycle}} \left|\begin{matrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{matrix}\right| \xrightarrow{\text{Second cycle}}$$

$$\left|\begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{matrix}\right| \quad \text{Third, and fourth cycle. No further augmentation.}$$

The corresponding steps of the reduction algorithm are:

1. First, write the full matrix $\mathbf{Q}_a$ of four attributes in binary representations (only non-zero columns are given);

$$\mathbf{Q}_a = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

2. Since the first element of each column of the reachability matrix **R** is 1, columns 1 to 7 in $\mathbf{Q}_a$ should be removed;

3. Columns 9 and 13 in $\mathbf{Q}_a$ do not conform to the logical order of **R** and should be removed;

4. Only columns 8, 10, 11, 12, 14, and 15 are retained in $\mathbf{Q}_a$.

It can be seen that the results produced by the reduction algorithm are the same as those produced by the augment algorithm when only non-zero columns are considered.

#### 2.1.3 Augmenting Based on the Test Q-Matrix

A test **Q**-matrix can be obtained from the retrofitting data. Let the basis matrix be the test **Q**-matrix.

**Proposition 2. If the test Q-matrix is a necessary matrix, then all the columns in the potential Q-matrix can be augmented based on the test Q-matrix.**

**Proof.** Because the potential **Q**-matrix can be obtained by the augment algorithm based on the reachability matrix, the other non-essential columns are contained in the potential **Q**-matrix. Furthermore, the matrix obtained by the augment algorithm based on the necessary **Q**-matrix must still be the potential **Q**-matrix.

**Proposition 3. If the test Q-matrix (denoted as H) is not a necessary Q-matrix, then the set of columns based on H-augment must be a proper subset of the potential Q-matrix.**

**Proof.** Because **H** is not a necessary **Q**-matrix, it lacks at least one column that is present in the reachability matrix (e.g., column **x**). If we can show that **x** cannot be augmented from the other columns in **H**, then we prove that the true sub-matrix of the potential **Q**-matrix can be obtained by augmenting based on **H**. This is because the reachability matrix **R** can always be represented as an upper triangular matrix, and **x** cannot be augmented from the columns excluding **x** in **R**. If the non-essential columns contain **x**, the number of non-zero elements must be more than the number of non-zero elements in **x** (i.e., the "length" is greater than **x**'s length). In this case, **x** cannot be augmented as well. Lastly, **x** cannot be augmented from the non-essential columns that do not contain **x**. The proof is complete.

The necessary **Q**-matrix (Ding et al., 2011; Ding, Luo, & Wang, 2012) is named to distinguish it from sufficient **Q**-matrix. It's also sometimes called complete **Q**-matrix (Cai et al., 2018). Clearly, the necessary **Q**-matrix must be sufficient, while the sufficient **Q**-matrix may not be a necessary **Q**-matrix. It is easy to find an example that a sufficient **Q**-matrix (Tatsuoka, 1995, 2009) is not a necessary **Q**-matrix (see Example 2), and the augment based on the sufficient **Q**-matrix may not be able to obtain all non-zero columns derived from the reachability matrix.

**Example 2.** Suppose that there are three attributes with independent structure, whose reachability matrix is a identity matrix of size three, corresponding to seven non-zero knowledge states. Let us consider a **Q**-matrix $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. This **Q**-matrix is sufficient because if we compare the rows, we see the three attributes are not prerequisites for each other, and this implies the reachability matrix. However, this matrix is not a necessary **Q**-matrix because it does not contain the reachability matrix. Based on this sufficient matrix, only four non-zero columns can be obtained, not seven non-zero columns. It can be seen from Proposition 3 that given the attributes and their hierarchical relationships, the result based on the sufficient **Q**-matrix (not the necessary **Q**-matrix) is different from that based on the augment of the necessary **Q**-matrix.

#### 2.1.4 The Incremental Augment Algorithm

The augment algorithm involves performing Boolean union operations between column $j$ and all columns to the right of column $j$ in the basic **Q**-matrix. In comparison, the incremental augment algorithm (Yang et al., 2010) involves performing Boolean union operations for all columns to the left of column j in the basic **Q**-matrix, and finding and inserting the new columns before column $j$. The specific steps of this algorithm are listed on the next page.

**Algorithm** The Incremental Augment Algorithm

```
Output: The potential Q-matrix Q_p
  R = (r_1, r_2, ..., r_K)          # Divide R into column vectors
  m = 1;
  for j = 1 to K do
      m = m + 1;
      q_m = r_j        # q_m is a new column vector added in Q
      for t = m - 1 to 1 do
          if r_j ⊕ q_t is not equal to any existing column then
              m = m + 1
              q_m = r_j ⊕ q_t
                        # ⊕ is the Boolean addition operation
```

### 2.2 The Basis Matrix is a Polytomous Q-Matrix

A **Q**-matrix whose elements are non-negative integers is called a polytomous **Q**-matrix (Chen & de la Torre, 2013; J. Sun et al., 2013). One advantage of the augment algorithm is that it can easily be extended from the Boolean matrix case to the polytomous **Q**-matrix case once we rewrite the Boolean union of *a* and *b* as *max*{a, b}. To obtain the polytomous potential **Q**-matrix (all non-zero polytomous knowledge states), the basis matrix should be the quasi-reachability matrix (Ding, Wang, Luo, & Xiong, 2015; Ding, Luo, Wang, & Xiong, 2016). The polytomous quasi-reachability matrix $\mathbf{R}_p$ is defined and calculated as follows.

According to J. Sun et al. (2013), we suppose there are $K$ attributes $A_1, A_2, \cdots, A_K$, where the highest level of attribute $A_i$ is an integer denoted by $w_i$, $w_i \geq 1$, $i = 1, 2, \ldots, K$. First, the binary reachability matrix $\mathbf{R}_2$ of shape $K \times K$ is given according to the attributes and their hierarchical relationships. Second, the elements of row $i$ and column $i$ of $\mathbf{R}_2$ (i.e., the diagonal elements) are expanded into a $w_i$-dimension row vector $(1, 2, \ldots, w_i)$, while the other elements in column $i$ of $\mathbf{R}_2$ are multiplied by a $1 \times w_i$ vector of ones (scalar multiplication), $i = 1, 2, \ldots, K$. In this case, the $\mathbf{R}_2$ matrix of shape $K \times K$ is expanded into a polytomous reachability matrix $\mathbf{R}_p$ of shape $K \times (\sum_{i=1}^{K} w_i)$. This $\mathbf{R}_p$ is a quasi-reachability matrix. The term "quasi" is added because the polytomous reachability matrix is not necessarily a square matrix, whereas the reachability matrices are all square matrices when using a Boolean matrix. An illustrative example is given in Appendix A.

### 2.3 The Inverse Algorithm of the Augment Algorithm: The Reduction Algorithm

Next, we consider whether the necessary **Q**-matrix can represent the cognitive model. Can the adjacency matrix be extracted from the necessary **Q**-matrix? Gierl et al. (2000, p. 40) expressed concerns about the possibility of obtaining attributes' direct and indirect relationships from the reachability matrices. This is an important issue to address because it is concerned with whether the reachability matrix can represent the cognitive model.

The reduction algorithm (Ding, Mao, et al., 2012) can reduce the potential **Q**-matrix $\mathbf{Q}_p$ to the reachability matrix **R**, and the reachability matrix can be obtained through the reduction algorithm based on any necessary **Q**-matrix. The principle of the reduction algorithm is deleting all the non-essential columns in the **Q**-matrix one by one, retaining only the columns that cannot be yielded by the Boolean union of all columns of the **Q**-matrix. Note that for some **Q**-matrices, if the number of columns retained by the reduction algorithm is small than the number of attributes, then the **Q**-matrix indeed cannot contain the reachability matrix. Even if the number of remaining columns is equal to the number of attributes, the **Q**-matrix does not necessarily contain a reachability matrix (e.g., the **Q**-matrix is a sufficient but not necessary **Q**-matrix, see Example 2). In addition to the reduction algorithm, the cleansing algorithm can be used to obtain the adjacency matrix from the reachability matrix. This algorithm first sets all the diagonal elements of the reachability matrix to 0 and then converts all elements that may be transferred to 0 (Ding & Luo, 2005, 2013). With the two aforementioned algorithms, both the necessary **Q**-matrix and the reachability matrix can represent the cognitive model.

For an introduction and examples of the reduction algorithm and the cleansing algorithm, please refer to Appendix B.

## 3 The Role and Application of the Augment Algorithm in Cognitive Diagnosis

### 3.1 Establishing the Important Role of Reachability Matrix in Cognitive Diagnosis

In the translators' preface of the Chinese Version of *Psychological Testing*, Zhu states: "Psychological test is an indirect measurement. ... To make accurate and reliable inferences, we need the behaviors of the respondents to be representative..., and the test items that cause the behaviors related to the psychological attributes to be representative" (Anastasi & Urbina, 2001, p. 3).

Theoretically speaking, the attribute vector of an item in a cognitive diagnostic test corresponds to a column of the potential **Q**-matrix. According to the augment algorithm, a column in the potential **Q**-matrix can be represented by the Boolean union of the columns of the reachability matrix, which means that the items that correspond to the columns in the reachability matrix are representative.

The knowledge state corresponds to the columns of student **Q**-matrix. The knowledge state, and the item attribute vector are in the same "space." These have important implications for cognitive diagnosis test design and item selection strategies (Wu et al., 2011; Ding et al., 2011, 2010). Also, as discussed above, the reachability matrix can represent the cognitive model. Thus, items that correspond to the reachability matrix can be seeded into a test, which is of great benefit to the improvement of the validity of a cognitive diagnostic test. In addition, the reachable matrix **R** also plays a compensatory role when other columns of the test **Q**-matrix are miscalibrated (Gan et al., 2014).

### 3.2 Determining the Set of Knowledge States

Given the attributes and their hierarchical relationships, Tatsuoka's (1995, 2009) reduction algorithm can be used to determine all the knowledge states, which lays the foundation for the diagnostic analysis, classification of respondents, and subsequent tailored remediation. The augment algorithm based on the reachability matrix can also be applied to determine the set of non-zero knowledge states directly. For example, for a convergent structure of four attributes (Leighton et al., 2004), Gierl et al. (2000) considered that the attributes and their hierarchy corresponded to six non-zero knowledge states; according to the corresponding reachability matrix, four non-zero knowledge states were obtained by using the augment algorithm. Inferred from the footnote of Gierl et al. (2007, p. 256), the result obtained via the augment algorithm was correct.

Thus, while determining the set of knowledge states,the reachability matrix can be used as a "sieve" in the reduction algorithm and as a "seed" in the augment algorithm. Under certain conditions, the augment algorithm shows greater efficiency than the reduction algorithm (Yang & Ding, 2011).

### 3.3 Constructing Non-Zero Knowledge States—Two Expressions

The representation of non-zero knowledge states is not necessarily unique. There are two expressions: a redundant expression and a concise expression (Ding, Luo, Wang, Xiong, Duan, & Song, 2018).

**Definition 2. Assume that x is a non-zero knowledge state, let $\mathbf{S}_X = \{\mathbf{r}|\mathbf{r} \leq \mathbf{x}$ and r is a column of the reachability matrix$\}$, the vector in $\mathbf{S}_X$ is called the component of x. The Boolean union for all the vectors in $\mathbf{S}_X$ is a redundant expression of x. If both $r_1$ and $r_2$ are components of x, and $r_1 \leq r_2$, then delete $r_1$ (called "delete small and keep large"). After "deleting small and keeping large" for all components of x, the Boolean union of the remaining components of x is called the concise expression formula of x.**

**Example 3 (cont. Example 1).** To briefly explain the above two expressions of the non-zero knowledge states, let us consider the reachability matrix in Example 1. It is assumed that $\mathbf{x} = \begin{pmatrix}1\\1\\1\\0\end{pmatrix}$, then

$$\mathbf{S}_X = \left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}\right\}.$$

According to the above definitions, the redundant expression of $\mathbf{x} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \vee \begin{pmatrix}1\\1\\0\\0\end{pmatrix} \vee \begin{pmatrix}1\\0\\1\\0\end{pmatrix}$, and the concise expression of $\mathbf{x} = \begin{pmatrix}1\\1\\0\\0\end{pmatrix} \vee \begin{pmatrix}1\\0\\1\\0\end{pmatrix}$. If all non-zero knowledge states apply redundant expressions (or concise expressions), the representation of non-zero knowledge states is unique.

When the attributes are non-compensable and adopt the 0-1 scoring, the reachability matrix can be used as the test **Q**-matrix, which maps the knowledge states to itself. An important generalization of this conclusion is that the necessary **Q**-matrix can create a one-to-one correspondence between the sets of knowledge states and ideal response patterns. This conclusion plays an essential role in the development of the implicit cognitive diagnosis model. It is often used in the simulation studies that verify the accuracy rates of newly developed cognitive diagnostic models, such as cognitive diagnosis with the state transition diagram in terms of computer science (Lin & Ding, 2007), the Hamming Distance Discriminating method (Z. Luo et al., 2015), the Generalized Distance Discriminant method (J. Sun et al., 2011), the Polytomous Generalized Distance Discriminant method (PGDD; J. Sun et al., 2013), a non-parametric cognitive diagnostic method of mixed scoring—Manhattan Distance Discriminating method (Kang et al., 2019), and so forth. The proof of this conclusion is tedious; however, it is relatively easy to prove it using the concept of the redundant expression (Ding, Luo, Wang, Xiong, Duan, & Song, 2018).

Moreover, the concept of redundant expression plays an important role in the development of simple **Q**-matrix calibration methods (Ding, Luo, Wang, & Xiong, 2018; W. Wang, Wang, Song, & Gao, 2018). This method can also be used to calibrate the polytomous **Q**-matrix.

### 3.4 Introducing Boolean Intersection and Boolean Union into the Column Set of Student Q-Matrix to Form Lattice

Determining the relationship in terms of magnitudes between vectors of column set of the student **Q**-matrix yields the so-called partially ordered set. Since the augment algorithm shows that the partial order set is closed for the Boolean union, any two elements have the least upper bound. The proof of the closure of Boolean intersection by the set $S$ of reachability matrix columns and zero columns are relatively is complicated. We refer interested readers to Yang and Ding (2011) for more details.

Every non-zero column of any **Q**-matrix is a Boolean union of a reachability matrix's columns, and the operations between the Boolean intersection and Boolean union meet the distributive property, so the Boolean intersection of any two columns of the **Q**-matrix is close to the column set of the student **Q**-matrix. The Boolean intersection is the greatest lower bound of any two elements in this partially ordered set, so any two elements in this partially ordered set have the greatest lower bound.

According to the definition of Lattice (see Zuo et al., 1982), the set of columns of the student **Q**-matrix forms the Lattice. This conclusion lays a theoretical foundation for generating a new method of **Q**-matrix calibration—the cross-difference method (W. Wang et al., 2011; W. Wang & Song, 2015), and for developing item selection strategies for online multi-stage computerized adaptive tests (CD-OMST) with cognitive diagnosis (F. Luo et al., 2018, 2016).

### 3.5 Deriving Theoretical Construct Validity (TCV)

Let **Q** denote the test **Q**-matrix and **R** denote the reachability matrix, and assume the numbers of columns of the **Q**-matrices obtained based on **R** and augmented from the test **Q**-matrix to be $m_1$ and $m_2$, respectively. Then the theoretical construct validity (TCV) of **Q** is $(m_2+1)/(m_1+1)$ under the evenly distributed knowledge states; otherwise it can be calculated based on the distribution of knowledge states. The 1 in the numerator and denominator indicates that the knowledge states of the zero-vector are considered (Ding, Mao, et al., 2012). According to Proposition 2, if the test **Q**-matrix is necessary, then TCV = 1. From Proposition 3, if the test **Q**-matrix is not necessary, then TCV $< 1$, and at this time, the more columns of the reachability matrix the test **Q**-matrix contains, the closer the TCV value is to 1.

Therefore, the TCV can be used to evaluate a test before it is administered. If the TCV is too low, the test should be revised, such as adding columns in the reachability matrix to the test **Q**-matrix, and adding corresponding test items.

## 4 Summary and Discussion

The non-zero columns obtained by Tatsuoka's reduction algorithm are the same as those derived from the augment algorithm based on the reachability matrix. However, the content of the augment algorithm is more informative.

The augment algorithm, reduction algorithm, and cleansing algorithm indicate the special role of the reachability matrix. There is a one-to-one correspondence between the reachability matrix and the adjacency matrix, indicating that the reachability matrix can represent the cognitive model and have significant implications on the design of the cognitive diagnostic test. Further research indicated that the role of reachability matrix in the design of cognitive diagnosis test is irreplaceable (Ding et al., 2016). A simple method for the calibration of **Q**-matrix can be developed based on the reachability matrix and the redundant expression (Ding, Luo, Wang, Xiong, Duan, & Song, 2018; W. Wang, Song, & Ding, 2018; W. Wang, Wang, et al., 2018). The augment algorithm reveals that the reachability matrix is the basis of constructing the **Q**-matrix. The augment algorithm also shows the algebraic properties of columns of the **Q**-matrix. From the reachability matrix and the test **Q**-matrix, the TCV index derived from the augment algorithm is used to evaluate the quality of the test **Q**-matrix from a new perspective; the augment algorithm based on the reachability matrix can also be used in the development of cognitive diagnosis model and so on.

This paper also shows that the validity of a test that is designed based on cognitive diagnosis can be evaluated by calculating the test TCV. If the diagnostic analysis is applied for a revised test, do the results require further correction? For example, based on a revised test, some methods (e.g., asking experts to identify) can be used to obtain the associated test **Q**-matrix (which corresponds to some attribute hierarchy, such as $X$). Based on the cognitive diagnosis analysis, a specific attribute hierarchy (such as $Y$) can also be extracted from the set of knowledge states of the examinees. A question that worth asking is that if $X$ and $Y$ are consistent? If not, what causes the inconsistency? (see Ding, Mao, et al., 2012; X. Wang et al., 2019). It is also worth studying that if the test **Q**-matrix (Y. Sun et al., 2014) extracted directly from the response matrix is consistent with the **Q**-matrix obtained from the set of knowledge states
.

The column set $S$ of the student **Q**-matrix forms a lattice. The applications of lattices in cognitive diagnosis need further study. Applications of the fact that combining $S$ and the redundant expression makes the columns of the reachability matrix similar to the "basis" of linear space (that is, the columns of the reachability matrix can uniquely represent the other columns of the **Q**-matrix, if the redundant expression of a non-zero knowledge state is employed) also requires more investigation, since the "basis" of linear space in linear algebra is of great significance.

This paper emphasizes attributes and their hierarchical relationships because they can represent the cognitive models. However, it is well-known that the accuracy of a cognitive model is difficult to guarantee. Therefore, in general, field experts tend to assume that the hierarchical attribute relationship is independent. In this way, it will not be problematic as long as the attributes and their numbers can be defined appropriately, and items that correspond to these attributes can be written. This is because the set of knowledge states cannot go beyond the set of attributes relative to the independent attributes. It is the safest way to deal with the hierarchical relationships. However, if certain attribute cannot be used alone, the relationships among these attributes are not independent because at least one of them has a prerequisite.

One of the reasons why it is difficult to correctly specify the hierarchical relationships is that the hierarchy can change dynamically. One of the common trends is "a rising tide lifts all boats," that is to say, while the cognitive levels of the respondents increase, the division of attributes may change accordingly. Attributes that are considered "fundamental" (usually the shared prerequisite attributes) would have the most significant changes. For example, when a student starts learning multiplication (e.g., "$3 \times 4$"), addition operator should be viewed as the prerequisite of multiplication operator. The **Q**-matrix should reflect that addition is the prerequisite attribute of multiplication. However, after introducing the concept of multiplication, it is possible to test only the multiplication table (multiplication) without labeling addition as a prerequisite attribute for multiplication. Of course, there are also examples where attributes in certain items cannot exist without prerequisite attributes. For example, while calculating $1/3-1/4$, "addition and subtraction with the same denominator" and "common multiples" are the prerequisite attributes of "addition and subtraction with different denominators."

The independence is also a kind of hierarchical relationship. The convergence, divergence, unstructured, and linear structures introduced by Leighton et al. (2004) are closer relationships. Their related reachability and adjacent matrices are more complex than those of the independent attributes.

The adjacent and reachability matrices are easy to obtain based on the attributes and their hierarchical relationships. However, for retrofitting data, the test **Q**-matrices are difficult to specify. Even if a test **Q**-matrix can be defined, it is unknown whether we can apply the reduction algorithm to this **Q**-matrix and find a reachability matrix. This is because sometimes the number of columns (the number of items that can be separated independently, which are the item-attribute vectors) of the **Q**-matrix obtained via the reduction algorithm is smaller than the number of its rows (the number of attributes). Even if the reachability matrix can be extracted from the **Q**-matrix, its hierarchical relationships (or adjacent matrix) can be unclear. Further investigation is required, such as using line-by-line comparison methods (Tatsuoka, 1995, 2009) to obtain the hierarchical relationships of the attributes.

The algorithms described in this paper do not require big data. These algorithms are intuitive and easy to understand, and some of them can be extended to derive other conclusions. We have provided the mathematical proofs for some the algorithms. Compared with other complex algorithms, these algorithms have certain advantages. As it is say that "everything has its merits and demerits," we hope that these algorithms can provide useful references.

## References

Anastasi, A., & Urbina, S. (2001). *Psychological testing (in Chinese; X. Miao & P. Zhu, Trans.)*. Zhejiang Education Publishing Group. (Original work published 1997).

Cai, Y., Tu, D., & Ding, S. (2018). Theorems and methods of a complete Q-matrix with attribute hierarchies under restricted Q-matrix design. *Frontiers in Psychology*, *9*, 1413.

Chen, J., & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, *37*(6), 419–437.

Ding, S., & Luo, F. (2005). Algorithm: From poset to Hasse diagram (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *29*, 150–152.

Ding, S., & Luo, F. (2013). An efficient algorithm of deriving Hasse diagram from the reachibility matrix of a partial order relation (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *37*, 441–444.

Ding, S., Luo, F., Cai, Y., Lin, H., & Wang, X. (2008). Complement to Tatsuoka's Q-matrix theory. In A. Okada, T. Imaizumi, & T. Hoshino (Eds.), *New trends in psychometrics* (pp. 417–424). Universal Academy Press.

Ding, S., Luo, F., & Wang, W. (2012). Extension to Tatsuoka's Q-matrix theory (in Chinese). *Psychological Exploration*, *32*, 417–422.

Ding, S., Luo, F., Wang, W., & Xiong, J. (2016). Dichotomous and polytomous Q-matrix theory. In L. A. van der Ark, D. M. Bolt, W.-C. Wang, J. A. Douglas, & M. Wiberg (Eds.), *Quantitative psychology* (Vol. 167, pp. 277–289). Springer.

Ding, S., Luo, F., Wang, W., & Xiong, J. (2018). A simple method to specify Q-matrix (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *42*, 130–133.

Ding, S., Luo, F., Wang, W., Xiong, J., Duan, H., & Song, L. (2018). Different expressions of a knowledge state and their applications. In M. Wiberg, S. Culpepper, R. Janssen, J. Gonzalez, & D. Molenaar (Eds.), *Quantitative psychology* (Vol. 233, pp. 377–383). Springer.

Ding, S., Mao, M., Wang, W., Luo, F., & Cui, Y. (2012). Evaluating the consistency of test items relative to the cognitive model for educational cognitive diagnosis (in Chinese). *Acta Psychologica Sinica*, *44*, 1535–1546.

Ding, S., Wang, W., & Luo, F. (2012). Q-matrix and Q-matrix theory in cognitive diagnosis (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *36*, 441–445.

Ding, S., Wang, W., Luo, F., & Xiong, J. (2015). The polytomous Q-matrix theory (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *39*, 365–370.

Ding, S., Wang, W., & Yang, S. (2011). The design of cognitive diagnostic test blueprints (in Chinese). *Journal of Psychological Science*, *34*, 258–265.

Ding, S., Yang, S., & Wang, W. (2010). The importance of reachability matrix in constructing cognitively diagnostic testing (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *34*, 490–494.

Ding, S., Zhu, Y., Lin, H., & Cai, Y. (2009). Modification of Tatsuoka's Q-matrix theory (in Chinese). *Acta Psychologica Sinica*, *41*, 175–181.

Duda, R. O., Hart, P. E., & Stork, D. G. (2003). *Pattern classification* (2nd ed.). CITIC Publishing House.

Gan, Z., Wang, W., & Ding, S. (2014). The research on the remedial effects of reachability matrix when identifying an item attribute incorrectly (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *38*, 600–604.

Gierl, M. J., Leighton, J. P., & Hunka, S. (2000). Exploring the logic of Tatsuoka's rule-space model for test development and analysis. *Educational Measurement: Issues and Practice*, *19*, 34–44.

Gierl, M. J., Leighton, J. P., & Hunka, S. (2007). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 242–274). Cambridge University Press.

Kang, C., Yang, Y., & Zeng, P. (2019). An approach to cognitive diagnosis: The manhattan distance discriminating method (in Chinese). *Journal of Psychological Science*, *42*, 455–462.

Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule-space approach. *Journal of Educational Measurement*, *41*, 205–237.

Lin, H., & Ding, S. (2007). An exploration and realization of computerized adaptive testing with cognitive diagnosis (in Chinese). *Acta Psychologica Sinica*, *39*, 747–753.

Luo, F., Ding, S., Wang, X., & Xiong, J. (2016). Application study on online multistage intelligent adaptive testing for cognitive diagnosis. In L. A. van der Ark, D. M. Bolt, W.-C. Wang, J. A. Douglas, & M. Wiberg (Eds.), *Quantitative psychology research* (pp. 265–274). Springer.

Luo, F., Wang, X., Ding, S., & Xiong, J. (2018). The design and selection strategies of adaptive multi-group testing for cognitive diagnosis (in Chinese). *Journal of Psychological Science*, *41*, 720–726.

Luo, Z., Li, Y., Yu, X., Gao, C., & Peng, Y. (2015). A simple cognitive diagnosis method based on Q-matrix theory (in Chinese). *Acta Psychologica Sinica*, *47*, 264–272.

Sun, J., Xin, T., Zhang, S., & de la Torre, J. (2013). A polytomous extension of generalized distance discriminating method. *Applied Psychological Measurement*, *37*, 503–521.

Sun, J., Zhang, S., Xin, T., & Bao, Y. (2011). A cognitive diagnosis method based on Q-matrix and generalized distance (in Chinese). *Acta Psychologica Sinica*, *43*, 1095–1102.

Sun, Y., Ye, S., Inoue, S., & Sun, Y. (2014). Alternating recursive method for Q-matrix learning. In *Proceedings of the 7th international conference oneducational data mining (edm)* (pp. 14–20).

Tatsuoka, K. K. (1995). Architecture of knowledge structure and cognitive diagnosis: A statistical pattern recognition and classification approach. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (pp. 327–361). Erlbaum.

Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the rule space method*. Routledge.

Wang, C., & Lu, J. (2021). Learning attribute hierarchies from data: two exploratory approaches. *Journal of Educational and Behavioral Statistics*, *46*, 58–84.

Wang, W., Ding, S., & You, X. (2011). Online item attribute identification in cognitive diagnostic computerized adaptive testing (in Chinese). *Acta Psychologica Sinica*, *43*, 964–976.

Wang, W., & Song, L. (2015). *Research of theory and technology in educational cognitive diagnosis assessment (in Chinese)*. Beijing Normal University Publishing Group.

Wang, W., Song, L., & Ding, S. (2018). A method for Q-matrix specification based on the reachability matrix (in Chinese). *Journal of Psychological Science*, *41*, 968–975.

Wang, W., Wang, T., Song, L., & Gao, P. (2018). The method for compensatory model's Q-matrix specification based on the reachability matrix (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *42*, 441–446.

Wang, X., Ding, S., & Luo, F. (2019). Q-matrix and its applications in cognitive diagnosis (in Chinese). *Journal of Psychological Science*, *42*, 739–746.

Wu, Z., Gan, D., & Ding, S. (2011). The research on reachability matrix in item-selection strategy of cognitive diagnosis (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *35*, 422–426.

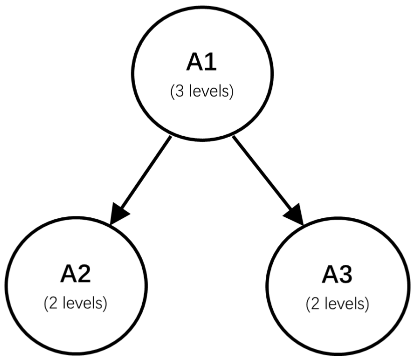
Yang, S., Cai, S., Ding, S., Lin, H., & Ding, Q. (2008). Augment algorithm for reduced Q-matrix (in Chinese). *Journal of Lanzhou University (Natural Sciences)*, *44*, 87–91,96.

Yang, S., & Ding, S. (2011). Theory and method for predicating valid objects (in Chinese). *Journal of Jiangxi Normal University (Natural Science)*, *35*, 1–4.

Yang, S., Ding, S., & Ding, Q. (2010). The incremental augment algorithm of Qr matrix. *Transactions of Nanjing University of Aeronautics and Astronautics*, *27*, 183–189.

Zuo, X., Li, W., & Liu, Y. (1982). *Discrete mathematics*. Shanghai Scientific and Technological Literature Press.

## Appendix A: Example of Constructing a Quasi-Reachability Matrix




$$\mathbf{R}_2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [r_{ij}]$$

where $r_{21} = r_{31} = 0,\ r_{12} = 1,\ r_{32} = 0,\ r_{13} = 1,\ r_{23} = 0.$

$$\mathbf{R}_2 \to \begin{bmatrix} 1 & 2 & 3 & 1(1 & 1) & 1(1 & 1) \\ 0(1 & 1 & 1) & 1 & 2 & 0(1 & 1) \\ 0(1 & 1 & 1) & 0(1 & 1) & 1 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 2 & 3 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$
$$= \mathbf{R}_p$$

## Appendix B: Introduction and Examples of the Reduction Algorithm and the Cleansing Algorithm

The reduction algorithm means that if a column **q** of **Q**-matrix can be represented by the Boolean union of other columns in the **Q**-matrix, then **q** will be deleted from the **Q**-matrix. This process continues until each column in **Q** cannot be represented by the Boolean union of other columns.This algorithm is called the reduction algorithm.

Note that the length of the Boolean union of two different Boolean vectors (whose elements are either 0 or 1) must be longer than any of the two given vectors. Therefore, we first arrange the columns of the **Q**-matrix according to the number of non-zero elements in each corresponding column in ascending order. Suppose that the **Q**-matrix has $K$ rows and $m$ columns and $s \leq m$, then start with the $s^{th}$ column (denoted as **x**) and check whether the Boolean union of some columns to the left yields the result equal to **x**. If yes, then delete **x**; otherwise, then $s = m - 1$. Follow the above steps to check the $s^{th}$ column, decide whether it should be deleted, and continue until all columns cannot be represented by the Boolean columns added on the left side.

This algorithm is the inverse of the augment algorithm. Through the reduction algorithm, the potential **Q**-matrix ($\mathbf{Q}_p$) can be reduced to the reachability matrix **R**. Note here the potential **Q**-matrix can be reduced to the reachability matrix, but not all **Q**-matrices can be reduced to the reachability matrix; some **Q**-matrices after the reduction can have fewer columns than rows. Then, according to the reachability matrix **R**, the hierarchical relationships among the attributes are obtained by pairwise comparisons of its rows (Ding & Luo, 2013).

Ding and Luo (2005) presented a cleansing algorithm for drawing a partial order relation Hasse graph. For the completeness of the discussion, we describe steps of this algorithm below, where **E** is the identity matrix.

**Algorithm** The Cleansing Algorithm

**Output:** The relation matrix corresponding to the Hasse graph (**M** is a $n \times n$ partial order relation matrix)

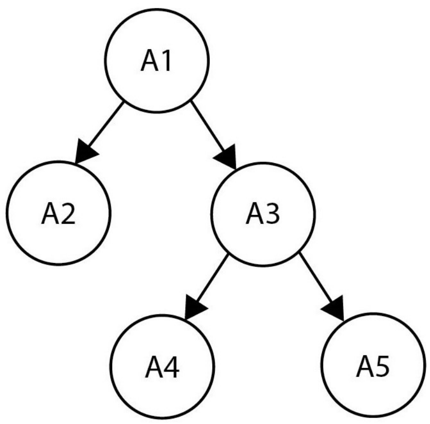
```
U := M - E
for i = 1 to n do
    for j = 1 to n do
        for k = 1 to n do
            u_ik = u_ik − u_ik * u_ij * u_jk
  # U = (u_ij) is the relation matrix corresponding to the Hasse graph
```

The above algorithm has $2n^3$ multiplications and $n^3$ subtractions. Because $u_{ij} \in \{0,1\}$, $u_{ik} * u_{ij} * u_{jk}$ can also be represented by the "logical AND" (i.e., Boolean intersection).

The relation matrix corresponding to the Hasse graph is the adjacency matrix of the cognitive model, or these attributes and the hierarchies among them, which is anti-reflexive, anti-symmetric, and anti-transitive.

Figure 2 presents a hierarchy of five attributes.

Figure 2
*A Hierarchical Structure Involving Five Attributes*




The adjacency matrix corresponding to the hierarchical relationships shown in Figure 2 is $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$,

the reachability matrix is $\mathbf{R} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

From the augment algorithm, we get

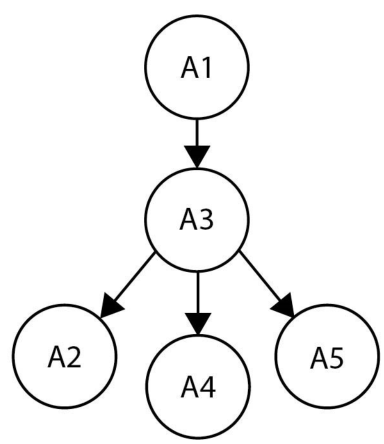
$$\mathbf{Q}_p = \left[\begin{array}{ccccc|ccc|c|cc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & & 1 & 1 \end{array}\right]$$

If we delete the second column in $\mathbf{Q}_p$ (which also corresponds to the second column in the reachability matrix), the resulting $\mathbf{Q}$ is not a necessary $\mathbf{Q}$-matrix. We then apply the reduction algorithm to the $\mathbf{Q}$-matrix:

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} = \mathbf{Q}_1$$

Columns 9, 8, 7, and 6 of the given $\mathbf{Q}$ satisfy the requirements of $\mathbf{q}_9 = \mathbf{q}_7 \vee \mathbf{q}_8$, $\mathbf{q}_8 = \mathbf{q}_3 \vee \mathbf{q}_4$, $\mathbf{q}_7 = \mathbf{q}_4 \vee \mathbf{q}_5$, $\mathbf{q}_6 = \mathbf{q}_3 \vee \mathbf{q}_5$, such that $\mathbf{Q}_1$ can be obtained by reducing columns 6, 7, 8, and 9 of $\mathbf{Q}$. Pairwise comparisons of rows in $\mathbf{Q}_1$ show that A1 is a prerequisite for attributes A2, A3, A4, and A5, while A3 is a prerequisite for attributes A2, A4, and A5. Note that $\mathbf{Q}_1$ and $\mathbf{R}$ are different. The hierarchical attribute relationships derived from $\mathbf{Q}_1$ is shown in Figure 3. The difference vector formed between the third row and the second row of $\mathbf{Q}_1$ has all non-negative entries. In contrast, the difference vector formed between the third row and the second row of the reachability matrix $\mathbf{R}$ cannot satisfy the requirement of all components being non-negative.

Figure 3
*A "Compromised" Structure Involving Five Attributes*




From this example, we see that if the adjacency matrix $\mathbf{Q}$ does not satisfy the requirement that the reachability matrix $\mathbf{R}$ is its submatrix (i.e., it is not a necessary $\mathbf{Q}$-matrix), the method introduced by Tatsuoka (1995, 2009) may cause the attribute hierarchy extracted from $\mathbf{Q}$ to be inconsistent with those extracted from the reachability matrix.