Journal of Classification 36:541-565 (2019)
DOI: 10.1007/s00357-018-9278-6

# Attribute Hierarchy Models in Cognitive Diagnosis: Identifiability of the Latent Attribute Space and Conditions for Completeness of the Q-Matrix

Hans-Friedrich Köhn

University of Illinois at Urbana-Champaign

Chia-Yi Chiu

Rutgers, the State University of New Jersey

**Abstract:** Educational researchers have argued that a realistic view of the role of attributes in cognitively diagnostic modeling should account for the possibility that attributes are not isolated entities, but interdependent in their effect on test performance. Different approaches have been discussed in the literature; among them the proposition to impose a hierarchical structure so that mastery of one or more attributes is a prerequisite of mastering one or more other attributes. A hierarchical organization of attributes constrains the latent attribute space such that several proficiency classes, as they exist if attributes are not hierarchically organized, are no longer defined because the corresponding attribute combinations cannot occur with the given attribute hierarchy. Hence, the identification of the latent attribute space is often difficult—especially, if the number of attributes is large. As an additional complication, constructing a complete Q-matrix may not at all be straightforward if the attributes underlying the test items are supposed to have a hierarchical structure. In this article, the conditions of identifiability of the latent space if attributes are hierarchically organized and the conditions of completeness of the Q-matrix are studied.

**Keywords:** Cognitive diagnosis; Attribute hierarchy; Latent attribute apace; Q-Matrix; Completeness; DINA model; General DCMs.

---

This material is based upon work supported by the National Science Foundation under Grant No. 1552563.

Corresponding Author's Address: H-F. Köhn, University of Illinois at Urbana-Champaign, 425 Psychology Bldg., 603 E. Daniel Street M/C 716, Champaign, IL 61820, email: hkoehn@illinois.edu.

Published online: 16 November 2018

## 1. Introduction

In the past decade, cognitive diagnosis has emerged as a new paradigm of educational measurement that seeks to combine rigorous psychometric standards with the goals of formative assessment. (DiBello, Roussos, and Stout, 2007; Haberman and von Davier, 2007; Leighton and Gierl, 2007; Rupp, Templin, and Henson, 2010). Mastery of the instructional content is explicitly targeted to provide immediate feedback on students' strengths and weaknesses in terms of skills learned and skills needing study. Within the cognitive diagnosis framework, skills, specific knowledge, aptitudes—any cognitive characteristic required to perform tasks—are collectively referred to as "attributes" that an examinee may or may not posses. Ability in a domain is perceived as a composite of these attributes such that different combinations define profiles of distinct proficiency classes. Examinees are to be assigned to the proficiency classes based on their test performance. The test items themselves are also characterized by individual attribute profiles that determine which specific skills are required to respond correctly to an item. These item-attribute associations define the Q-matrix of a test (Tatsuoka, 1985). The entire set of realizable attribute profiles, given a particular set of attributes, is called the latent attribute space (e.g., Tatsuoka, 2009). Thus, the Q-matrix is a subset of the latent attribute space. The identification of the latent attribute space and the correct specification of the Q-matrix are key conditions of the validity of cognitive diagnosis. As another essential requirement, the Q-matrix must be complete, which means it must guarantee the identifiability of all realizable proficiency classes among examinees (Chiu, Douglas, and Li, 2009). An incomplete Q-matrix causes examinees to be assigned to proficiency classes to which they do not belong.

Educational researchers have argued that a realistic view of the role of attributes in cognitively diagnostic modeling should account for the possibility that attributes are not isolated entities, but interdependent in their effect on test performance. Different approaches have been discussed in the literature. De la Torre and Douglas (2004) proposed a higher-order model linking the latent attribute space to an underlying multivariate normal distribution with possibly correlated dimensions. Haertel and Wiley (1994), and Leighton, Gierl, and Hunka (2004) (see also Leighton and Gierl, 2007; Tatsuoka, 2009; Templin and Bradshaw, 2014) developed a different approach to account for potential relations/interdependencies among attributes by imposing a hierarchical structure so that mastery of one or more attributes is a prerequisite of mastering one or more other attributes. A hierarchical organization of attributes constrains the latent attribute space such that several proficiency classes, as they exist if attributes are not hierarchically organized, are no longer defined because the corresponding attribute combinations cannot occur with the given attribute hierarchy. Hence, the identification of the latent attribute space is often difficult—especially, if the number of attributes is large. As an additional complication, constructing a complete Q-matrix may not at all be straightforward if the attributes underlying the test items are supposed to have a hierarchical structure.

In this article, the conditions of identifiability of the latent space if attributes are hierarchically organized and the conditions of completeness of the Q-matrix are studied. The next section briefly reviews definitions and technical key concepts of the cognitive diagnosis framework and attribute hierarchy models in particular. Theoretical propositions and proofs are presented in the subsequent sections. The discussion section addresses two related theoretical questions and concludes with some practical recommendations.

## 2. Review of Technical Key Concepts

### 2.1 Cognitive Diagnosis and Diagnostic Classification Models

Suppose ability in a given domain is modeled as a composite of $K$ latent binary attributes $\alpha_1, \alpha_2, \ldots, \alpha_K$. The $K$-dimensional binary vector $\boldsymbol{\alpha}_m = (\alpha_{m1}, \alpha_{m2}, \ldots, \alpha_{mK})^T$ denotes the attribute profile of proficiency class $\mathcal{C}_m$, $m = 1, 2, \ldots, M$, where the $k^{th}$ entry, $\alpha_{mk} \in \{0, 1\}$, indicates (non-)mastery of the corresponding attribute. If the attributes do not have a hierarchical structure, then there are $2^K = M$ distinct proficiency classes. The entire set of their attribute vectors defines the latent attribute space (Tatsuoka, 2009). The attribute profile $\boldsymbol{\alpha}_{i \in \mathcal{C}_m}$ of examinee $i \in \mathcal{C}_m$, $i = 1, 2, \ldots, N$, is usually written as $\boldsymbol{\alpha}_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{iK})^T$. (Throughout the text, the terms "profile" and "vector" are used interchangeably; the transpose of vectors or matrices is denoted by a superscripted $T$; the "prime" notation is reserved for distinguishing between vectors or their scalar entries. For brevity, the examinee index $i$ is omitted if the context permits; for example, $\boldsymbol{\alpha}_i$ is simply written as $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)^T$.)

Models of cognitive diagnosis—henceforth, diagnostic classification models (DCMs)—are constrained latent class models such that the latent variable proficiency-class membership—associated with mastery of a particular attribute set—determines the probability of a correct item response. A plethora of DCMs has been proposed in the literature (e.g., Fu and Li, 2007; Rupp and Templin, 2008). They differ in how the functional relation between mastery of attributes and the probability of a correct item response is modeled. DCMs have been distinguished based on criteria like compensatory versus non-compensatory (lacking certain attributes can/cannot be compensated for by possessing other attributes), or conjunctive (all attributes required for an item must be mastered; mastering only a subset of them results in a success probability equal to that of an examinee mastering none of the attributes) versus disjunctive (mastery of a subset of the required attributes is a sufficient condition for maximizing the probability of a correct item response) (de la Torre and Douglas, 2004; Henson, Templin, and Willse, 2009; Maris, 1999). The Deterministic Inputs, Noisy "AND" Gate (DINA) Model (Haertel, 1989; Junker and Sijtsma, 2001; Macready and Dayton, 1977) is the standard example of a conjunctive DCM. Let $Y_{ij}$ denote the manifest response of examinee $i$ to binary item $j$, $j = 1, 2, \ldots, J$. For the DINA model, the conditional probability of examinee $i$ having attribute profile $\boldsymbol{\alpha}_i$ answering item $j$ correctly is given by the item response function (IRF)

$$P(Y_{ij} = 1 \mid \boldsymbol{\alpha}_i) = (1 - s_j)^{\eta_{ij}} g_j^{(1-\eta_{ij})} \tag{1}$$

subject to $0 < g_j < 1 - s_j < 1 \; \forall j$. The conjunction parameter $\eta_{ij}$, defined as $\eta_{ij} = \prod_{k=1}^{K} \alpha_{ik}^{q_{jk}}$ indicates whether examinee $i$ has mastered all the attributes needed to answer item $j$ correctly. The item-related parameters $s_j = P(Y_j = 0 \mid \eta_j = 1)$ and $g_j = P(Y_j = 1 \mid \eta_j = 0)$ formalize the probabilities of slipping (failing to answer item $j$ correctly despite having the skills required to do so) and guessing, (answering item $j$ correctly despite lacking the skills required to do so), respectively. Thus, $\eta_{ij}$ can be interpreted as the ideal item response when neither slipping nor guessing occur. The Deterministic Inputs, Noisy "OR" gate (DINO) model (Templin and Henson, 2006) is the prototypic disjunctive DCM (i.e., mastery of a subset of the required attributes is a sufficient condition for maximizing the probability of a correct item response). Define the disjunction parameter $\omega_{ij} = 1 - \prod_{k=1}^{K} (1 - \alpha_{ik})^{q_{jk}}$ that indicates whether at least one of the attributes associated with item $j$ has been mastered. (Like $\eta_{ij}$ in the DINA model, $\omega_{ij}$ is the ideal item response.) The IRF of the DINO model is $P(Y_{ij} = 1 \mid \boldsymbol{\alpha}_i) = (1 - s_j)^{\omega_{ij}} g_j^{(1-\omega_{ij})}$.

## 2.2 The Q-Matrix and the Completeness Property

Consider a test of $J$ items for assessing ability in the domain. Each individual item $j$, $j = 1, 2, \ldots, J$, is associated with a $K$-dimensional binary vector $\mathbf{q}_j$ called the item-attribute profile, where $q_{jk} = 1$ if a correct answer requires mastery of the $k^{th}$ attribute, and 0 otherwise. Item-attribute profiles consisting entirely of zeroes are inadmissible, because they correspond to items that require no attributes at all. Hence, given $K$ attributes, there are at most $2^K - 1$ distinct item-attribute profiles. The item-attribute profiles of a test are collected into the $J \times K$ Q-matrix, $\mathbf{Q} = \{q_{jk}\}_{(J\times K)}$, that summarizes the constraints specifying the associations between items and attributes (Tatsuoka, 1985).

Fitting educational data with a DCM requires that the Q-matrix underlying the test items be known and complete. A Q-matrix is said to be complete if it allows for the identification of the attribute profiles of all realizable proficiency classes—formally, $\boldsymbol{S}(\boldsymbol{\alpha}) = \boldsymbol{S}(\boldsymbol{\alpha}') \Rightarrow \boldsymbol{\alpha} = \boldsymbol{\alpha}'$, where $\boldsymbol{S}(\boldsymbol{\alpha}) = E(\boldsymbol{Y} \mid \boldsymbol{\alpha})$ denotes the expectation of the item response vector $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_J)'$, given attribute profile $\boldsymbol{\alpha}$ (Chiu et al., 2009; Chiu and Köhn, 2015). As a small-scale example, consider the two Q-matrices $\mathbf{Q}_{1:3}$ and $\mathbf{Q}_{4:6}$, with columns corresponding to the attributes $\alpha_1$, $\alpha_2$, and $\alpha_3$ (so, $K = 3$) and rows containing the item attribute profiles $\mathbf{q}_j$; the matrix subscripts, $1 : 3$ and $4 : 6$, refer to the item indices $j = 1, 2, 3$ and $j = 4, 5, 6$, respectively.

$$\mathbf{Q}_{1:3} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad \mathbf{Q}_{4:6} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Completeness of $\mathbf{Q}_{1:3}$ and $\mathbf{Q}_{4:6}$ is evaluated by checking the definition $\boldsymbol{S}(\boldsymbol{\alpha}) = \boldsymbol{S}(\boldsymbol{\alpha}') \Rightarrow \boldsymbol{\alpha} = \boldsymbol{\alpha}'$ for each pair, $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$, of the $M = 8$ proficiency classes. For the DINA model, the elements of the $J$-dimensional vector of expected item responses, $\boldsymbol{S}(\boldsymbol{\alpha})$, are defined as $S_j(\boldsymbol{\alpha}) = E(Y_j \mid \boldsymbol{\alpha}) = P(Y_j \mid \boldsymbol{\alpha}) = (1 - s_j)^{\eta_j} g_j^{(1-\eta_j)}$. Hence,

$$S_j(\boldsymbol{\alpha}) = E(Y_j \mid \boldsymbol{\alpha}) = \begin{cases} 1 - s_j & \text{if } \eta_j = 1 \\ g_j & \text{if } \eta_j = 0 \end{cases}$$

For the DINA model, $\mathbf{Q}_{1:3}$ is not complete, whereas $\mathbf{Q}_{4:6}$ is complete, as the computation of the expected item-response profiles $\boldsymbol{S}(\boldsymbol{\alpha})$ shows:

| $\boldsymbol{\alpha}$ | $\mathbf{Q}_{1:3}$ | | | $\mathbf{Q}_{4:6}$ | | |
|---|---|---|---|---|---|---|
| | $\mathbf{q}_1 = (011)$ | $\mathbf{q}_2 = (101)$ | $\mathbf{q}_3 = (110)$ | $\mathbf{q}_4 = (100)$ | $\mathbf{q}_5 = (010)$ | $\mathbf{q}_6 = (001)$ |
| | $S_1(\boldsymbol{\alpha})$ | $S_2(\boldsymbol{\alpha})$ | $S_3(\boldsymbol{\alpha})$ | $S_4(\boldsymbol{\alpha})$ | $S_5(\boldsymbol{\alpha})$ | $S_6(\boldsymbol{\alpha})$ |
| (000) | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ | $g_6$ |
| (100) | $g_1$ | $g_2$ | $g_3$ | $1 - s_4$ | $g_5$ | $g_6$ |
| (010) | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $1 - s_5$ | $g_6$ |
| (001) | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ | $1 - s_6$ |
| (110) | $g_1$ | $g_2$ | $1 - s_3$ | $1 - s_4$ | $1 - s_5$ | $g_6$ |
| (101) | $g_1$ | $1 - s_2$ | $g_3$ | $1 - s_4$ | $g_5$ | $1 - s_6$ |
| (011) | $1 - s_1$ | $g_2$ | $1 - s_3$ | $g_4$ | $1 - s_5$ | $1 - s_6$ |
| (111) | $1 - s_1$ | $1 - s_2$ | $1 - s_3$ | $1 - s_4$ | $1 - s_5$ | $1 - s_6$ |

$\mathbf{Q}_{1:3}$ does not allow to distinguish between all $\boldsymbol{\alpha}$; for example, $\boldsymbol{S}(\boldsymbol{\alpha}_1) = \boldsymbol{S}(\boldsymbol{\alpha}_2) = (g_1, g_2, g_3)^T$ but $\boldsymbol{\alpha}_1 = (000) \neq \boldsymbol{\alpha}_2 = (100)$. Thus, $\mathbf{Q}_{1:3}$ is not complete. However, if the single-attribute items 4–6 in $\mathbf{Q}_{4:6}$ are included, then $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}' \Rightarrow \boldsymbol{S}(\boldsymbol{\alpha}) \neq \boldsymbol{S}(\boldsymbol{\alpha}')$. In fact, Chiu et al. (2009) proved that $\mathbf{Q}$ is complete for the DINA model if and only if each attribute is represented by at least one single-attribute item—that is, $\mathbf{Q}$ has rows, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_K$, among its $J$ rows, where $\mathbf{e}_k$ denotes a $1 \times K$ vector, with the $k^{th}$ element $e_k$ equal to 1, and all other entries equal to 0 (For the DINO model too, Q-completeness is guaranteed if and only if $\mathbf{Q}$ includes all $K$ single-attribute items; Chiu and Köhn, 2015.)

## 3. Hierarchical Attribute Structures and Attribute Hierarchy Models

Leighton, Gierl, and Hunka (2004) described in detail the concept of an attribute hierarchy and various forms of attribute hierarchies. Imposing a hierarchy on the attributes means that mastery of certain attributes is a prerequisite for the mastery of other attributes. Attribute hierarchies are often displayed in the form of tree-shaped graphs, where the vertices represent attributes, and the directed edges connecting the vertices—the "branches" of the tree—signify the prerequisite relation. Leighton, Gierl, and Hunka (2004) presented four types of tree-shaped attribute hierarchies that they called "linear" (the attributes are ordered along a line such that the prerequisite relation translates into precedence), "convergent" (one attribute is required for mastering two subsequent attributes that, in turn, are required for one or more succeeding attribute(s)), "divergent" (two attributes share one preceding attribute as prerequisite), and "unstructured" (one attribute is the prerequisite for multiple attributes that themselves are not hierarchically ordered). As was mentioned earlier, the existence of a hierarchy among the attributes immediately affects the structure of the latent attribute space because certain proficiency classes, which would be defined were there no attribute hierarchy, are no longer defined simply because the corresponding attribute profiles cannot occur with the given attribute hierarchy.

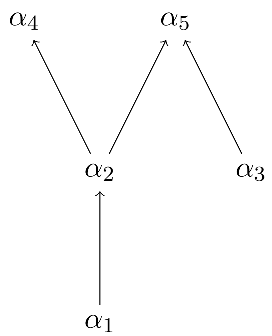
**Example 1.** As a simple example, consider the "divergent" attribute hierarchy among $\alpha_1$, $\alpha_2$, and $\alpha_3$:

$$
\begin{tikzcd}
\alpha_2 & & \alpha_3 \\
& \alpha_1 \arrow[ul] \arrow[ur] &
\end{tikzcd}
$$

Instead of $M = 8$ proficiency classes when attributes are not hierarchically organized, now there are only $M = 5$ proficiency classes $(000)$; $(100)$; $(110)$; $(101)$; and $(111)$; proficiency classes $(010)$, $(001)$, and $(011)$ are no longer defined. The constraints imposed on the latent attribute space by the attribute hierarchy also affect the conditions for Q-completeness. As an example, consider again $\mathbf{Q}_{1:3}$ and $\mathbf{Q}_{4:6}$. Neither of them is complete if the divergent hierarchy is imposed on $\alpha_1$, $\alpha_2$, and $\alpha_3$. As an important detail, notice that within this convergent attribute hierarchy Items 1, 5, and 6 are no longer admissible, because they do not include $\alpha_1$ as a required attribute. However, $\alpha_1$ is a prerequisite for attributes $\alpha_2$ and $\alpha_3$; casually speaking—and perhaps more to the point—$\alpha_2$ and $\alpha_3$ "cannot be had" without $\alpha_1$:

| | $\mathbf{Q}_{1:3}$ | | | $\mathbf{Q}_{4:6}$ | | |
|---|---|---|---|---|---|---|
| $\boldsymbol{\alpha}$ | $\mathbf{q}_1 = (011)$ | $\mathbf{q}_2 = (101)$ | $\mathbf{q}_3 = (110)$ | $\mathbf{q}_4 = (100)$ | $\mathbf{q}_5 = (010)$ | $\mathbf{q}_6 = (001)$ |
| | $S_1(\boldsymbol{\alpha})$ | $S_2(\boldsymbol{\alpha})$ | $S_3(\boldsymbol{\alpha})$ | $S_4(\boldsymbol{\alpha})$ | $S_5(\boldsymbol{\alpha})$ | $S_6(\boldsymbol{\alpha})$ |
| $(000)$ | | $g_2$ | $g_3$ | $g_4$ | | |
| $(100)$ | **not** | $g_2$ | $g_3$ | $1 - s_4$ | **not** | **not** |
| $(110)$ | **defined** | $g_2$ | $1 - s_3$ | $1 - s_4$ | **defined** | **defined** |
| $(101)$ | | $1 - s_2$ | $g_3$ | $1 - s_4$ | | |
| $(111)$ | | $1 - s_2$ | $1 - s_3$ | $1 - s_4$ | | |

**Example 2.** The next example is more complex and involves $K = 5$ attributes. Without a hierarchy imposed on the attributes, there are $2^5 = 32 = M$ proficiency classes. Assume that the five attributes are organized in the hierarchy displayed by this tree graph:




Due to the complex prerequisite structure, most of the theoretically realizable proficiency classes are not defined. For example, the single-attribute profiles $(01000) = \mathbf{e}_2$, $(00010) = \mathbf{e}_4$, and $(00001) = \mathbf{e}_5$ are no longer defined and must be replaced by $(11000)$, $(11010)$, and $(11001)$, respectively, so that the prerequisite relations defining the hierarchy among attributes are satisfied. In fact, out of the original 32 proficiency classes only 10 are defined:

| No. | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ | $\alpha_4$ | $\alpha_5$ |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 1 | 0 | 0 |
| 3 | 1 | 0 | 0 | 0 | 0 |
| 4 | 1 | 0 | 1 | 0 | 0 |
| 5 | 1 | 1 | 0 | 0 | 0 |
| 6 | 1 | 1 | 1 | 0 | 0 |
| 7 | 1 | 1 | 0 | 1 | 0 |
| 8 | 1 | 1 | 1 | 1 | 0 |
| 9 | 1 | 1 | 1 | 0 | 1 |
| 10 | 1 | 1 | 1 | 1 | 1 |

These 10 proficiency class attribute vectors define the latent attribute space associated with the attribute hierarchy displayed in the tree graph shown above. Thus, the item attribute vectors that define the complete Q-matrix must be a subset of the 10 proficiency class attribute vectors. The crucial question, however, remains how to find them among the vectors that define the latent attribute space of a hierarchy—and how to determine these attribute vectors in the first place? In summary, cognitively diagnostic modeling using an attribute hierarchy requires: (a) the identification of the latent attribute space—that is, the set of proficiency class profiles that are realizable/legitimate given the postulated attribute hierarchy; (b) the identification of the subset of item attribute vectors of the latent attribute space that guarantee a complete Q-matrix.

In the next section, a framework for addressing these tasks is developed based on lattice theory.

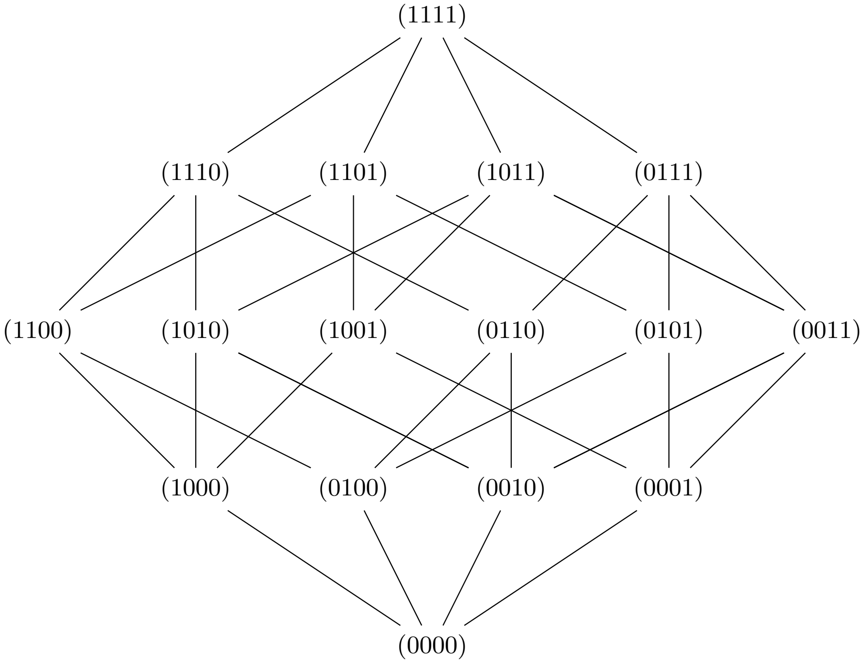
## 4. Attribute Hierarchies: A Lattice-Theoretical Approach

Several definitions are required. First, let $\mathcal{L}$ denote the set of binary $K$-dimensional attribute vectors of all realizable proficiency classes called the latent attribute space. If there is no hierarchy defined on the set of attributes $\{\alpha_1, \alpha_2, \ldots, \alpha_K\}$, then $\mathcal{L} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_M\}$, where $\boldsymbol{\alpha}_1$ denotes the $K$-dimensional all-zeroes vector and $\boldsymbol{\alpha}_M$ the $K$-dimensional all-ones vector. Second, the $K$ attributes $\alpha_k \in \{0, 1\}$ are called Boolean variables; the $M$ vectors $\boldsymbol{\alpha}$ are called Boolean vectors. Third, the order relation $\leq$ for two binary $K$-dimensional attribute vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ is defined such that $\boldsymbol{\alpha}' \leq \boldsymbol{\alpha}$ if and only if $\alpha'_k \leq \alpha_k \ \ \forall k$. The relation $\leq$ is reflexive, antisymmetric, and transitive; hence, it is a partial order. Thus, $\mathcal{L}$ is a partially ordered set (poset) written $\langle \mathcal{L}, \leq \rangle$. (If the partial ordering has been well-identified and misunderstandings are impossible, then $\langle \mathcal{L}, \leq \rangle$ is often written simply as $\mathcal{L}$.) Fourth, $\langle \mathcal{L}, \leq \rangle$ is called a lattice if each of its (finite) subsets has an infimum and a supremum. In general, for Boolean vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, the infimum is defined as the conjunction $\mathbf{u} \wedge \mathbf{v} = \mathbf{u} \cdot \mathbf{v} = (u_1 v_1, u_2 v_2, \ldots, u_n v_n)$. The supremum is defined as the disjunction $\mathbf{u} \vee \mathbf{v} = \mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$. The Boolean operators $\wedge$ and $\vee$ are idempotent; for example, $1 \vee 1 = 1$. (Thus, the infimum and the supremum of the singletons, $\{\boldsymbol{\alpha}_1\}, \{\boldsymbol{\alpha}_2\}, \ldots, \{\boldsymbol{\alpha}_M\}$, just equal $\boldsymbol{\alpha}$; for example, $\inf\{\boldsymbol{\alpha}_1\} = \sup\{\boldsymbol{\alpha}_1\} = (00\ldots0)$, or $\inf\{\boldsymbol{\alpha}_M\} = \sup\{\boldsymbol{\alpha}_M\} = (11\ldots1)$, and so on.) Fifth, a lattice is called complete if it has universal bounds—that is, a least element $0$ and a largest element $I$; in case of $\mathcal{L}$, $O = \mathbf{0}_K = \boldsymbol{\alpha}_1 = (00\ldots0)$ and $I = \mathbf{1}_K = \boldsymbol{\alpha}_M = (11\ldots1)$. (Thus, the infimum and supremum of every subset of $\mathcal{L}$ must be also in $\mathcal{L}$.—The authoritative reference on lattice theory is Birkhoff (1970); the second chapter in Heijmans (1994) provides a succinct but thorough introduction to basic concepts.)

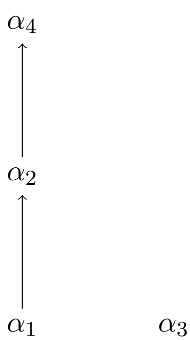
**Example 3.** As an example, consider the case of $K = 4$ attributes, $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, that do not have a hierarchical structure—here is the corresponding graph:

$$\begin{array}{cc} \alpha_1 & \alpha_2 \\ \\ \alpha_3 & \alpha_4 \end{array}$$

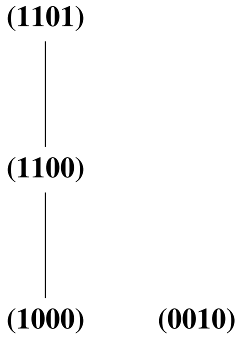
In total, there are $2^4 = 16$ realizable proficiency classes. The latent attribute space $\mathcal{L}$ is defined by the set of the 16 binary four-dimensional attribute vectors. The poset $\langle \mathcal{L}, \leq \rangle$ is a complete lattice, with $O = \mathbf{0}_4 = (0000)$ and $I = \mathbf{1}_4 = (1111)$. Infimum and supremum are defined for all subsets of $\mathcal{L}$; for example, $\inf\{(1000), (0100)\} = (1000) \wedge (0100) = (0000)$ and $\sup\{(1000), (0100)\} = (1000) \vee (0100) = (1100)$. The lattice $\mathcal{L}$ is displayed by the Hasse diagram below, with the proficiency classes vertically ordered and connected by an edge if they are in the relation $\leq$; because order relations are transitive, any relation between proficiency classes can be deduced by following the edges upward. The Hasse diagram further shows how all realizable proficiency classes can be obtained through the Boolean operations $\wedge$ and $\vee$ performed on the four single-attribute profiles $(1000) = \mathbf{e}_1$, $(0100) = \mathbf{e}_2$, $(0010) = \mathbf{e}_3$, and $(0001) = \mathbf{e}_4$ that can be derived from the graph by inspection.



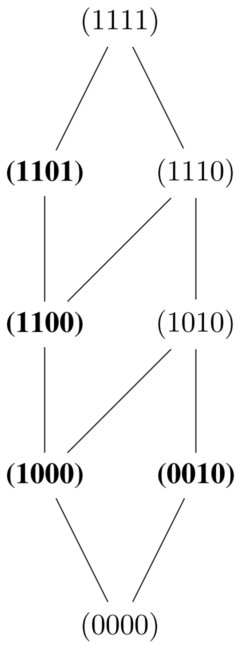

**Example 4.** Now, assume that the four attributes have a hierarchical structure such that mastery of $\alpha_1$ is the prerequisite of mastering $\alpha_2$, and both are prerequisites of mastering $\alpha_4$; $\alpha_3$ is not part of the hierarchy. The tree graph of this hierarchy is



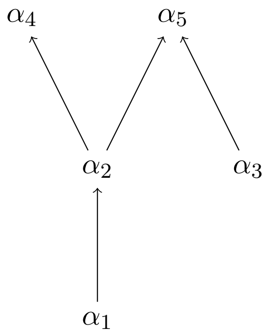

The proficiency class $(0100)$ is no longer defined, because mastery of $\alpha_1$ is a prerequisite of mastering $\alpha_2$. Thus, $(0100)$ is replaced by $(1100)$. From the tree graph, four basic proficiency classes are derived by inspection:



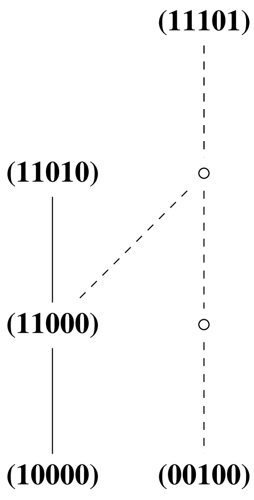

But how to construct the entire latent attribute space $\mathcal{L}$? $\mathcal{L}$ is still a lattice. Then, the infimum and supremum of any subset of the four basic attribute profiles are defined; they must be also in $\mathcal{L}$. For example, $\inf\{(1000), (0010)\} = (1000) \wedge (0010) = (0000) \in \mathcal{L}$, which is the "all-zeros" attribute vector (i.e., $\mathbf{0}_4$); or $\sup\{(1000), (0010)\} = (1000) \vee (0010) = (1010) \in \mathcal{L}$; $\sup\{(1100), (1010)\} = (1100) \vee (1010) = (1110) \in \mathcal{L}$, and so on. The Hasse diagram of the lattice of the entire latent attribute space is



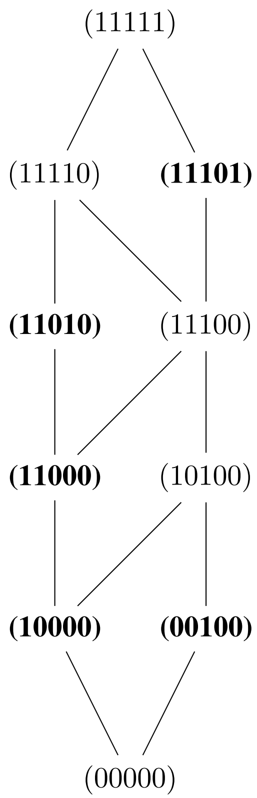

**Example 2 Revisited.** Consider again Example 2 involving $K = 5$ attributes; the tree graph of the attribute hierarchy is repeated here for convenience:




From the tree graph, five basic attribute vectors, (10000), (11000), (00100), (11010), and (11001), can be derived by inspection:




(The dashed lines and the two "$\circ$" indicate that the hierarchical relations between the basic attribute vectors "leap" across levels characterized by mastery of two and three attributes; see also the Hasse diagram of the complete lattice on the next page.) The lattice of the entire latent attribute space is reconstructed from the five attribute vectors; for example, $(10000) \vee (00100) = (10100)$, $(11000) \vee (00100) = (11100)$, $(00100) \vee (11010) = (11110)$, and $(11010) \vee (11101) = (11111)$. Note that $(10000) \wedge (00100) = (11000) \wedge (00100) = (00100) \wedge (11010) = (00000)$.




The insights from the four examples presented can be summarized in the following claim. A set of attribute profiles called “basic attribute vectors” can be derived by inspection from the tree graph of any attribute hierarchy. These basic attribute vectors must be a subset of the latent attribute space $\mathcal{L}$. Because $\mathcal{L}$ is a lattice, the latent attribute space can be reconstructed in its entirety from the basic attribute vectors using the operations $\vee$ and $\wedge$. In addition, it can be shown that any Q-matrix that contains the unique $K \times K$ submatrix formed by the basic attribute vectors is complete for the DINA model, given any attribute hierarchy. Consider again Example 2, with a hierarchy involving $K = 5$ attributes. The Q-matrix derived from the basic attribute vectors is

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

Computing the expected item response vectors $\boldsymbol{S}(\boldsymbol{\alpha})$ confirms its completeness:

| $\boldsymbol{\alpha}$ | $\mathbf{Q}$ | | | | |
|---|---|---|---|---|---|
| | $\mathbf{q}_1 = (00100)$ | $\mathbf{q}_2 = (10000)$ | $\mathbf{q}_3 = (11000)$ | $\mathbf{q}_4 = (11001)$ | $\mathbf{q}_5 = (11110)$ |
| | $S_1(\boldsymbol{\alpha})$ | $S_2(\boldsymbol{\alpha})$ | $S_3(\boldsymbol{\alpha})$ | $S_4(\boldsymbol{\alpha})$ | $S_5(\boldsymbol{\alpha})$ |
| $(00000)$ | $g_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ |
| $(00100)$ | $1-s_1$ | $g_2$ | $g_3$ | $g_4$ | $g_5$ |
| $(10000)$ | $g_1$ | $1-s_2$ | $g_3$ | $g_4$ | $g_5$ |
| $(11000)$ | $g_1$ | $1-s_2$ | $1-s_3$ | $g_4$ | $g_5$ |
| $(10100)$ | $1-s_1$ | $1-s_2$ | $g_3$ | $g_4$ | $g_5$ |
| $(11100)$ | $1-s_1$ | $1-s_2$ | $1-s_3$ | $g_4$ | $g_5$ |
| $(11010)$ | $g_1$ | $1-s_2$ | $1-s_3$ | $g_4$ | $g_5$ |
| $(11110)$ | $1-s_1$ | $1-s_2$ | $1-s_3$ | $g_4$ | $1-s_5$ |
| $(11101)$ | $1-s_1$ | $1-s_2$ | $1-s_3$ | $1-s_4$ | $g_5$ |
| $(11111)$ | $1-s_1$ | $1-s_2$ | $1-s_3$ | $1-s_4$ | $1-s_5$ |

The next section presents formal proofs of these claims.

## 5. Theorems and Proofs: Latent Attribute Space and Q-Completeness for Attribute Hierarchy Models

Recall that the infimum and the supremum of the latent attribute space —the universal bounds $\mathbf{0}_K$ and $\mathbf{1}_K$—are included in $\mathcal{L}$. The only exception occurs when the attributes have a linear order (called a "chain" in lattice theory); then, $\mathbf{0}_K \notin \mathcal{L}$. In this article, however, the assumption is that $\mathbf{0}_K = (00\ldots0) \in \mathcal{L}$.

**Lemma 1.** *Define $\mathcal{H}_k = \{l \mid \alpha_l$ is a prerequisite of $\alpha_k\} \cup \{k\}$ and $\mathbf{E}_k = \sum_{l\in\mathcal{H}_k} \mathbf{e}_l$ for attribute $k$. Let $\mathcal{E} = \{\boldsymbol{E}_1, \boldsymbol{E}_2, \cdots, \boldsymbol{E}_K\}$; the power set of $\mathcal{E}$ is denoted by $\mathcal{P}(\mathcal{E})$. If $\mathcal{S}$ is defined as $\mathcal{S} = \{\sup(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathcal{P}(\mathcal{E})\}$, then the latent space $\mathcal{L} = \mathcal{S}$.*

*Proof.* The least and the largest element, $\mathbf{0}_K$ and $\mathbf{1}_K$, respectively, are in $\mathcal{L}$. Hence, $\mathcal{L}$ is a complete lattice—which implies that the supremum and the infimum of every subset of $\mathcal{L}$ exist and are also in $\mathcal{L}$. Now, because $\mathcal{E} \subset \mathcal{L}$, for every $\mathbf{u} \in \mathcal{P}(\mathcal{E})$—hence, every $\sup(\mathbf{u}) \in \mathcal{S}$—it must be true that $\mathbf{u} \subset \mathcal{L}$. Also, as $\mathcal{L}$ is complete, $\mathbf{u} \subset \mathcal{L}$ implies $\sup(\mathbf{u}) \in \mathcal{L}$ and thus, $\mathcal{S} \subseteq \mathcal{L}$. On the other hand, the definition suggests that $\mathbf{E}_k$ is the least q-vector measuring attribute $k$ for all $k$. Now, suppose some $\mathbf{v} \in \mathcal{L}$, but $\mathbf{v} \notin \mathcal{S}$. Further, assume there exists a $\mathbf{w} \notin \mathcal{P}(\mathcal{E})$ such that $\mathbf{v} = \sup(\mathbf{w})$. Therefore, there must exist some $k' \in \{1, 2, \ldots, K\}$ such that $\mathbf{E}_{k'} \leq \mathbf{v}$; but $\mathbf{E}_{k'} \notin \mathcal{E}$, which contradicts the assumption that $\mathcal{E}$ contains all $\mathbf{E}_k$. Hence, $\mathbf{v}$ must be in $\mathcal{S}$ implying that $\mathcal{L} \subseteq \mathcal{S}$. Due to the results $\mathcal{L} \supseteq \mathcal{S}$ and $\mathcal{L} \subseteq \mathcal{S}$, $\mathcal{L} = \mathcal{S}$.
■

Alternatively, Lemma 1 states that the latent attribute space $\mathcal{L}$ of any hierarchical attribute structure can be reconstructed as a set of linear combinations of the basic attribute vectors in $\mathcal{E}$. (Recall that $\mathbf{E}_k$ are Boolean vectors; hence, $\mathbf{E}_k \vee \mathbf{E}_{k'} = \mathbf{E}_k + \mathbf{E}_{k'}$.)

Theorem 1 states that the basic attribute vectors of any attribute hierarchy, $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K$, form a Q-matrix that is complete for the DINA model.

**Theorem 1.** *$\mathcal{H}_k$ and $\mathbf{E}_k$ are as defined in Lemma 1. For the DINA model, $\mathbf{Q}$ is complete if and only if it contains among its $J$ rows the $K$ vectors $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K$.*

*Proof.* ($\Rightarrow$) Consider latent classes $\boldsymbol{\alpha} = \boldsymbol{E}_k$ and $\boldsymbol{\alpha}' = \sum_{l \in \mathcal{H}_k/\{k\}} \mathbf{e}_l$.

$$S_j(\boldsymbol{\alpha}) \quad = \quad \begin{cases} 1 - s_j & \text{if } \mathbf{q}_j = \boldsymbol{\alpha} \\ 1 - s_j & \text{if } \mathbf{q}_j \leq \boldsymbol{\alpha}' \\ g_j & \text{otherwise} \end{cases} \tag{2}$$

and

$$S_j(\boldsymbol{\alpha}') \quad = \quad \begin{cases} g_j & \text{if } \mathbf{q}_j = \boldsymbol{\alpha} \\ 1 - s_j & \text{if } \mathbf{q}_j \leq \boldsymbol{\alpha}' \\ g_j & \text{otherwise}\,. \end{cases} \tag{3}$$

Suppose that $\mathbf{E}_k$ is missing from the Q-matrix. According to Equations 2 and 3, $S_j(\boldsymbol{\alpha}) = S_j(\boldsymbol{\alpha}')$ for all $j$ implying that $\boldsymbol{S}(\boldsymbol{\alpha}) = \boldsymbol{S}(\boldsymbol{\alpha}')$. Hence, $\mathbf{Q}$ is not complete.

($\Leftarrow$) Suppose that the vectors $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K$ are in the Q-matrix. Reorder the rows of $\mathbf{Q}$ such that $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K$ are moved to the first $K$ rows. Suppose $\boldsymbol{\alpha} \in \mathcal{L}$. Due to Lemma 1, there exists a set $\mathcal{B} \subseteq \{1, \ldots, K\}$ such that $\boldsymbol{\alpha} = \bigvee_{k \in \mathcal{B}} \mathbf{E}_k = \sum_{k \in \mathcal{B}} \mathbf{E}_k = \sum_{k \in \bigvee_{l \in \mathcal{B}} \mathcal{H}_l} \mathbf{e}_k$ (recall that $\mathbf{E}_k$ are Boolean vectors). Equivalently, for any $k$,

$$\alpha_k = \begin{cases} 1 & \text{if } k \in \bigvee_{l \in \mathcal{B}} \mathcal{H}_l \\ 0 & \text{otherwise}\,. \end{cases} \tag{4}$$

For the DINA model, the expected response of item $k$, given q-vector $\mathbf{E}_k$ and $\boldsymbol{\alpha}$, can be expressed as

$$S_k(\boldsymbol{\alpha}) = \begin{cases} 1 - s_k & \text{if } \boldsymbol{\alpha} \geq \boldsymbol{E}_k \\ g_k & \text{otherwise.} \end{cases} \tag{5}$$

Because $\boldsymbol{\alpha} = \bigvee_{k \in \mathcal{B}} \mathbf{E}_k = \sum_{k \in \bigvee_{l \in \mathcal{B}} \mathcal{H}_l} \mathbf{e}_k$, Equation 5 can be rewritten as

$$S_k(\boldsymbol{\alpha}) = \begin{cases} 1 - s_k & \text{if } k \in \bigvee_{l \in \mathcal{B}} \mathcal{H}_l \\ g_k & \text{otherwise.} \end{cases} \tag{6}$$

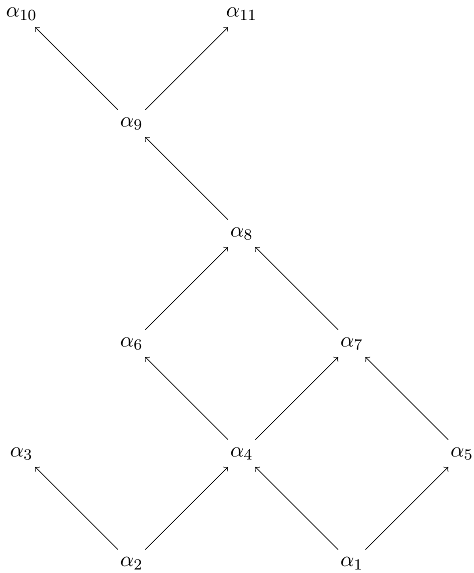
Equations 4 and 6 then imply that

$$S_k(\boldsymbol{\alpha}) \quad = \quad \begin{cases} 1 - s_k & \text{if } \alpha_k = 1 \\ g_k & \text{if } \alpha_k = 0\,. \end{cases}$$

Recall that any two latent profiles in $\mathcal{L}$ are distinct: $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$. Thus, it must be true that $\boldsymbol{S}_{1:K}(\boldsymbol{\alpha}) \neq \boldsymbol{S}_{1:K}(\boldsymbol{\alpha}')$. Therefore, $\boldsymbol{S}(\boldsymbol{\alpha}) \neq \boldsymbol{S}(\boldsymbol{\alpha}')$ regardless of whether $S_j(\boldsymbol{\alpha}) = S_j(\boldsymbol{\alpha}')$ for $j = (K+1), \ldots, J$.
■

## 6. How to Use These Results in Practice: A Complex Example

The results of Lemma 1 and Theorem 1 translate directly into a procedure useful in practice for identifying the latent attribute space for any attribute hierarchy and the corresponding Q-matrix that is complete for the DINA model. As an illustration, consider the rather complex hierarchy involving $K = 11$ attributes displayed in the tree graph below.

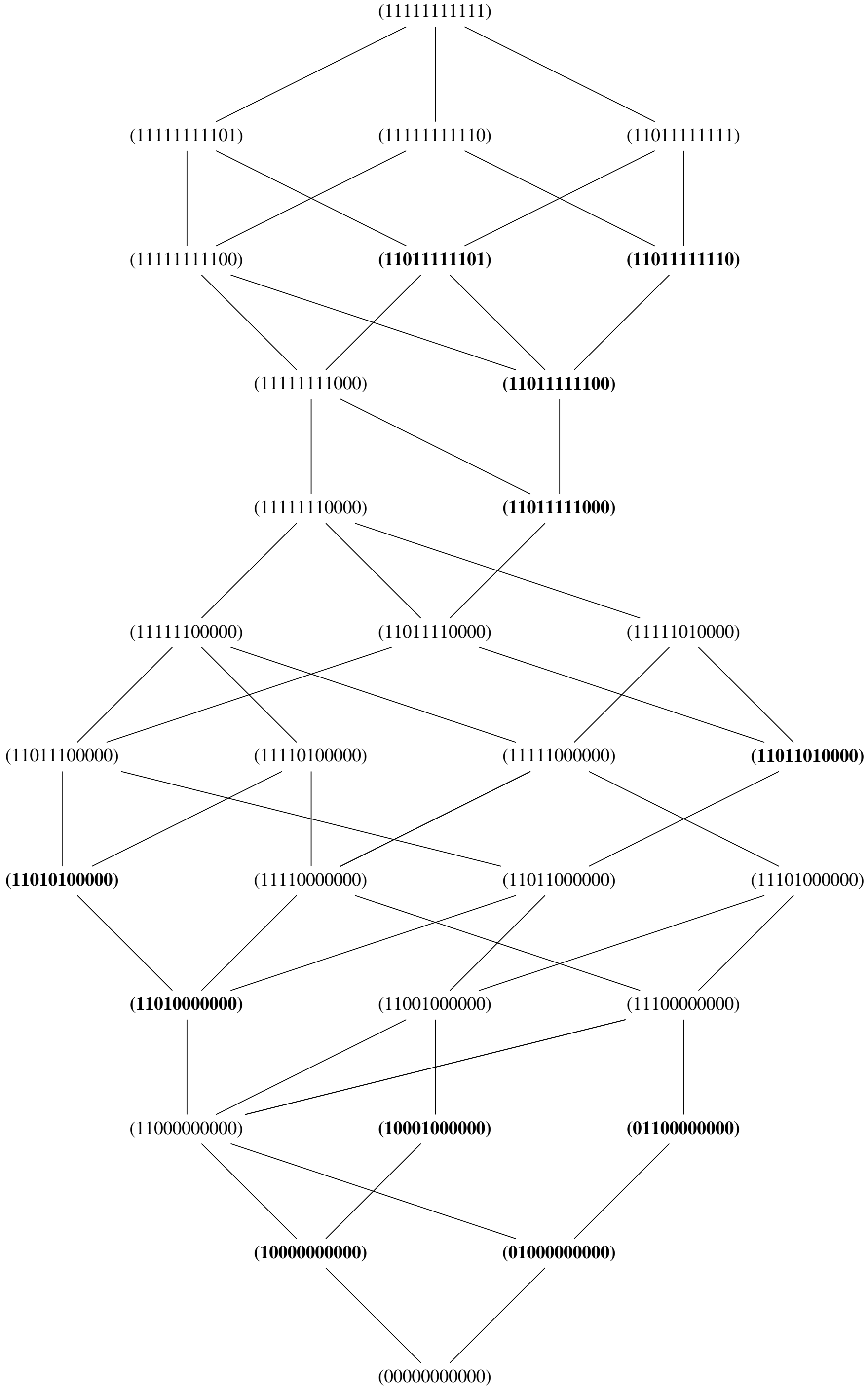



Without a hierarchy imposed on the attributes, there would be $M = 2^{11} = 2048$ proficiency classes. Most of these proficiency classes, however, are no longer defined. From the tree graph, 11 basic attribute vectors can be derived by inspection:

| Basic Attribute Vectors | Length $\|\|\mathbf{E}_k\|\|$ |
|---|---|
| $\mathbf{E}_1 = (10000000000)$ | 1 |
| $\mathbf{E}_2 = (01000000000)$ | 1 |
| $\mathbf{E}_3 = (01100000000)$ | 2 |
| $\mathbf{E}_4 = (11010000000)$ | 3 |
| $\mathbf{E}_5 = (10001000000)$ | 2 |
| $\mathbf{E}_6 = (11010100000)$ | 4 |
| $\mathbf{E}_7 = (11011010000)$ | 5 |
| $\mathbf{E}_8 = (11011111000)$ | 7 |
| $\mathbf{E}_9 = (11011111100)$ | 8 |
| $\mathbf{E}_{10} = (11011111110)$ | 9 |
| $\mathbf{E}_{11} = (11011111101)$ | 9 |

The lattice of the latent attribute space $\mathcal{L}$ shown in the Hasse diagram on the next page contains the attribute vectors of 31 proficiency classes. They were constructed from the 11 basic attribute vectors using $\wedge$ and $\vee$ as stated in Lemma 1. The Hasse diagram has 12 different levels that can be numbered in ascending order from 0 to 11 according to the length of the attribute vectors, $||\boldsymbol{\alpha}||$, at each level. Beginning with the two shortest basic vectors, $\mathbf{E}_1 = (10000000000)$ and $\mathbf{E}_2 = (01000000000)$, the attribute vector at level 0 is obtained by $\mathbf{E}_1 \wedge \mathbf{E}_2$. Attribute vector $(11000000000)$ is obtained by $\mathbf{E}_1 \vee \mathbf{E}_2$; the other two $\boldsymbol{\alpha}$ at level 2 are given as $\mathbf{E}_5$ and $\mathbf{E}_3$. At level 3, $\mathbf{E}_4$ is given, the remaining two 3-attribute $\boldsymbol{\alpha}$ are obtained as the supremum of the pairs from level 2, $\sup\{(11000000000), \mathbf{E}_5\}$ and $\sup\{(11000000000), \mathbf{E}_3\}$. Note that $\sup\{\mathbf{E}_3, \mathbf{E}_5\} = (11101000000) \in \mathcal{L}$ but "jumps" from level 2 to 4. The procedure of obtaining the $\boldsymbol{\alpha}$ at the different levels of $\mathcal{L}$ is summarized as follows:

| $\|\|\boldsymbol{\alpha}\|\|$ | $\boldsymbol{\alpha}$ | | Operation resulting in $\boldsymbol{\alpha}$ |
|---|---|---|---|
| 0 | $(00000000000)$ | $=$ | $\inf\{\mathbf{E}_1, \mathbf{E}_2\} = (10000000000) \wedge (01000000000)$ |
| 1 | $(10000000000)$ | | **$\mathbf{E}_1$ given** |
| 1 | $(01000000000)$ | | **$\mathbf{E}_2$ given** |
| 2 | $(11000000000)$ | $=$ | $\sup\{\mathbf{E}_1, \mathbf{E}_2\} = (10000000000) \vee (01000000000)$ |
| 2 | $(10001000000)$ | | **$\mathbf{E}_5$ given** |
| 2 | $(01100000000)$ | | **$\mathbf{E}_3$ given** |
| 3 | $(11010000000)$ | | **$\mathbf{E}_4$ given** |
| 3 | $(11001000000)$ | $=$ | $\sup\{(11000000000), \mathbf{E}_5\} = (11000000000) \vee (10001000000)$ |
| 3 | $(11100000000)$ | $=$ | $\sup\{(11000000000), \mathbf{E}_3\} = (11000000000) \vee (01100000000)$ |
| 4 | $(11010100000)$ | | **$\mathbf{E}_6$ given** |
| 4 | $(11110000000)$ | $=$ | $\sup\{\mathbf{E}_4, (11100000000)\} = (11010000000) \vee (11100000000)$ |
| 4 | $(11011000000)$ | $=$ | $\sup\{\mathbf{E}_4, (11001000000)\} = (11010000000) \vee (11001000000)$ |
| 4 | $(11101000000)$ | $=$ | $\sup\{(11010000000), (11100000000)\} = (11010000000) \vee (11100000000)$ |
| $\vdots$ | | | |
| 11 | $(11111111111)$ | $=$ | $\sup\{(11111111110), (11111111101), (11011111111)\} = (11111111110) \vee (11111111101) \vee (11011111111)$ |


(11111111111)
(11111111101)
(11111111110)
(11011111111)
(11111111100)
(11011111101)
(11011111110)
(11111111000)
(11011111100)
(11111110000)
(11011111000)
(11111100000)
(11011110000)
(11111010000)
(11011100000)
(11110100000)
(11111000000)
(11011010000)
(11010100000)
(11110000000)
(11011000000)
(11101000000)
(11010000000)
(11001000000)
(11100000000)
(11000000000)
(10001000000)
(01100000000)
(10000000000)
(01000000000)
(00000000000)

Finally, recall that including the unique submatrix formed by the 11 basic attribute vectors $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_{11}$ in $\mathbf{Q}$ guarantees its completeness for the DINA model.

## 7. Discussion and Conclusion

Hierarchical attribute models have been developed to account for potential relations/interdependencies among the attributes supposed to underlie a given test. Attributes are said to have a hierarchical structure if mastery of one or more attributes is a prerequisite of mastering one or more other attributes. If attributes are hierarchically organized, then the latent attribute space is constrained such that many proficiency classes are no longer defined because their attribute profiles cannot occur, given the prerequisite structure associated with the particular attribute hierarchy. Hence, the identification of the latent attribute space and the specification of a complete Q-matrix—both key requirers for the validity of cognitive diagnosis—can be difficult; especially, if the number of attributes is large and the hierarchy is complex. As was proven in this article, the latent attribute space can be constructed from a collection of basic attribute vectors that can be derived by inspection from the tree graph of the attribute hierarchy in question. Also, it was proven that any Q-matrix that contains the unique $K \times K$ submatrix formed by these basic attribute vectors is guaranteed to be complete for the given attribute hierarchy if the underlying DCM is the DINA model. Based on these results, procedures for identifying the latent attribute space and establishing Q-completeness can be readily devised that facilitate the use of attribute hierarchies in practice. Two theoretical questions remain to be addressed.

First, in recent years, general DCMs have received considerable attention as a framework for expressing the specific functional relation between attribute mastery and the probability of a correct item response of diverse DCMs in a unified mathematical form and parameterization (see von Davier, 2005, 2008; Henson et al., 2009; de la Torre, 2011). The immediate question seems why not extend the results of this study to general DCMs? The answer is two-fold. First, the procedure for identifying the latent attribute space suggested by the theoretical results of this study can also be applied to general DCMs. Second, however, because completeness is not an intrinsic property of the Q-matrix (i.e., completeness does not depend on the chosen attribute q-vectors alone), but can only be assessed in reference to the specific DCM supposed to underlie the data—that is, the Q-matrix of a given test can be complete for one model but incomplete for another—the results on Q-completeness for the DINA model when attributes have a hierarchical structure do not immediately apply to general DCMs as well. The complex parameterization of general DCMs causes anomalies in Q-completeness that cannot be accounted for at present. As an illustration, consider a relatively simple instance of a general DCM, de la Torre's (2011) Additive Cognitive Diagnosis Model (A-CDM). The IRF of the A-CDM is defined as a linear combination of $K$ attributes

$$P(Y_j = 1 \mid \boldsymbol{\alpha}) = \beta_{j0} + \sum_{k=1}^{K} \beta_{jk} q_{jk} \alpha_{ik} \, ,$$

where $q_{jk}$ indicates whether mastery of attribute $\alpha_k$ is required for item $j$. (Additional constraints on the coefficients $\beta_{jk}$—not described here—guarantee $0 \leq P(Y_j = 1 \mid \boldsymbol{\alpha}) \leq 1$.) Suppose a test involves $K = 5$ attributes that have a convergent hierarchy, as shown in the following tree graph:

$$\begin{tikzcd}
 & & \alpha_5 & & \\
 & \alpha_2 \arrow[ur] & & \alpha_4 \arrow[ul] & \\
\alpha_1 \arrow[ur] & & & & \alpha_3 \arrow[ul]
\end{tikzcd}$$

Theorem 1 states that for the DINA model the five basic vectors that can be derived by inspection from the tree graph form the unique complete Q-matrix:

$$\mathbf{Q} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$\mathbf{Q}$ is still complete for the A-CDM, as can be shown by computing the vectors of expected item responses $\mathbf{S}(\boldsymbol{\alpha})$ (for the A-CDM, the expected response to item $j$ is $S_j(\boldsymbol{\alpha}) = P(Y_j|\boldsymbol{\alpha}) = \beta_{j0} + \sum_{k=1}^{5} \beta_{jk}\alpha_k q_{jk}$). However, $\mathbf{Q}$ is no longer uniquely complete for the A-CDM, as the example of the two Q-matrices $\boldsymbol{Q}_{1:2}$ and $\boldsymbol{Q}_{3:5}$ shows

$$\mathbf{Q}_{1:2} \;=\; \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad \mathbf{Q}_{3:5} \;=\; \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

The vectors of expected item responses for the A-CDM corresponding to $\mathbf{Q}_{1:2}$ and $\mathbf{Q}_{3:5}$ are:

| | $\mathbf{Q}_{1:2}$ | |
|---|---|---|
| $\boldsymbol{\alpha}$ | $\mathbf{q}_1 = (11000)$ | $\mathbf{q}_2 = (11111)$ |
| | $S_1(\boldsymbol{\alpha})$ | $S_2(\boldsymbol{\alpha})$ |
| (00000) | $\beta_{10}$ | $\beta_{20}$ |
| (10000) | $\beta_{10} + \beta_{11}$ | $\beta_{20} + \beta_{21}$ |
| (11000) | $\beta_{10} + \beta_{11} + \beta_{12}$ | $\beta_{20} + \beta_{21} + \beta_{22}$ |
| (00100) | $\beta_{10}$ | $\beta_{20} + \beta_{23}$ |
| (10100) | $\beta_{10} + \beta_{11}$ | $\beta_{20} + \beta_{21} + \beta_{23}$ |
| (11100) | $\beta_{10} + \beta_{11} + \beta_{12}$ | $\beta_{20} + \beta_{21} + \beta_{22} + \beta_{23}$ |
| (00110) | $\beta_{10}$ | $\beta_{20} + \beta_{23} + \beta_{24}$ |
| (10110) | $\beta_{10} + \beta_{11}$ | $\beta_{20} + \beta_{21} + \beta_{23} + \beta_{24}$ |
| (11110) | $\beta_{10} + \beta_{11} + \beta_{12}$ | $\beta_{20} + \beta_{21} + \beta_{22} + \beta_{23} + \beta_{24}$ |
| (11111) | $\beta_{10} + \beta_{11} + \beta_{12}$ | $\beta_{20} + \beta_{21} + \beta_{22} + \beta_{23} + \beta_{24} + \beta_{25}$ |

| | $\mathbf{Q}_{3:5}$ | | |
|---|---|---|---|
| $\boldsymbol{\alpha}$ | $\mathbf{q}_3 = (00100)$ | $\mathbf{q}_4 = (11100)$ | $\mathbf{q}_5 = (11111)$ |
| | $S_3(\boldsymbol{\alpha})$ | $S_4(\boldsymbol{\alpha})$ | $S_5(\boldsymbol{\alpha})$ |
| (00000) | $\beta_{30}$ | $\beta_{40}$ | $\beta_{50}$ |
| (10000) | $\beta_{30}$ | $\beta_{40} + \beta_{41}$ | $\beta_{50} + \beta_{51}$ |
| (11000) | $\beta_{30}$ | $\beta_{40} + \beta_{41} + \beta_{42}$ | $\beta_{50} + \beta_{51} + \beta_{52}$ |
| (00100) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{43}$ | $\beta_{50} + \beta_{53}$ |
| (10100) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{41} + \beta_{43}$ | $\beta_{50} + \beta_{51} + \beta_{53}$ |
| (11100) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{41} + \beta_{42} + \beta_{43}$ | $\beta_{50} + \beta_{51} + \beta_{52} + \beta_{53}$ |
| (00110) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{43}$ | $\beta_{50} + \beta_{53} + \beta_{54}$ |
| (10110) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{41} + \beta_{43}$ | $\beta_{50} + \beta_{51} + \beta_{53} + \beta_{54}$ |
| (11110) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{41} + \beta_{42} + \beta_{43}$ | $\beta_{50} + \beta_{51} + \beta_{52} + \beta_{53} + \beta_{54}$ |
| (11111) | $\beta_{30} + \beta_{33}$ | $\beta_{40} + \beta_{41} + \beta_{42} + \beta_{43}$ | $\beta_{50} + \beta_{51} + \beta_{52} + \beta_{53} + \beta_{54} + \beta_{55}$ |

Remarkably, $\mathbf{Q}_{1:2}$ is complete although it contains only two items. Two items, however, do not automatically guarantee completeness for the A-CDM, as the columns $\mathbf{q}_4 = (11100)$ and $\mathbf{q}_5 = (11111)$ of $\mathbf{Q}_{3:5}$ demonstrate. Completeness depends on the specific item parameter values and is not guaranteed because, for example, $\beta_{41} = \beta_{43}$ and $\beta_{51} = \beta_{53}$ cannot be ruled out so that $\boldsymbol{S}_{4:5}(10000) = \boldsymbol{S}_{4:5}(00100)$ is possible. Thus, a Q-matrix formed just by $\mathbf{q}_4$ and $\mathbf{q}_5$ cannot be guaranteed to be complete. The inconclusiveness of $\mathbf{q}_4$ and $\mathbf{q}_5$ can be resolved by adding item $\mathbf{q}_3 = (00100)$—indeed, $\mathbf{Q}_{3:5}$ is complete. (Note that the alternative additions $(00110)$ and $(11111)$, instead of $\mathbf{q}_3$, would also make $\mathbf{Q}_{4:5}$ guaranteed complete.) In summary, this small-scale example of the A-CDM shows that for a given attribute hierarchy there are at least three different ways to construct a complete Q-matrix for a general DCM. At present, it seems not possible to identify common rules of Q-completeness for general DCMs that can be used with any attribute hierarchy.

The second question to be addressed concerns an interesting theoretical coincidence. Within a graph-theoretic perspective, the hierarchical structure among attributes corresponds to an acyclic directed graph (digraph) (Harary, Norman, and Cartwright, 1965; Hubert, 1974). Remarkably, the unique $K \times K$ submatrix that guarantees Q-completeness for the DINA model and that can be formed by the basic attribute vectors, $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K$, is equal to the transpose of the reachability matrix of the digraph representation of the given attribute hierarchy. (Recall that $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K$ can be derived by inspection from the tree graph of the given attribute hierarchy.)

The $K \times K$ reachability matrix $\mathbf{R}$ of a digraph has entries $r_{kk'}$ defined as $r_{kk'} = 1$ if $\alpha_{k'}$ is reachable from $\alpha_k$, and $r_{kk'} = 0$ otherwise. ("Reachable" is to be taken literally: there exists a path leading from $\alpha_k$ to $\alpha_{k'}$; hence, $\alpha_{k'}$ can be "reached" from $\alpha_k$.) By definition, each attribute is reachable from itself. Hence, the diagonal entries of the reachability matrix $\mathbf{R}$ are all equal to one. The reachability matrix of a digraph can be obtained from its adjacency matrix $\mathbf{A}$, which is defined as a $K \times K$ binary matrix, with entries $a_{kk'} = 1$ if $\alpha_k$ and $\alpha_{k'}$ are connected by a single directed edge and $a_{kk'} = 0$ otherwise. Let the matrix $\mathbf{A}^2$ denote the matrix $\mathbf{AA}$ obtained by Boolean arithmetic; similarly, $\mathbf{A}^n$ can be obtained through Boolean arithmetic: $\mathbf{AA} \ldots \mathbf{A}$. The entry $a_{ij}^n$ of $\mathbf{A}^n$ is 1 if there is at least one sequence of directed edges of length $n$ from $\alpha_k$ to $\alpha_{k'}$ in the digraph (Theorem 5.5, p. 120; Harary et al., 1965). Thus, the reachability matrix $\mathbf{R}$ of the digraph of a given hierarchy involving $K$ attributes can be obtained as

$$\mathbf{R} = (\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^n) = (\mathbf{I} + \mathbf{A})^n$$

where $n \leq K-1$ (Theorem 5.7, Corollary 5.7a, p. 122; Harary et al., 1965). As an illustration, consider again Example 4, where

$$\mathbf{I} + \mathbf{A} \quad = \quad \begin{array}{c} \begin{array}{ccccccccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} & \alpha_{11} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Hence, $\mathbf{R} = (\mathbf{I} + \mathbf{A})^n$, with $n = 5$ (note that $n$ is determined if $(\mathbf{I} + \mathbf{A})^n - (\mathbf{I} + \mathbf{A})^{n+1} = \mathbf{0}$):

$$\mathbf{R} \quad = \quad \begin{array}{c} \begin{array}{ccccccccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} & \alpha_{11} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\Rightarrow \quad \mathbf{R}^T \quad = \quad \begin{array}{c} \begin{array}{ccccccccccc} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_5 & \alpha_6 & \alpha_7 & \alpha_8 & \alpha_9 & \alpha_{10} & \alpha_{11} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix} \end{array}$$

Indeed, the transpose of $\mathbf{R}$ is $\mathbf{Q} = (\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_K)$, as it was obtained earlier by inspection from the tree graph of the attribute hierarchy (see Section 6).

## References


BIRKHOFF, G. (1970), *Lattice Theory*, New York: Macmillan.

CHIU, C.-Y., and KÖHN, H.-F. (2015), "Consistency of Cluster Analysis for Cognitive Diagnosis: The DINO Model and the DINA Model Revisited", *Applied Psychological Measurement, 39*, 465–479.

CHIU, C.-Y., DOUGLAS, J.A., and LI, X. (2009), "Cluster Analysis for Cognitive Diagnosis: Theory and Applications", *Psychometrika, 74*, 633–665.

DE LA TORRE, J. (2011), "The Generalized Dina Model Framework", *Psychometrika, 76*, 179–199.

DE LA TORRE, J., and DOUGLAS, J.A. (2004), "Higher-Order Latent Trait Models for Cognitive Diagnosis", *Psychometrika, 69*, 333–353.

DIBELLO, L.V., ROUSSOS, L.A., and STOUT, W.F. (2007), "Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models", in *Handbook of Statistics: Vol. 26. Psychometrics*, eds. C.R. Rao and S. Sinharay, Amsterdam: Elsevier, pp. 979–1030.

FU, J., and LI, Y. (2007), "An Integrative Review of Cognitively Diagnostic Psychometric Models", paper presented at the April Annual Meeting of the National Council on Measurement in Education, Chicago, IL.

HABERMAN, S.J., and VON DAVIER, M. (2007), "Some Notes on Models for Cognitively Based Skill Diagnosis", in *Handbook of Statistics: Vol. 26. Psychometrics*, eds. C.R. Rao and S. Sinharay, Amsterdam: Elsevier, pp. 1031–1038.

HAERTEL, E.H. (1989), "Using Restricted Latent Class Models to Map the Skill Structure of Achievement Items", *Journal of Educational Measurement, 26*, 333–352.

HAERTEL, E.H., and WILEY, D.E. (1994), "Representation of Ability Structure: Implications for Testing", in *Testing Theory for a New Generation of Tests*, eds. N. Fredriksen, R. Mislevy, and I. Bejar, Hillsdale, NJ: Erlbaum, pp. 359–384.

HARARY, F., NORMAN, R.Z., and CARTWRIGHT, D. (1965), *Structural Models: An Introduction to the Theory of Directed Graphs*, New York: Wiley.

HEIJMANS, H.J.A.M. (1994), *Morphological Image Operators*, Boston, MA: Academic Press/Harcourt, Brace and Company, Publishers.

HENSON, R.A., TEMPLIN, J.L., and WILLSE, J.T. (2009), "Defining a Family of Cognitive Diagnosis Models Using Log-Linear Models with Latent Variables", *Psychometrika, 74*, 191–210.

HUBERT, L.J. (1974), "Some Applications of Graph Theory to Clustering", *Psychometrika, 39*, 283–309.

JUNKER, B.W., and SIJTSMA, K. (2001), "Cognitive Assessment Models with Few Assumptions, and Connections with Nonparametric Item Response Theory", *Applied Psychological Measurement, 25*, 258–272.

KÖHN, H.-F., and CHIU, C.-Y. (2017), "A Procedure for Assessing the Completeness of the Q-Matrices of Cognitively Diagnostic Tests", *Psychometrika, 82(1)*, 112–132.

LEIGHTON, J., and GIERL, M. (2007), *Cognitive Diagnostic Assessment for Education: Theory and Applications*, Cambridge, UK: Cambridge University Press.

LEIGHTON, J.P., GIERL, M.J., and HUNKA, S. (2004), "The Attribute Hierarchy Model: An Approach for Integrating Cognitive Theory with Assessment Practice", *Journal of Educational Measurement, 41*, 205–236.

MACREADY, G.B., and DAYTON, C.M. (1977), "The Use of Probabilistic Models in the Assessment of Mastery", *Journal of Educational Statistics, 33*, 379–416.

MARIS, E. (1999), "Estimating Multiple Classification Latent Class Models", *Psychometrika, 64*, 187–212.

RUPP, A.A., TEMPLIN, J.L., and HENSON, R.A. (2010), *Diagnostic Measurement. Theory, Methods, and Applications*, New York: Guilford.

RUPP, A.A., and TEMPLIN, J.L. (2008), "Effect of Q-matrix Misspecification on Parameter Estimates and Misclassification Rates in the DINA Model", *Educational and Psychological Measurement, 68*, 78–96.

TATSUOKA, K.K. (1985), "A Probabilistic Model for Diagnosing Misconception in the Pattern Classification Approach", *Journal of Educational and Behavioral Statistics, 12*, 55–73.

TATSUOKA, K.K. (2009), *Cognitive Assessment. An Introduction to the Rule Space Method*, New York: Routledge/Taylor and Francis.

TEMPLIN, J.L., and HENSON, R.A. (2006), "Measurement of Psychological Disorders Using Cognitive Diagnosis Models", *Psychological Methods, 11*, 287–305.

TEMPLIN, J., and BRADSHAW, L. (2014), "Hierarchical Diagnostic Classification Models: A Family of Models for Estimating and Testing Attribute Hierarchies", *Psychometrika, 79*, 317–339.

VON DAVIER, M. (2005), "A General Diagnostic Model Applied to Language Testing Data", Research Report No. RR-05-16, September, Princeton, NJ: Educational Testing Service.

VON DAVIER, M. (2008), "A General Diagnostic Model Applied to Language Testing Data", *British Journal of Mathematical and Statistical Psychology, 61*, 287–301.