Skill Set Analysis in Knowledge Structures*

Günther Gediga†

Institut für Evaluation und Marktanalysen

Brinkstr. 19

49143 Jeggen, Germany

gediga@eval-institut.de

Ivo Düntsch†

Dept of Computer Science

Brock University

St. Catherines, Ontario, Canada, L2S 3AI

duentsch@cosc.brocku.ca

Abstract

We extend the theory of knowledge structures by taking into account information about the skills a subject has. In the first part of the paper we exhibit some structural properties of the skill-problem relationship and consequences for the interpretation of concurrent theories in terms of the skill theory. The second part of the paper offers a test theory based on skill functions: We present measurements for the data consistency of the skill-problem relationship, and estimate abilities in terms of lower and/or upper boundaries of problem states and skills, given a special instance of the skill-problem relationship.

Some practical considerations are discussed, which enable the user of a skill based system to optimise a partial theory about the skill based behaviour of subjects based on empirical results.

*Co-operation for this paper was supported by EU COST Action 274 "Theory and Applications of Relational Structures as Knowledge Instruments" (TARSKI); http://www.tarski.org/

†Equal authorship is implied.

1 Introduction

The concept of skill assignments in knowledge spaces was introduced by Falmagne et al. (1990). In subsequent independent development, Korossy (1993), Doignon (1994) and, Düntsch and Gediga (1995) proposed various approaches to skills and knowledge structures; further work in the area can be found in the recent collection edited by Albert and Lukas (1999).

While the discussion of the dependency of cognitive abilities (skills) and observed knowledge states took place within the theoretical frame of knowledge spaces, application of the skill assignment theory has met with a limited response, although its practical usefulness as a test theory had been demonstrated by Düntsch and Gediga (1998). The present article aims at positioning the theory of skill assignment within a framework of related treatments of data analysis and data representation such as knowledge spaces or concept analysis; it offers applicable procedures to set up a practical test theory based on skill assignments through extending the theory by estimating the consistency between a theoretical model and observed data; it also includes a component of uncertainty handling.

[Figure 1 about here.]

Our basic model is pictured in Figure 1: First, the teacher chooses a domain of skills in which (s)he is interested. An empirical model in form of a set of test questions is constructed by an operationalisation which assigns to sets M of skills those problems, which can be solved with the skills in M. In a third step, representation (or numerical) models are induced by a scaling which can be used for assessment in various ways. This is an instance of the data model put forward by Gigerenzer (1981).

Just as knowledge assessment by modern scaling theory such as knowledge structures can be regarded as a qualitative way of measurement, we aim at a qualitative description of a subject's skill state. This philosophy is in the spirit of non–invasive data analysis (Düntsch and Gediga, 2000) which

- Minimises model assumptions, and
- Admits ignorance when no conclusion can be drawn.

As a consequence, we will concentrate on building the empirical model and the properties of the operationalisation; formally, our scaling will be an isomorphism. This is based on the conviction that a sensible diagnostic cannot be a point estimate of the knowledge of an individual such as a test score, but that, in most cases, only a range of skills can be given, which are more or less mastered by an individual.

The paper is organised as follows: We start with the definitions and notations which will be used in the sequel. This is followed by exhibiting some structural properties of various types of skill – problem operationalisations, which show a connection to mathematical structures such as modal logic and information systems, as well as to the theories of knowledge spaces (Doignon and Falmagne, 1999) and concept analysis (Wille, 1982). Several types of consistency between the theoretical and observed states are investigated in Section 4. The paper closes with an example which illustrates the concepts previously introduced.

2 Definitions and notation

2.1 Lattices and relations

First, let us briefly recall some set theoretic concepts from lattice theory and the theory of relations. The *symmetric difference* of two sets A, B is defined as

$$A \oplus B = (A \setminus B) \cup (B \setminus A).$$

In the finite case, the cardinality of $A \oplus B$ is sometimes called the *Hamming distance* of A and B, denoted by $H(A,B)$.

A *closure operator* on a set Q is a mapping $\mathrm{cl} : 2^Q \to 2^Q$ such that for all $X, Y \subseteq Q$,

$$X \subseteq Y \subseteq Q \Rightarrow \mathrm{cl}(X) \subseteq \mathrm{cl}(Y), \qquad \text{i.e. cl is monotone,} \tag{1}$$

$$X \subseteq \mathrm{cl}(X), \qquad \text{i.e. cl is expansive,} \tag{2}$$

$$\mathrm{cl}(X) = \mathrm{cl}(\mathrm{cl}(X)) \qquad \text{i.e. cl is idempotent.} \tag{3}$$

A *closure system* is a family of subsets of Q which is closed under intersection. It is well known that there are one-one correspondences between closure operators, closure systems, and $\cap$ - congruences, the latter, if Q is finite.

Dually, an *interior operator* is a mapping $\mathrm{int} : 2^Q \to 2^Q$ such that for all $X, Y \subseteq Q$,

$$X \subseteq Y \subseteq Q \Rightarrow \mathrm{int}(X) \subseteq \mathrm{int}(Y), \qquad \text{i.e. int is monotone,}$$

$$\mathrm{int}(X) \subseteq X, \qquad \text{i.e. int is contracting,}$$

$$\mathrm{int}(X) = \mathrm{int}(\mathrm{int}(X)), \qquad \text{i.e. int is idempotent.}$$

An *interior system* is a family of subsets of Q which is closed under union. Depending on the context, these are also known as *dependence spaces* (Novotný, 1997) and *knowledge spaces* (Doignon and Falmagne, 1985).

Let $\mathcal{K}$ be a family of subsets Q, and $X \subseteq Q$. We call X *join-irreducible in* $\mathcal{K}$, if X is not the union of proper subsets which are in $\mathcal{K}$. If $\mathcal{K}$ is understood, we call X just *join-irreducible.* Dually, we define *meet-irreducible* elements in $\mathcal{K}$. If $\mathcal{K}$ is a closure system, each element of $\mathcal{K}$ is an intersection of meet-irreducible elements. Dually, if $\mathcal{K}$ is an interior system, each element of $\mathcal{K}$ is a union of join-irreducible elements.

If $R \subseteq A \times B$ is a binary relation between elements of the sets A and B we will usually write xRy for $\langle x,y\rangle \in R$. The *relational converse of* R, denoted by $R^{\smile}$ is the relation

$$\{\langle y,x\rangle : xRy\}.$$

The *range of* x *in* R is the set $\{y \in B : xRy\}$, and we denote it by $R(x)$. The *range of* R is

$$\operatorname{ran} R = \{y \in B : (\exists x \in A) xRy\} = \bigcup_{x \in A} R(x).$$

With each $R \subseteq A \times B$ we associate several mappings from $2^A \to 2^B$: If $X \subseteq A$, then we set

$$[R](X) = \{b \in B : (\forall a \in A)[aRb \Rightarrow a \in X]\}, \qquad \text{(Necessity operator)}$$

$$\langle R\rangle(X) = \{b \in B : (\exists a \in A)[a \in X \text{ and } aRb]\}, \qquad \text{(Possibility operator)}$$

$$[[R]](X) = \{b \in B : (\forall a \in A))[a \in X \Rightarrow aRb]\}, \qquad \text{(Sufficiency operator)}$$

$$\langle\langle R\rangle\rangle(X) = \{b \in B : (\exists a \in A)[a \notin X \text{ and } \text{ not } aRb]\}. \qquad \text{(Dual sufficiency operator)}$$

These mappings arise in modal logics (Fitting, 1993) and information algebras (Düntsch and Orłowska, 2000, Orłowska, 1995). As an example, consider a relation $R \subseteq A \times B$, where A is a set of students, B is a set of problems, and aRb is interpreted as “Student a solves problem b”. If $X \subseteq A$ is a set of students, then for a problem b we have

$$b \in [R](X) \iff \text{Each student who solves } b \text{ is in } X$$

$$b \in \langle R \rangle(X) \iff \text{Some student in } X \text{ solves } b,$$

$$b \in [[R]](X) \iff b \text{ is solved by each student in } X$$

$$b \in \langle\langle R \rangle\rangle(X) \iff \text{Not all students in } -X \text{ solve } b$$

It may also be worthy to point out that $[[R]]$ is the derivation operator used in formal concept lattices.

Clearly,

$$b \in [R](X) \iff R\breve{}(b) \subseteq X,$$

$$b \in \langle R \rangle(X) \iff R\breve{}(b) \cap X \neq \emptyset,$$

$$b \in [[R]](X) \iff (-R)\breve{}(b) \subseteq -X,$$

$$b \in \langle\langle R \rangle\rangle(X) \iff (-R)\breve{}(b) \cap -X \neq \emptyset.$$

If $f : 2^A \to 2^B$ is a mapping, then its dual f^∂ is defined by

$$f^\partial(X) = -f(-X).$$

It is easy to see that $[R]$ and $\langle R \rangle$, as well as $[[R]]$ and $\langle\langle R \rangle\rangle$, are dual to each other.

2.2 Skills and knowledge structures

The qualitative analysis of knowledge via observable knowledge states was introduced by Doignon and Falmagne (1985), and an up to date account can be found in Doignon and Falmagne (1999).

Throughout this paper, we suppose that Q is a finite nonempty set of problems, and that U is a finite nonempty set of subjects. We also let $\mathrm{solv}_Q : U \to 2^Q$ be a function which assigns to a subject t the set of problems $\mathrm{solv}_Q(t)$ which (s)he can solve. Each $\mathrm{solv}_Q(t)$ is an *empirical knowledge state*, and the structure $\langle U, Q, \mathrm{solv}\rangle$ is called an *empirical knowledge structure* (EKS). If Q is understood, we will just write $\mathrm{solv}(t)$, and call $\{\mathrm{solv}(t) : t \in U\}$ an EKS, denoted by $\mathcal{K}_{\mathrm{solv}_Q}$, or just by $\mathcal{K}_{\mathrm{solv}}$.

The set Q of problems can be considered (in the sense of (Gigerenzer, 1981)) as an empirical model of those skills whose presence (or absence) in a student we want to ascertain, and we need to relate the problems to sets of skills with which the problems can be solved; in other words, we have to provide an operationalisation of the domain of interest. It may be worthy of mention that the operationalisation is a first source of uncertainty, since it is not always clear whether the empirical model truly reflects the properties and relations of the domain. In many instances, this will be the case; nevertheless, one needs to distinguish carefully between the set of problems a student is able to solve and the skills (s)he possesses in the area of interest.

Throughout the paper, we let S be a finite set of skills in a student's mastery of which we are interested. We assume that every problem in Q needs one or more nonempty sets of skills of S to be solved; this condition can be achieved by a simple pre-processing procedure.

Skills and problems can be related in the following way: Let Γ be a relation between Q and $2^S \setminus \{\emptyset\}$, i.e. $\Gamma \subseteq Q \times 2^S \setminus \{\emptyset\}$; note that $\Gamma(q)$ is either empty or a family of nonempty subsets of S. We call Γ a *skill relation* if $\Gamma(q) \neq \emptyset$ for each $q \in Q$, and the elements of $\Gamma(q)$ are pairwise incomparable with respect to $\subseteq$. We interpret $q\Gamma X$ as

X is minimal with respect to $\subseteq$ for the collection of all skill sets sufficient to solve q.

In other words, $q\Gamma X$, if q can be solved with the skills in X, but no proper subset of X is sufficient to solve q; we call each $X \in \Gamma(q)$ a *strategy for q*. At this stage, we do not put any (theoretical) restrictions on the strategies for q; in practice, Γ might be obtained by questioning an expert about the skills necessary and sufficient to solve q. Observe that for $X \subseteq S$, $\{q \in Q : (\exists Y \subseteq X) q\Gamma Y\}$ is the set of exactly those problems which can be solved with the skills in X. This induces a mapping $\delta : 2^S \to 2^Q$ defined by

$$X \overset{\delta}{\mapsto} \{q \in Q : (\exists Y \subseteq X) q\Gamma Y\}. \tag{4}$$

In particular,

$$(\forall X \in \Gamma(q)) q \in \delta(X). \tag{5}$$

More generally, a *problem function* is a mapping $\delta : 2^S \to 2^Q$ such that

1. δ is normal, i.e. $\delta(\emptyset) = \emptyset$.
2. δ is monotone.

The concepts of skill relation and problem function are equivalent (Düntsch and Gediga, 1995). Indeed, given a problem function $\delta : 2^S \to 2^Q$, we obtain Γ by

$$q\Gamma X \Longleftrightarrow X \text{ is minimal with respect to } q \in \delta(X).$$

We call the triple $\langle S, Q, \delta \rangle$ a *skill knowledge structure* (SKS). If Q is understood, we will also call the set $\{\delta(X) : X \subseteq S\}$ an SKS and denote it by $\mathcal{K}_\delta$. This set can be interpreted as the set of knowledge states which should be observable given the operationalisation δ. It was shown by Doignon (1994) and Düntsch and Gediga (1995) that for any $\mathcal{K} \subseteq 2^Q$ there are a set S and a problem function $\delta : 2^S \to 2^Q$

with $\operatorname{ran}\delta = \mathcal{K}$.

3 Structural properties

In defining a problem function, we have not required that it preserves $\cap$ or $\cup$. Indeed, these properties correspond to very special skill assignments as the following result shows:

Theorem 3.1. *1. δ preserves $\cap$ if and only if $|\Gamma(q)| = 1$ for all $q \in Q$.*

2. δ preserves $\cup$ if and only if $|T| = 1$ for all $T \in \Gamma(q)$ and all $q \in Q$.

Proof. **1.** "$\Rightarrow$": Assume that $X, Y \in \Gamma(q)$ for some $q \in Q$, $X \neq Y$; then, $q \in \delta(X) \cap \delta(Y)$ by (5), and thus, $q \in \delta(X \cap Y)$. By (4) there is some $Z \in \Gamma(q)$ such that $Z \subseteq X \cap Y$. Since $X \neq Y$, we $Z \subsetneq X$ or $Z \subsetneq Y$. This contradicts the fact that the elements of $\Gamma(q)$ are pairwise incomparable.

"$\Leftarrow$": Let $\Gamma(q) = \{X_q\}$ for each $q \in Q$. It is well known that δ preserves $\cap$ if and only if $\delta^{-1}(P)$ has a smallest element for each $P \in \operatorname{ran}(\delta)$ (see e.g. Novotný, 1997). Let $P \in \operatorname{ran}(\delta)$ and $Z = \bigcap\{Y : \delta(Y) = P\}$; we need to show that $\delta(Z) = P$. Since $Z \subseteq Y$ for each $Y \in \delta^{-1}(P)$ and δ is monotone, we have $\delta(Z) \subseteq P$. Conversely, let $p \in P$; then, $X_p \subseteq Y$ for all $Y \in \delta^{-1}(P)$, and hence, $X_p \subseteq Z$. It follows that $p \in \delta(Z)$.

2. "$\Rightarrow$": Assume that $q\Gamma X$ and $|X| \geq 2$; then, there are $Y, Z \subsetneq X$ with $Y \cup Z = X$. Now, $q \in \delta(Y \cup Z)$, but $q \notin \delta(Z) \cup \delta(Y)$.

"$\Leftarrow$": We show that $\delta^{-1}(P)$ has a greatest element for each $P \in \operatorname{ran}(\delta)$. Thus, let $P = \delta(Y)$ for some Y, and set $Z = \bigcup \delta^{-1}(P)$. Since $X \subseteq Z$ for each $X \in \delta^{-1}(P)$, and δ is monotone, we have $P \subseteq \delta(Z)$. Conversely, let $q \in \delta(Z)$. By (4) there is some $W \subseteq Z$ such that $q\Gamma W$. By our hypothesis, we can suppose that $W = \{t\}$, and by the definition of Z there is some $X \in \delta^{-1}(P)$ such that $t \in X$. Hence, $q \in \delta(X) \subseteq P$. $\square$

In the sequel, a skill function which preserves $\cap$ will be called *conjunctive*, and one which preserves $\cup$ will be called *disjunctive.*

As a Corollary of Theorem 3.1 we obtain

Corollary 3.2. *(Doignon, 1994)*

1. *If δ preserves $\cap$, then $\mathcal{K}_\delta$ is a closure system.*

2. *If δ preserves $\cup$, then $\mathcal{K}_\delta$ is a knowledge space.*

The converses are not true: Suppose that $Q = \{p,q\}$, $\Gamma = \{\langle p,S\rangle, \langle q,Y\rangle, \langle q,Z\rangle\}$, where Y,Z are incomparable with each having more than one element, and $S = Y \cup Z$. It is straightforward to see that

$$\mathcal{K}_\delta = \{\emptyset, \{q\}, Q\},$$

and δ preserves neither $\cap$ nor $\cup$.

The question arises whether a representation as in Theorem 3.1 can always be achieved for a closure system, respectively, a knowledge space. A positive answer was given by Doignon and Falmagne (1999); below we give somewhat different constructions.

Theorem 3.3. 1. *If $\mathcal{K}$ is a closure system, then there exists a skill representation with $|\Gamma(q)| = 1$ for all $q \in Q$.*

2. *If $\mathcal{K}$ is a knowledge space, then there a exists a skill representation with $|T| = 1$ for all $T \in \Gamma(q)$ and all $q \in Q$.*

Proof. **1.** Since $\mathcal{K}$ is finite and closed under intersection, it is generated by its meet-irreducible elements $M(\mathcal{K})$. For each nonempty $M_i \in M(\mathcal{K})$ choose a skill s_i, and let S be the collection of all these skills. Now we set

$$q\Gamma X \Longleftrightarrow X = \{s_i : q \notin M_i\}.$$

If $X \subseteq S$, then

$$\begin{aligned} q \in \delta(X) &\Longleftrightarrow q\Gamma X, \\ &\Longleftrightarrow \{s_i : q \notin M_i\} \subseteq X, \\ &\Longleftrightarrow (\forall s_i)[q \notin M_i \Rightarrow s_i \in X], \\ &\Longleftrightarrow (\forall s_i)[s_i \notin X \Rightarrow q \in M_i], \\ &\Longleftrightarrow q \in \bigcap\{M_i : s_i \notin X\}. \end{aligned}$$

Since each nonempty state of $\mathcal{K}$ is the intersection of meet irreducible elements, we have $\mathcal{K} = \mathcal{K}_\delta$.

2. Since $\mathcal{K}$ is finite and closed under union, it is generated by its join-irreducible elements $J(\mathcal{K})$. Suppose that the nonempty elements of $J(\mathcal{K})$ are $K_1, \ldots, K_n$, and choose a set $S = \{s_1, \ldots, s_n\}$. For each $q \in Q$ we let

$$q\Gamma X \iff (\exists s_i) X = \{s_i\} \text{ and } q \in K_i.$$

Suppose that $X = \{s_{i_0}, s_{i_1}, \ldots, s_{i_k}\}$. Then,

$$\begin{aligned} q \in \delta(X) &\iff q\Gamma X \\ &\iff (\exists s_{i_j})[s_{i_j} \in X \text{ and } \langle q, \{s_{i_j}\}\rangle \in \Gamma_{\min}], \\ &\iff (\exists s_{i_j})[s_{i_j} \in X \text{ and } q \in K_{i_j}], \\ &\iff q \in K_{i_1} \cup \ldots \cup K_{i_k}. \end{aligned}$$

Since the nonempty states of $\mathcal{K}$ are exactly the unions of K_is, we have $\mathcal{K} = \mathcal{K}_\delta$. □

Let us look at these "extreme" cases more closely: Suppose that $\Gamma \subseteq Q \times S$, and set $\Delta = \Gamma^{\smile}$. We can interpret $q\Gamma s$ in two different ways:

It is possible to solve q with skill s.

Skill s is necessary to solve q, and $\Gamma(q)$ is minimally sufficient to solve q.

These are, respectively, the disjunctive and conjunctive skill assignments of (Doignon, 1994). It will turn out that the knowledge structures arising from these two interpretations can be neatly described by the modal possibility and necessity operators. We will denote the corresponding problem functions by δ_d and δ_c, and the resulting knowledge structures by $\mathcal{K}_d$ and $\mathcal{K}_c$. Then, for all $X \subseteq S$,

$$\begin{aligned} \delta_c(X) &= \{q \in Q : \Gamma(q) \subseteq X\} &&= \{q \in Q : (\forall s \in S)[s\Delta q \Rightarrow s \in X]\} &&= [\Delta](X), \\ \delta_d(X) &= \{q \in Q : \Gamma(q) \cap X \neq \emptyset\} &&= \{q \in Q : (\exists s \in S)[s\Delta q \text{ and } s \in X]\} &&= \langle\Delta\rangle(X). \end{aligned}$$

If $K \subseteq Q$, we write $-K$ for $Q \setminus K$. We now have

Theorem 3.4. *1.* $[\Delta]\langle\Gamma\rangle$ *is a normal closure operator on* 2^Q*, and for each* $P \subseteq Q$, $[\Delta]\langle\Gamma\rangle(P)$ *is the* *smallest element of* $\mathcal{K}_c$ *containing* P.

2. $\langle\Delta\rangle[\Gamma]$ *is a normal interior operator on* 2^Q*, and for each* $P \subseteq Q$*,* $\langle\Delta\rangle[\Gamma](P)$ *is the largest element of* $\mathcal{K}_d$ *contained in* P.

3. $\mathcal{K}_d = \{-K : K \in \mathcal{K}_c\}$.

Proof. 1. Since both $\langle\Gamma\rangle$ and $[\Delta]$ are normal and monotone, so is $[\Delta]\langle\Gamma\rangle$. Let $q \in P$; then,

$$\begin{aligned} q \in [\Delta]\langle\Gamma\rangle(P) &\iff \Gamma(q) \subseteq \langle\Gamma\rangle(P), \\ &\iff (\forall s \in S)[q\Gamma s \Rightarrow (\exists p \in P) p\Gamma s]. \end{aligned}$$

Since $p \in P$, we may set $q = p$. For (3) we have

$$\begin{aligned} q \in [\Delta]\langle\Gamma\rangle[\Delta](X) &\Rightarrow (\forall s \in S)[q\Gamma s \Rightarrow s \in \langle\Gamma\rangle[\Delta](X), \\ &\Rightarrow (\forall s \in S)[q\Gamma s \Rightarrow (\exists p \in Q)[p\Gamma s \wedge (\forall t \in S)[p\Gamma t \Rightarrow t \in X]]], \\ &\Rightarrow (\forall s \in S)[q\Gamma s \Rightarrow s \in X], \\ &\Rightarrow q \in [\Delta](X). \end{aligned}$$

Suppose that $P \subseteq [\Delta](X)$ for some $X \subseteq S$. Then,

$$[\Delta]\langle\Gamma\rangle(P) \subseteq [\Delta]\langle\Gamma\rangle[\Delta](X) = [\Delta](X).$$

2. and 3. follow immediately from the duality of $\langle\Delta\rangle$ and $[\Delta]$, respectively, $\langle\Gamma\rangle$ and $[\Gamma]$. □

It is well known that for $R \subseteq A \times B$ the function $[[R^{\smile}]][[R]]$ is a closure operator as well, and it is not hard to see that it is just the $''$ operator used in formal context analysis. The following result, (whose simple – if somewhat tedious – proof is left to the reader) shows how the various closures differ:

Theorem 3.5. *For all $q \in Q$, $P \subseteq Q$,*

$$q \in [\Delta]\langle\Gamma\rangle(P) \Longleftrightarrow \bigcap\{\Gamma(p) : p \in P\} \subseteq \Gamma(q).$$

$$q \in [[\Delta]][[\Gamma]](P) \Longleftrightarrow \Gamma(q) \subseteq \bigcup\{\Gamma(p) : p \in P\}.$$

In the following sections we will investigate the question how well an SKS $\mathcal{K}_\delta$ approximates an EKS $\mathcal{K}_{\text{solv}}$. Theorem 3.4 shows that in case of δ_c, each $P \in \mathcal{K}_{\text{solv}}$ is contained in a smallest $P' \in \mathcal{K}_c$, which we will denote by $\overline{P}^c$. For a lower bound, we take in this case

$$\underline{P}_c = \begin{cases} P, \text{if } P \in \mathcal{K}_c, \\ \bigcap\{T \in \mathcal{K}_c : T \text{ is maximal with } T \subseteq P\}. \end{cases}$$

Observe that $\underline{P}_c \in \mathcal{K}_c$, since the latter is closed under intersection. Similarly, for δ_d, we set $\underline{P}_d = \langle\Delta\rangle[\Gamma](P)$, and

$$\overline{P}^d = \begin{cases} P, \text{if } P \in \mathcal{K}_d, \\ \bigcup\{T \in \mathcal{K}_d : T \text{ is minimal with } T \supseteq P\}. \end{cases}$$

We now have, not surprisingly,

Theorem 3.6. *For all $P \subseteq Q$,*

$$\underline{P}_c = -\overline{(-P)}^d, \; \overline{P}^c = -\underline{(-P)}_d.$$

Proof. First, note that by Theorem 3.4.3, we have for all $P, K \subseteq Q$

$$\begin{aligned}&K \text{ is maximal with respect to } K \in \mathcal{K}_c \text{ and } K \subseteq P \iff \\ &-K \text{ is minimal with respect to } K \in \mathcal{K}_d \text{ and } K \supseteq -P.\end{aligned}$$

If $P \in \mathcal{K}_c$, then $-P \in \mathcal{K}_d$, and the conclusion follows. Otherwise,

$$\begin{aligned}\underline{P}_c &= \bigcap\{K : K \text{ is maximal with } K \in \mathcal{K}_c \text{ and } K \subseteq P\},\\ &= \bigcap\{[\Delta]\langle\Gamma\rangle(K) : K \text{ is maximal with } K \in \mathcal{K}_c \text{ and } K \subseteq P\},\\ &= \bigcap\{-\langle\Delta\rangle(-\langle\Gamma\rangle)(K) : K \text{ is maximal with } K \in \mathcal{K}_c \text{ and } K \subseteq P\},\\ &= \bigcap\{-\langle\Delta\rangle([\Gamma])(-K) : -K \text{ is minimal with } -K \in \mathcal{K}_d \text{ and } K \supseteq -P\},\\ &= -\bigcup\{\langle\Delta\rangle([\Gamma])(-K) :: -K \text{ is minimal with } -K \in \mathcal{K}_d \text{ and } K \supseteq -P\},\\ &= -\overline{(-P)}^d.\end{aligned}$$

The second case is shown analogously. □

If $P = \mathrm{solv}(t)$, then, in case of a conjunctive δ, we interpret $\langle\Gamma\rangle(P)$ as the upper bound of the skills which t possibly has, and $\overline{P}^c$ as an upper bound of the problems which t is capable of solving. Note that this model assumption restricts the occurrence of careless errors, and does not touch upon lucky guesses.

With regard to the choice of knowledge spaces v. closure systems, the preceding duality results reinforce the view of Doignon (1994) that

> "There is no formal reason to prefer one kind of stability over the other when working with skill assignments. Only a careful study of the learning situation can give a hint to

> what should be the most relevant assumptions on the family of states."

Indeed, each of knowledge space and closure system covers just one side of a coin: While knowledge spaces seem to give more information about the skills a subject certainly has, closure systems delineate the skills a subject possibly has (and thus, which skills are not possessed).

It may be worth to look at knowledge structures, which are closure systems as well as knowledge spaces. A special example are linear orders of clusters of problems, an example of which will be considered in Section 5.1: Suppose that $\{A_i : i \leq k\}$ is a partition of Q induced by the equivalence relation $\sim$, and that Q is quasi-ordered by

$$p \prec q \iff p \in A_i,\ q \in A_j \text{ for some } i \lneq j \leq k.$$

We interpret the ordering as

> If q is solved, every problem $r \prec q$ must be solved as well, and any problem $p \sim q$ may or may not be solved.

This leads to the following knowledge structure: For each $i \leq k$, let $B_i = \bigcup_{j \leq i} A_j$. We also set $A_{k+1} = \emptyset$. Now,

$$\mathcal{K} = \{B_i \cup T : i \leq k, T \subseteq A_{i+1}\} \cup \{\emptyset\}.$$

$\mathcal{K}$ is $\cap$ and $\cup$ stable, and a conjunctive and a disjunctive skill assignment can be constructed in the following way: For each $i \leq k$, let $A_i = \{q_{i,j} : j \leq n_i\}$, so that $Q = \{q_{i,j} : i \leq k,\ j \leq n_i\}$. Our set of skills is $S = \{s_{i,j} : i \leq k, j \leq n_i\}$; observe that the assignment $s_{i,j} \mapsto q_{i,j}$ is a bijection from S to Q.

- For all $i \leq k$, $j \leq n_i$ set

$$\Gamma_0(q_{i,j}) = \{\{s_{r,t} : r \lneq i,\ t \leq n_r\} \cup \{s_{i,j}\}\}.$$

It is straightforward to see that Γ_0 leads to a conjunctive skill assignment, and that the resulting SKS is just $\mathcal{K}$.

- For all $i \leq k$, $j \leq n_i$ set

$$\Gamma_1(q_{i,j}) = \{\{s_{r,t}\} : i \lneq r,\ t \leq n_r\} \cup \{\{s_{i,j}\}\}.$$

Then, the skill assignment is disjunctive, and $\delta_1(\{s_{i,j}\}) = \bigcup_{t \lneq i} A_i \cup \{q_{i,j}\}$ shows that $\mathcal{K}_{\delta_1} = \mathcal{K}$.

Therefore, incompatible theoretical ideas may lead to the same empirical results.

4 Towards a test theory based on skill functions

Once we have fixed our operationalisation δ of a domain of skills, we need to gauge how well the empirical data and the theoretical structure interact. For global test consistency, we need to compare the EKS $\mathcal{K}_{\text{solv}}$ and the SKS $\mathcal{K}_{\delta}$, and our plan is to find indices which provide a closer analysis of the compatibility of the observed states and the predicted states.

In the literature there are several – quite diverging – approaches to construct estimation procedures to measure the fit between $\mathcal{K}_{\text{solv}}$ and the SKS $\mathcal{K}_{\delta}$. Classical latent traits modelling approaches such as the conjunctive latent task model (using a conjunctive assignment of tasks to problems, (Jannarone, 1986, 1991, 1997)) and the disjunctive latent task model (Lord, 1984) show a remarkably different behaviour; this is somewhat astonishing, because there are situations in which both modelling approaches should come to approximately the same results as our construction of Γ_0 and Γ_1 above. In this Section we describe a test theory based on skill functions, which does not use the strict model assumptions of a Birnbaum or Rasch model such as tracelines of task behaviour or, additionally, weighted sums of cognitive demands as in cognitive design systems (Embretson, 1998, 1999). Instead, we use indirect checks of model consistency. Furthermore, the proposed test theory does not – at least, in principle – rely on the special cases of conjunctive or disjunctive functions.

Within the frame of knowledge space theory, probabilistic assumptions for modelling errors in $\mathcal{K}_{solv}$ have been considered as well in what may be called a *probabilistic knowledge structure* (PKS) approach. This technique uses a probability α for lucky guesses and a probability β for careless errors to describe the discrepancies of $\mathcal{K}_{solv}$ and $\mathcal{K}_{\delta}$ (Kambouri et al., 1994). There are, however, several problems: First of all, the PKS approach is not specified unless the parameter restrictions are given. If no restrictions are set, one needs to specify α- and β-parameters for all state$\times$problem combinations, which is by far too much to be estimated using a reasonable number of subjects. Restricting the parameters is possible, but results in models whose properties are largely unknown. Some approaches assign all lucky – guess – parameters to **one** parameter α, and all careless – error – parameters to **one** parameter β. This is a very restrictive probabilistic knowledge structure, and it is not clear in which real life situations it can be fruitfully employed. The second origin of problems stems from the fact that the PKS does not reflect the nature of the skill function: We have shown that certain knowledge structures can be constructed from conjunctive or disjunctive skill functions, but the PKS is identical for both – a situation which is not satisfactory, if one assumes the probability of an error when solving a problems is a function of the “missing skill” probabilities. Given a conjunctive skill function, the ”problem solved” probability will be mixture of convolutions of elementary probability distributions, whereas the disjunctive skill function results in a complicated mixture of elementary components.

Therefore, one can say with some justification that there is – up to now – no simple and elegant way to set up a probabilistic version of the deterministic model of skill assignment. In the next two subsections we investigate the problem in a very elementary way, which can be described by the term "bump hunting", ie. looking for unevenness in the texture. The main assumption is that α- and β-parameters are small, and that the states of the deterministic model are a substantial part of $\mathcal{K}_{\text{solv}}$. A statistical test procedure to check this assumption is provided.

4.1 Global consistency

In the sequel, we call any observed state $\text{solv}(t) \in \mathcal{K}_{\text{solv}} \setminus K_\delta$ *inconsistent* (with respect to $\langle S, Q, \delta \rangle$).

The *consistency index*

$$\gamma = \frac{|\mathcal{K}_{\text{solv}} \cap \mathcal{K}_\delta|}{|\mathcal{K}_{\text{solv}}|}$$

offers a first look at the relative consistency of $\mathcal{K}_{\text{solv}}$ and $\mathcal{K}_\delta$, since $0 \leq \gamma \leq 1$ and

$$\gamma = 1 \Longleftrightarrow \mathcal{K}_{\text{solv}} \subseteq \mathcal{K}_\delta.$$

A statistic to measure consistency of the EKS and the SKS which takes into account the weighting of the states by subjects is

$$\gamma(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_\delta) = \frac{|\bigcup_{K \in \mathcal{K}_{\text{solv}} \cap \mathcal{K}_\delta} \text{solv}^{-1}(K)|}{|\bigcup_{K \in \mathcal{K}_{\text{solv}}} \text{solv}^{-1}(K)|} = \frac{|\{t \in U : \text{solv}(t) \in \mathcal{K}_\delta\}|}{|U|},$$

If $\mathcal{K}_{\text{solv}}, \mathcal{K}_\delta$ are understood we will usually just write γ_U.

If γ_U is not far from 1, it seems obvious that the empirical structure is more or less captured by the skill function. However, even $\gamma_U = 1$ is not a guarantee that the skill function has any significance for the data: Suppose that $f : S \to Q$ is a bijection, and set $\delta(X) = \{f(x) : x \in X\}$. In this case, the resulting $\mathcal{K}_\delta$ is simply 2^Q and any state is consistent with such skill assignment.

In terms of a PKS we count the relative number of elements in states with the highest probabilities. But, for large $|U|$, the value of γ_U is expected to be very small, because even small deviations of EKS and SKS due to lucky guesses and careless errors will result in remarkable number of misfits. For both reasons, there is the need for a statistical procedure to derive the significance of the intersection of EKS and SKS. This computation needs some knowledge about a plausible benchmark distribution of the fit of $\mathcal{K}_\delta$ with respect to $\mathcal{K}_{\text{solv}}$. Following the road of non-invasive data analysis (Düntsch and Gediga, 2000), we want to minimise model assumptions; therefore, in the absence of other information, we suggest to employ a randomisation technique since it is valid for any kind of sample (Manly, 1997). We assume the relation Γ to be given, and we randomise the problems within the relation. Let Σ be the set of all permutations of Q. For each $\sigma \in \Sigma$ we define Γ_σ by

$$\sigma(q)\Gamma_\sigma X \iff q\Gamma X.$$

Since σ is a bijection, Γ_σ is well defined, and we let $\mathcal{K}_{\delta(\sigma)}$ be the resulting skill knowledge structure. Note that σ is just a re-labelling of Q, and thus, the set theoretic structure of $\mathcal{K}_\delta$ is not changed. We can now compute the position of the empirical value $\gamma(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_\delta)$ in the distribution

$$\{\gamma(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_{\delta(\sigma)}) : \sigma \in \Sigma_Q\},$$

which enables us to measure the significance of the null hypothesis H_0

"$\gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta)$ is the result of a random assignment of problems to skill sets"

by

$$p = \frac{|\{\gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_{\delta(\sigma)}) \geq \gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta)\}|}{|Q|!}.$$

and to compare this to a fixed α value. Since the computation of p is not feasible for larger $|Q|$, we have developed a sequential randomisation procedure which dramatically reduces the computational effort (Düntsch and Gediga, 2000, Chapter 4.3). We have found, that in most cases less than 100 simulations were required. Note, that the randomisation procedure is a conditional test and that even a small value of γ_U may have a large effect, for example, if the number of items is large and the number of states is small. Therefore, in order to achieve an effect measure, γ_U has to be adjusted. Using the arguments of Cohen (1960) and Scott (1955), who took into account that agreement may be due to chance, we result in a *test consistency measure* which is a correction of γ_U for consistency expectation:

$$\kappa(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta) = \frac{\gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta) - \mathcal{E}[\gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta)]}{1 - \mathcal{E}[\gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta)]}.$$

The expectation $\mu(\gamma_U|H_0)$ of a random assignment of problems to skill sets is defined by

$$\mu(\gamma_U|H_0) = \frac{\sum_{\sigma \in \Sigma} \gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_{\delta(\sigma)})}{|Q|!} = \mathcal{E}_{\sigma \in \Sigma_Q}[\gamma(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_{\delta(\sigma)})], \tag{6}$$

and can be approximated by simulation methods as well.

$\kappa(U, \mathcal{K}_{\mathrm{solv}}, \mathcal{K}_\delta) = 1$ indicates that $\mathcal{K}_{\mathrm{solv}} \subseteq \mathcal{K}_\delta$ holds. If κ is near 0, an overlap of $\mathcal{K}_{\mathrm{solv}}$ and $\mathcal{K}_\delta$ may be due to chance.

The randomisation procedure above can also be used for the statistical evaluation of additional structure: Suppose that $\delta_1, \delta_2 : 2^S \to 2^Q$ are operationalisations of S, and suppose that $\mathcal{K}_{\delta_1} \subseteq \mathcal{K}_{\delta_2}$. We then compute $\mathcal{K}^*_{\text{solv}} = \mathcal{K}_{\text{solv}} \setminus \mathcal{K}_{\delta_1}$ and $\mathcal{K}^*_{\delta_2} = \mathcal{K}_{\delta_2} \setminus \mathcal{K}_{\delta_1}$ and $\gamma(U, \mathcal{K}^*_{\text{solv}}, \mathcal{K}^*_{\delta_2})$. Its position in the distribution $\gamma(U, \mathcal{K}^*_{\text{solv}}, \mathcal{K}_{\delta^*_2(\sigma)})$ gives us some insight whether the empirical observations in the additional states can be attributed to a random process or not.

4.2 Item consistency

Given a sufficiently homogeneous context, item analysis can be done by checking the change of the test consistency when removing an item q from the problem set. More formally, we set

$$\mathcal{K}^q_\delta = \{\delta(X) \setminus \{q\} : X \subseteq S\},$$

$$\mathcal{K}^q_{\text{solv}} = \{\text{solv}(t) \setminus \{q\} : t \in U\},$$

$$\gamma^q(U, \mathcal{K}^q_{\text{solv}}, \mathcal{K}^q_\delta) = \frac{|\{t \in U : \text{solv}^q(t) \in \mathcal{K}^q_\delta\}|}{|U|},$$

Observing that $\gamma^q(U, \mathcal{K}^q_{\text{solv}}, \mathcal{K}^q_\delta) \geq \gamma(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_\delta)$, we result in a procedure which helps to evaluate the influence of a test item: If the test consistency changes considerably when an item is removed, then this is a strong advice to remove the item from the test or to reformulate the given theory.

A comparable construction can be done for the evaluation of skills: Removing a set T of skills from S

will result in

$$\delta^T = \delta \restriction (S \setminus T), \tag{7}$$

$$\mathcal{K}_{\delta^T} = \{\delta^T(X) : X \subseteq S \setminus T\}, \tag{8}$$

$$\gamma^T(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_{\delta^T}) = \frac{|\{t \in U : \text{solv}(t) \in \mathcal{K}_{\delta^T}\}|}{|U|}, \tag{9}$$

Observing that $\gamma^T(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_{\delta^T}) \leq \gamma(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_{\delta})$, we result in a procedure which helps to find redundant skills in terms of consistency. A *skill reduct* is a set T of skills such that

$$\gamma^{(S\setminus T)}(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_{\delta^{(S\setminus T)}}) = \gamma(U, \mathcal{K}_{\text{solv}}, \mathcal{K}_{\delta}).$$

Any skill reduct has the same expressive power – in terms of consistency – as the full skill set.

4.3 Distance to boundaries

Consistency measures deal with the exact match of theory and data. However, there are situations which cannot be tackled by such a crisp zero-or-one-statistic. As a simple tool, the theory offers the usage of the lower and/or upper boundaries for the evaluation of the skill theory in case of suitable problem functions. As an example, suppose that the given problem function δ is conjunctive or disjunctive, so that $\overline{P}$ and $\underline{P}$ are meaningful. For every element $P \in \mathcal{K}_{\text{solv}}$ we compute the Hamming distances

$$\overline{H}(P) = |\overline{P} \setminus P|, \tag{10}$$

$$\underline{H}(P) = |P \setminus \underline{P}|. \tag{11}$$

Thus, every element of the EKS can be measured in terms of its distances to upper and lower approximation. A conjunctive skill function will result in a closure system, and a unique and simple upper bound of any element of the EKS can be computed. The analysis of the Hamming distance to the upper bound offers a simple and computationally feasible method for the evaluation of such closure system. Similarly, the lower bound of any element of the EKS using a disjunctive skill function (which generates a knowledge space) is unique and easy to compute as well, thus, providing a simple an computationally feasible method for the evaluation of such knowledge spaces.

The distribution of Hamming distances can be used for descriptive and inferential purposes. Their mean value provides a measure which is analogous to the coefficient of reproducibility (REP) for Guttman-scales by the definition

$$REP_{H^*} = 1 - \frac{\sum_{P \in \mathcal{K}_{\text{solv}}} H^*(P)}{|\mathcal{K}_{\text{solv}}||Q|},$$

letting $H^*(P)$ be either $\underline{H}(P)$ or $\overline{H}(P)$.

We think that – even in case of a Guttman scale – the analysis of Hamming distances to upper and lower state boundaries is more informative than an index such as reproducibility:

- The distances to lower and upper bound are applied to every empirical state, and they offer a characterisation of this state in terms of its boundaries.
- The concept of Hamming distances to upper and lower state boundaries is very general, and can be applied to far more situations than the Guttman-scale-based reproducibility approach. It should be noted that the reproducibility approach and the Hamming distance concept are not measuring the same thing: The reproducibility counts the number of element exchanges that have to be performed to result in a errorless model. This approach has not much in common

with the Hamming distance concept, because if the“hardest” item is solved alone (which has the largest Hamming distance to an upper boundary of a Guttman scale), the exchange of two elements would repair the error. If the “easiest” item is not solved, but all others are, we result in the largest Hamming distance to a lower boundary, but once again the error can be repaired by exchanging two elements.

Because the subject set U can be quasi-ordered by the Hamming distance $\overline{H(\mathrm{solv}(t))}$ and/or $\underline{H(\mathrm{solv}(t))}$, the distribution of the Hamming distances can be used to compare two or more skill functions. Two skill functions can be compared by using a Wilcoxon rank test, and for more than two skill functions one can use a Friedman analysis of ranks.

4.4 Comparing groups and explaining group differences

Comparing groups using test results has become a prominent topic in applications. There are – at least – two different questions when comparing groups using knowledge structures: First, the groups can be described by two different structures, and the knowledge structure is used for descriptive purposes (e.g. Janssens (1999)). Second, the task is to analyse which group is “better”, given the results within a representation of a common knowledge structure for all groups. For the second task, we have shown elsewhere (Düntsch and Gediga, 1998) that the results in $\mathcal{K}_{\mathrm{solv}}$ alone are sufficient to test whether one group outperforms another one. Here, we will present a slightly generalised version of this test procedure.

Suppose that the subject set U is partitioned in to groups $\mathfrak{G} = \{G_i : 1 \leq i \leq t_G\}$, and that for all

$G,H \in \mathfrak{G}$ we define $\mu_{G,H} : G \times H \to \{0,1\}$ by

$$\mu_{G,H}(x,y) = \begin{cases} 1, & \text{if } \mathrm{solv}(y) \subsetneq \mathrm{solv}(x), \\ 0, & \text{otherwise.} \end{cases}$$

The *group difference function* $D : \mathfrak{G} \times \mathfrak{G} \to \mathbb{N}$ is defined by

$$D(G,H) = \sum \{\mu_{G,H}(x,y) : \langle x,y \rangle \in G \times H\}.$$

The difference $D(G,H) - D(H,G)$ offers a possibility for the evaluation of the group differences based on $\mathcal{K}_{\mathrm{solv}}$. The *normalised difference* $r(G,H)$ is defined as

$$r(G,H) = \begin{cases} 0, & \text{if } D(G,H) + D(H,G) = 0, \\ \frac{D(G,H)-D(H,G)}{D(G,H)+D(H,G)}, & \text{otherwise.} \end{cases}$$

The normalised difference is bounded by $-1 \leq r(G,H) \leq 1$. If $r(G,H) = 0$ then there is no group difference, whereas $r(G,H) = 1$ indicates that group G dominates H in the highest possible way. The significance of the normalised difference $D(G,H)$ can be tested by a randomisation approach using the $\binom{|G|+|H|}{|G|}$ possible assignments σ of subjects to groups G^σ and H^σ, defining

$$D^\sigma(G,H) = \sum \{\mu_{G,H}(x,y) : \langle x,y \rangle \in G^\sigma \times H^\sigma\}.$$

The distribution $r^\sigma(G,H)$ can be generated as a basis of a one-sided significance test, and the distribution $|r^\sigma(G,H)|$ can be used as well for the statistical evaluation of $|r(G,H)|$.

The evaluation of the differences of more than two groups based on the randomisation approach is possible as well using

$$R = \frac{\sum_{G \neq H} |D(G,H) - D(H,G)|}{2 \cdot \sum_{G \neq H} D(G,H)}.$$

Observe that neither additional scaling assumptions nor distributional assumptions are being used for the construction of these tests, and that they are non-invasive as defined above. One might object, however, that the tests will not find group differences of the following kind:

- Group G: All subjects solve items A, B, C.
- Group H: All subjects solve item D.

The normalised difference $r(G,H)$ is zero, although the subjects of group G solve three problems, whereas all subjects in group H only solve one problem. This seems to be a drawback, but it is well within our non-invasive approach: The difference between both groups can only be stated on the basis of the (additional) scaling assumption

"Sets of solved problems can be scaled to numbers"

with all its implications. The proposed test procedure only uses the soft scaling assumption that the subset relation can be interpreted as $\leq$; in other words, subject x is regarded to be better than subject y if and only if $\mathrm{solv}(y) \subsetneq \mathrm{solv}(x)$.

There are several strategies to compare groups on the basis of the scaling assumptions. The general idea is to replace the observed $\mathrm{solv}(x)$ by a (in some sense) compatible state $s^*(x)$ and perform the statistical procedure with the new states. We demonstrate these strategies using a simple example: Consider

$$\mathcal{K} = \{\emptyset, \{A\}, \{A,B\}, \{A,B,C\}, \{A,B,C,D\}\} \text{ and } \mathrm{solv}(x) = \{B\}.$$

- The traditional "scale sum" approach is to replace the state $\mathrm{solv}(x)$ by the "nearest" state $s^*(x)$, which shows an identical scaling property, namely, the sum of the solved items. This is a property of a numerical system, which represents the data – and not a property of the empirical system. Obviously, $\mathrm{solv}(x)$ and $s^*(x)$ can be quite different, e.g. $s^*(x) = \{A\}$.

- Because the Hamming distance measure has several nice properties (Suck, 1999), the idea of taking a state $K \in \mathcal{K}_\delta$ which has a minimum Hamming distance to $\mathrm{solv}(x)$ is tempting. But even in the simple example the result of this procedure need not be unique, because

  $$H(\mathrm{solv}(x), \emptyset) = H(\mathrm{solv}(x), \{A, B\}) = 1.$$

  Therefore, such an approach is not useful for the purpose of group comparisons.

- In case of a conjunctive or disjunctive skill assignment, the upper bound approach using $s^*(x) = \overline{\mathrm{solv}(x)}$ will result in a unique representation of any element of $\mathcal{K}_{\mathrm{solv}}$. It offers an optimistic estimation for the set of problems a subject can solve, assuming that the problems in $s^*(x) \setminus \mathrm{solv}(x)$ have not been solved due to careless errors. In the example we observe $s^*(x) = \{A, B\}$.

  The lower bound approach using $s^*(x) = \underline{\mathrm{solv}(x)}$ will result in a unique representation of any element of $\mathcal{K}_{\mathrm{solv}}$ as well. Here, one assumes that the problems in $\mathrm{solv}(x) \setminus s^*(x)$ have been solved by lucky guesses. In the example we observe $s^*(x) = \emptyset$.

Because the upper and lower bound approaches offer an approximation which is

- Compatible with the theory,

- Located in the empirical system,

- Unique,

we favour the use of these strategies for skill theory based group comparison, if the operationalisation is conjunctive or disjunctive.

Using upper or lower bound alone may run in problems, as the following example shows:

$$\mathcal{K} = \{\emptyset, \{A\}, \{A,B\}, \{A,B,C\}, \{A,B,C,D\}\}, \text{ solv}(g) = \{D\}, \text{ solv}(h) = \{A,B,C\}.$$

We have learned from testing in K_{solv} that g and h will not differ. With the scale sum approach, we see that $g \precnsim h$, which is the same result which the lower bound approach offers. The upper bound approach results in $h \precnsim g$, because $\overline{\text{solv}(h)} = \{A,B,C\} \subsetneq \overline{\text{solv}(g)} = \{A,B,C,D\}$. This dissociation looks like a drawback, but it really is not, because the diverging results simply reflect the fact that the states given by the theory do not fit the data.

4.5 Skill based evaluation of groups

When using a conjunctively or disjunctively interpreted skill relation $\Gamma \subseteq Q \times 2^S \setminus \{\emptyset\}$, a skill based evaluation of groups is fairly simple, the reason being that the range of the skills of subject x can be retrieved from the upper and lower bounds of $\text{solv}(x)$ as described in Theorem 3.4 and the discussion afterwards. To show what can be done, we first define $R(s) = \{q \in Q : (\exists X)[s \in X \text{ and } q\Gamma X]\}$ for each skill s; regarding $|R(s)|$ as an indicator variable, we can count the resulting numbers within the groups. Assuming that every skill is tested by at least one problem, we are able to test group differences for skill s using a Chi-square test.

The skill set based approach of comparing groups using upper and lower bounds sometimes appears to be unsatisfactory. The reason for this are the crisp definitions of upper and lower bound, which turn out to mask group differences, if the skill function is not properly defined. To find out which skill assignments “often” work well and which do not, we can do the following in case of a conjunctive skill assignment: For each $s \in S$, the ratio

$$I(x|s) = \frac{|\,\mathrm{solv}(x) \cap R(s)|}{|R(s)|}$$

is called the *skill intensity* of s for subject x. If $I(x|s) \neq 0$, then s is an upper bound skill for x, and if $I(x|s) = 1$, s is a lower bound skill for x. Group differences represented within a “partially working” skill assignment can be found by inspecting the 95% confidence intervals of the means of $I(x|s)$ in every group.

5 An Example: Reanalysis of a well known Guttman Scale

Even for small scale situations, non-trivial examples for skill assignment procedures tend to be technically demanding, because, as an a-priori technique, the domain theory must be developed in detail, and its connection to the data must be stated in a precise manner.

In order to present an example with low technical overhead for describing the skill assignment technique, we have chosen one of the first applications of Guttman’s scaling technique (Guttman, 1944, 1950), in which Suchman (1950) investigates physical reactions to dangers of battle, experienced by soldiers who have been under fire. Obviously, physical reactions to danger are not “solved problems” and there can be no “skills” as a theoretical basis for explaining this reaction. Nevertheless, we will show below that one can re-interpret the given domain-data-connection in terms of the skill theory developed above. In this context, a “problem” is a physical symptom, and a “skill” is the representation of a stimulus which triggers such a symptom.

5.1 Structure of the example

Suchman (1950) showed that the patterns of symptoms experienced by the subjects form a Guttman scale with a coefficient of reproducibility of 0.92. Presence of symptoms in decreasing order was as follows:

q_1	Violent pounding of the heart	84%
q_2	Sinking feeling of the stomach	73%
q_3	Feeling sick at the stomach	57%
q_4	Shaking or trembling all over	52%
q_5	Feeling of stiffness	50%
q_6	Feeling of weakness or feeling faint	42%
q_7	Vomiting	35%
q_8	Loosing control of the bowels	21%
q_9	Urinating in pants	9%

Once one has found a scaling model for the data, the scale must be interpreted in theoretical terms. A simple interpretation is that

- The representation of the external danger stimuli ("skills") meets reaction thresholds for reactions q_1 to q_9,
- Reaction thresholds show the same ordinal relation for every subject,
- The representation of external danger stimuli grows smoothly.

A theory consistent with these assumption leads to a conjunctive skill assignment with the following

Γ relation:

$$\Gamma_G(q_1) = \{\{A_1\}\},$$

$$\Gamma_G(q_2) = \{\{A_1, A_2\}\},$$

$$\Gamma_G(q_3) = \{\{A_1, A_2, A_3\}\},$$

$$\Gamma_G(q_4) = \{\{A_1, A_2, A_3, A_4\}\},$$

$$\Gamma_G(q_5) = \{\{A_1, A_2, A_3, A_4, A_5\}\},$$

$$\Gamma_G(q_6) = \{\{A_1, A_2, A_3, A_4, A_5, A_6\}\},$$

$$\Gamma_G(q_7) = \{\{A_1, A_2, A_3, A_4, A_5, A_6, A_7\}\},$$

$$\Gamma_G(q_8) = \{\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8\}\},$$

$$\Gamma_G(q_9) = \{\{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9\}\}.$$

A "skill set" $\{A_1, \ldots, A_k\}$ is interpreted as

- The representation of external danger stimuli exceeds the reaction threshold of $q_1, \ldots, q_k$.
- There are no other possibilities to generate q_k.

The Guttman scaling – and the corresponding (skill) function interpretation – is constructed by

1. The rank order of the percentages of solved items.
2. The additional assumption that the dimension which is observed in the numerical system has a counterpart in the empirical system.

The idea of the skill function approach is to use a domain knowledge based construction, which does not use (post-hoc) information in the numerical system . The interpretation frame of the Guttman scale has two theoretical assumptions, which can be adopted:

- The representation of external danger stimuli meets reaction thresholds for reactions q_1 to q_9.
- The representation of external danger stimuli grows smoothly.

A look at the list of items shows that they can be assigned to three categories

- Slight somatic symptoms q_1, q_2.
- Medium to severe symptoms without excretion q_3, q_4, q_5, q_6.
- Excretion q_7, q_9, q_8.

Assuming an order among these three categories, and no order within the categories due to individual responses to stimuli representation, we can construct a clustered partial order of three groups, resulting in the following conjunctive relation:

$$\Gamma_3(q_1) = \{\{B_1\}\},$$
$$\Gamma_3(q_2) = \{\{B_2\}\},$$
$$\Gamma_3(q_3) = \{\{B_1, B_2, B_3\}\},$$
$$\Gamma_3(q_4) = \{\{B_1, B_2, B_4\}\},$$
$$\Gamma_3(q_5) = \{\{B_1, B_2, B_5\}\},$$
$$\Gamma_3(q_6) = \{\{B_1, B_2, B_6\}\},$$
$$\Gamma_3(q_7) = \{\{B_1, B_2, B_3, B_4, B_5, B_6, B_7\}\},$$
$$\Gamma_3(q_8) = \{\{B_1, B_2, B_3, B_4, B_5, B_6, B_8\}\},$$
$$\Gamma_3(q_9) = \{\{B_1, B_2, B_3, B_4, B_5, B_6, B_9\}\}.$$

We should like to point out that at this stage in the model building process, the data have not yet entered into the picture.

Whereas $\mathcal{K}_G$ has 10 theoretical states, $\mathcal{K}_{\delta_3}$ has 29 states which have one of the forms

$$\begin{array}{rl} P, & P \subseteq \{q_1, q_2\}, \\ \{q_1, q_2\} \cup P, & P \subseteq \{q_3, q_4, q_5, q_6\}, \\ \{q_1, q_2, q_3, q_4, q_5, q_6\} \cup P, & P \subseteq \{q_7, q_8, q_9\}. \end{array}$$

5.2 Consistency

Comparing the theoretical states with the observed answers from a sample U with size 100 (Suchman, 1950, p. 140), we obtain Table 1.

[Table 1 about here.]

There, a "hit" is an element of $\mathcal{K}_{\text{solv}} \cap \mathcal{K}_{\delta}$, and $\mathcal{E}$ is the expectation defined on p. 21.

The number of hits in both theories is significantly different from its expectation. The question arises, whether Γ_3 is a substantially better theory than Γ_G. Since $\mathcal{K}_{\delta_G} \subseteq \mathcal{K}_{\delta_3}$, we can use the technique for analysing partial γ values; the results are shown in Table 2.

[Table 2 about here.]

The number of additional hits is significantly different from the expectation, and we can conclude that Γ_G is substantially improved by Γ_3 in terms of consistency.

The leave-one-statistics for both theories are presented in Table 3.

[Table 3 about here.]

Both skill theories have higher support, when leaving out certain problems. In case of the Guttman scale, q_3 is a promising candidate, whereas in case of the 3 cluster assumption, item q_7 is a good choice for an even better model fit. It is trivial that leaving out $A1$ will not change the consistency, because $A1$ is really redundant, as it is a conjunctive part of every problem function. Whereas the profile of the A-skills is rather flat, the B-skills show (up to skill B1) a rather big difference to the start value (70). This indicates that most of skills for Γ_3 are necessary to result in such high consistency.

[Table 4 about here.]

5.3 Lower and upper bounds and the Hamming distance distributions

Table 4 presents lower and upper bounds of the elements of $\mathcal{K}_{\text{solv}}$, and an evaluation in terms of the distribution of Hamming distances to the upper and lower bounds is given in Table 5. This provides additional information about the misfits of the problem function in terms of the distribution of Hamming distances to the upper and lower boundaries.

[Table 5 about here.]

5.4 Group comparison

The example data do not contain a variable with group information. To demonstrate our approach, we define two groups by assigning the subjects to group 0 if item 5 is not present ($q_5 = 0$), and to group 1 if item 5 is present ($q_5 = 1$); item q_5 is then removed from the scale. The knowledge structure comparison – which is model independent – is presented in Table 6.

[Table 6 about here.]

Within the comparable elements item q_5 splits the empirical knowledge structure quite perfectly (r=0.81); the differences can hardly be attributed to random processes (sig($r \neq E[r]$) = 0.01).

The group comparison in terms of theoretical variables is presented in Table 7 for the basic terms in the data model given by Γ_3.

[Table 7 about here.]

The two groups show large differences in all indices of the theoretical terms $B1$ to $B9$. Table 7 also demonstrates a simple structural property of the upper approximation of skills: If a skill is a very basic one (like B1 or B2), it is often present in the upper approximation. The diagnostic can be done by comparing the mean value of the intensity with the upper bound percentages: If both differ remarkably, the skill is often added to the upper bound. For the most complicated skills (like B7, B8, B9), the mean of the intensity and the upper bound percentages are identical, which means that these skills were never added to an upper bound.

6 Summary and outlook

The theory of skill knowledge structures (TSKS) proposes a direct link from the researcher's theory to a tailored numerical system, using observed data (= the empirical system) as an intermediate medium. Its foundation is an a-priori scaling model, and therefore, there is a need for a precise formulation of the theory in logical or relational terms. In the model, the researcher starts by constructing a skill assignment, resp. a problem function, which is an explicit operationalisation of domain knowledge, mapped to an empirical system. There are no additional scaling assumption in TSKS – all assumptions are restrictions of the operationalisation. Even the model assumption for the operationalisation is rather soft: “More” in the domain should result in “more” in the empirical system – which makes sense for testing knowledge. We have shown that the additional scaling assumptions of closure under union or intersection can be expressed by restrictions on the operationalisation. Furthermore, we have positioned well known data analysis strategies such as knowledge spaces and concept analysis within the TSKS context.

Since the numerical system is constructed from the initial theory without the data, these must be interpreted in terms of the theory and its model assumptions; in this respect, Guttman scaling as an a-posteriori model is less demanding. In our context, for example, one may well ask, whether the assumption of a monotone operationalisation makes sense in the example of under-fire-symptoms; our results show that such a requirement is certainly consistent with the data.

If an expert formulates a theory (s)he will often not be as precise as necessary to result in an acceptable model fit. In this paper some procedures are offered to discover weaknesses of the problem function, and to optimise the theory of the data. The significance tests which we have suggested are a descriptive instrument only; however, by using cross validation, one has a powerful instrument for test construction. We have used the procedures introduced in this paper to construct a new intelligence test and have shown that the results compete well with results of conventional models of test construction (Preckel et al., 2001).

Because a sound theory does not allow every possible outcome, results of subjects are not always precisely given in the TSKS model, and only lower and upper bound of knowledge can be generated. We think that these bounds offer more information than a single point estimate, because the latter compresses the knowledge using a – sometimes not suitable – measurement instrument. The lower and upper bounds are not statistical bounds, but logical ones. A statistical error theory for upper and lower bounds still needs to be developed.

Acknowledgement

We would like to thank the anonymous referees for their constructive criticism and pointers to further literature.

References

Albert, D. and Lukas, J. (1999). *Knowledge Spaces: Theories, Empirical Research, and Applications.* Lawrence Erlbaum, Mahwah, NY.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:37–46.

Doignon, J.-P. (1994). Knowledge spaces and skill assignments. In Fischer, G. and Laming, D., editors, *Contributions to Mathematical Psychology, Psychometrics, and Methodology*. Springer, Berlin.

Doignon, J.-P. and Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man–Machine Studies*, 23:283–303.

Doignon, J. P. and Falmagne, J. C. (1999). *Knowledge spaces.* Springer-Verlag, Berlin.

Düntsch, I. and Gediga, G. (1995). Skills and knowledge structures. *British J. Math. Statist. Psych.*, 48:9–27.

Düntsch, I. and Gediga, G. (1998). Knowledge structures and their applications in CALL systems. In

Jager, S., Nerbonne, J., and van Essen, A., editors, *Language teaching and language technology*, pages 177–186. Swets & Zeitlinger.

Düntsch, I. and Gediga, G. (2000). *Rough set data analysis: A road to non-invasive knowledge discovery*, volume 2 of *Methoδos Primers*. Methoδos Publishers (UK), Bangor.

Düntsch, I. and Orłowska, E. (2000). Beyond modalities: Sufficiency and mixed algebras. In Orłowska, E. and Szałas, A., editors, *Relational Methods in Computer Science Applications*, pages 277–299, Berlin. Springer–Verlag.

Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, 3:380–396.

Embretson, S. E. (1999). Generating items during testing: Psychometric issues and models. *Psychometrika*, 64:407–433.

Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P., and Johannesen, J. (1990). Introduction to knowledge spaces: How to build, test and search them. *Psychological Review*, 97:202–234.

Fitting, M. (1993). Basic modal logic. In Gabbay, D. M., Hogger, C. J., and Robinson, J. A., editors, *Logical foundations*, volume 1 of *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 368–448. Clarendon Press, Oxford.

Gigerenzer, G. (1981). *Messung und Modellbildung in der Psychologie*. Birkhäuser, Basel.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, 9:139–150.

Guttman, L. (1950). The basis for scalogram analysis. In Stouffer et al. (1950), pages 60–90.

Jannarone, R. J. (1986). Conjunctive item response theory kernels. *Psychometrika*, 51:357–373.

Jannarone, R. J. (1991). Cognitive measurement theory: Cognitive research prospects. In Wilson, M., editor, *Objective measurement theory: Theory into practice.* Ablex, Norwood, NJ.

Jannarone, R. J. (1997). Models of locally dependent responses: Conjunctive item response theory. In van der Linden, W. J. and Hambleton, R. K., editors, *Handbook of modern item response theory*, pages 465–480. Springer, New York.

Janssens, R. (1999). A boolean approach to the measurement of group processes and attitudes. the concept of integration as an example. *Mathematical Social Sciences*, 38:275–293.

Kambouri, M., Koppen, M., Villano, M., and Falmagne, J. C. (1994). Knowledge assessment: Tapping human expertise by the QUERY routine. *International Journal of Human Computer Studies*, 40:119–151.

Korossy, K. (1993). *Modellierung von Wissen als Kompetenz und Performanz. Eine Erweiterung der Wissensstruktur–Theorie von Doignon und Falmagne.* Phd thesis, Universität Heidelberg.

Lord, F. M. (1984). Conjunctive and disjunctive item response functions. Technical Report RR-84-45-ONR, Educational Testing Service, Princeton, NJ.

Manly, B. F. J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology*. Chapman and Hall, London.

Novotný, M. (1997). Dependence spaces of information systems. In Orłowska, E., editor, *Incomplete Information – Rough Set Analysis*, pages 193–246. Physica – Verlag, Heidelberg.

Orłowska, E. (1995). Information algebras. In *Proceedings of AMAST 95*, volume 639 of *Lecture Notes in Computer Science*. Springer–Verlag.

Preckel, F., Gediga, G., Düntsch, I., and Holling, H. (2001). Constructing intelligence tests by deterministic and statistical test theory: What can be learned from different approaches? http://www.roughian.co.uk/tt/preckel.html.

Scott, W. A. (1955). Reliability of content analysis: the case of nominal scaling. *Public Opinion Quarterly*, 19:321–325.

Stouffer, S., Guttman, L., Suchman, E., Lazarsfeld, P., Star, S., and Clausen, J., editors (1950). *Measurement and Prediction.* Princeton University Press.

Suchman, E. (1950). The utility of scalogram analysis. In Stouffer et al. (1950), pages 122–171.

Suck, R. (1999). A dimensional-related metric on the lattice of knowledge spaces. *J. Math. Psych.*, 43:394–409.

Wille, R. (1982). Restructuring lattice theory: an approach based on hierarchies of concepts. In Rival, I., editor, *Ordered sets*, volume 83 of *NATO Advanced Studies Institute*, pages 445–470. Reidel, Dordrecht.

Figure 1: The data model

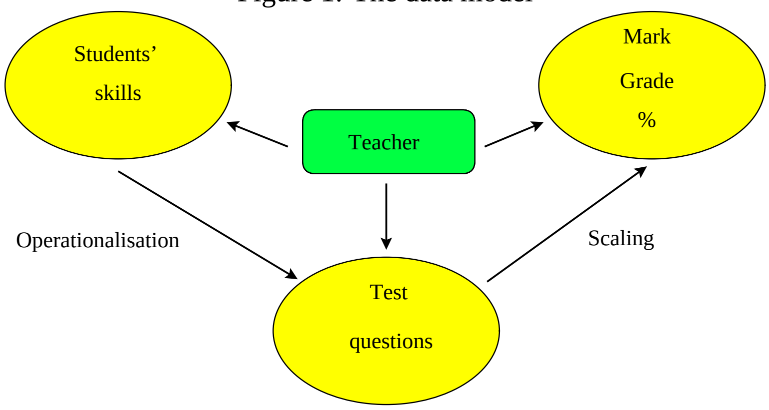

Table 1: Consistency of two skill theories

Theory	No. of hits	γ_U	$\mathcal{E}$[hits]	$p(\text{hits} \leq E[\text{hits}] \mid H_0)$	κ_U
Γ_G	51	0.51	16.28	≤ 0.001	0.415
Γ_3	70	0.70	20.65	≤ 0.001	0.622

Table 2: Comparison of consistency of two nested skill theories

No of additional hits	γ_U	$\mathcal{E}$[hits]	$p(\text{hits} \leq E[\text{hits}] \mid H_0)$	κ_U
19	0.39	1.55	≤ 0.001	0.613

Table 3: Leave-one-out statistic for soldier data

Item	Γ_G	Γ_3	Skill	Γ_G (A)	Γ_3 (B)
q_1	53	74	(A/B)1	51	69
q_2	54	74	(A/B)2	44	63
q_3	60	72	(A/B)3	44	54
q_4	58	73	(A/B)4	49	55
q_5	56	74	(A/B)5	49	50
q_6	58	72	(A/B)6	49	49
q_7	59	81	(A/B)7	45	61
q_8	53	71	(A/B)8	46	58
q_9	52	70	(A/B)9	44	51

Table 4: Lower and upper bounds in the soldier data

Pattern (X1.. X9)	Freq.	Γ_3 lower	Γ_3 upper	Γ_G lower	Γ_G upper
0 0 0 0 0 0 0 0 0	7	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
0 0 0 0 0 0 0 1 0	1	$\{q_2\}$	$\{q_2\}$	$\emptyset$	$\{q_1,q_2\}$
0 0 0 0 0 0 0 0 1	7	$\{q_1\}$	$\{q_1\}$	$\{q_1\}$	$\{q_1\}$
0 0 0 1 0 0 0 0 0	1	$\emptyset$	$\{q_1,q_2,q_6\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 0 1 0 0 0 0 0 0	1	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 0 0 1 0 1 0	1	$\{q_2\}$	$\{q_1,q_2,q_4\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4\}$
0 0 0 0 1 0 0 1 0	2	$\{q_2\}$	$\{q_1,q_2,q_5\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5\}$
0 0 0 0 0 0 0 1 1	7	$\{q_1,q_2\}$	$\{q_1,q_2\}$	$\{q_1,q_2\}$	$\{q_1,q_2\}$
0 1 0 0 0 0 0 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
0 0 1 0 0 0 0 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 0 1 0 0 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_5\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5\}$
0 0 0 0 0 0 1 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3\}$	$\{q_1\}$	$\{q_1,q_2,q_3\}$
0 0 0 0 0 1 0 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_4\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4\}$
0 0 0 1 0 0 1 0 0	1	$\emptyset$	$\{q_1,q_2,q_6\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 0 0 0 0 1 0 1 1	3	$\{q_1,q_2,q_4\}$	$\{q_1,q_2,q_4\}$	$\{q_1,q_2\}$	$\{q_1,q_2,q_3,q_4\}$
0 0 0 0 0 0 1 1 1	2	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3\}$
0 0 0 1 0 0 0 1 1	1	$\{q_1,q_2,q_6\}$	$\{q_1,q_2,q_6\}$	$\{q_1,q_2\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 0 1 0 1 0 1 1 0	1	$\{q_2\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 0 1 1 1 1 0	1	$\{q_2\}$	$\{q_1,q_2,q_3,q_4,q_5\}$	$\emptyset$	$\{q_1,q_2,q_3,q_4,q_5\}$
0 0 0 0 1 0 1 1 1	2	$\{q_1,q_2,q_3,q_5\}$	$\{q_1,q_2,q_3,q_5\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5\}$
0 0 0 1 0 0 1 1 1	3	$\{q_1,q_2,q_3,q_6\}$	$\{q_1,q_2,q_3,q_6\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 0 0 0 0 1 1 1 1	2	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4\}$
0 0 0 1 0 1 0 1 1	1	$\{q_1,q_2,q_3,q_6\}$	$\{q_1,q_2,q_3,q_6\}$	$\{q_1,q_2\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 0 0 0 1 1 0 1 1	3	$\{q_1,q_2,q_4,q_5\}$	$\{q_1,q_2,q_4,q_5\}$	$\{q_1,q_2\}$	$\{q_1,q_2,q_3,q_4,q_5\}$
0 0 1 0 0 0 1 1 1	1	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 1 0 1 0 0 1 1	1	$\{q_1,q_2,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 1 0 0 1 1 0 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
0 0 1 1 1 0 1 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
0 0 1 1 0 0 1 1 1	1	$\{q_1,q_2,q_3,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 1 0 1 1 1 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 1 1 1 0 1 1	1	$\{q_1,q_2,q_4,q_5,q_6\}$	$\{q_1,q_2,q_4,q_5,q_6\}$	$\{q_1,q_2\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 1 0 0 1 1 1 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
0 0 1 0 1 0 1 1 1	1	$\{q_1,q_2,q_3,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 0 1 1 1 1 1	2	$\{q_1,q_2,q_3,q_4,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5\}$
0 0 1 0 0 1 1 1 1	2	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 1 0 1 1 1 1	1	$\{q_1,q_2,q_3,q_4,q_6\}$	$\{q_1,q_2,q_3,q_4,q_6\}$	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
1 1 0 0 0 0 1 1 1	1	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8,q_9\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$
0 1 1 0 0 0 1 1 1	1	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
0 0 1 1 1 1 1 0 1	1	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 1 1 1 1 0 1 1	1	$\{q_1,q_2,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 1 1 0 1 1 1 1	1	$\{q_1,q_2,q_3,q_4,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3,q_4\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 1 0 1 1 1 1 1	1	$\{q_1,q_2,q_3,q_4,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3,q_4,q_5\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
0 0 0 1 1 1 1 1 1	6	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$
0 1 1 1 1 0 1 1 1	1	$\{q_1,q_2,q_3,q_5,q_6,q_7,q_8\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$	$\{q_1,q_2,q_3\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
0 0 1 1 1 1 1 1 1	5	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7\}$
1 0 0 1 1 1 1 1 1	1	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$
0 1 0 1 1 1 1 1 1	1	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$
1 1 0 1 1 1 1 1 1	1	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_8,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$
0 1 1 1 1 1 1 1 1	7	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$
1 1 1 1 1 1 1 1 1	6	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$	$\{q_1,q_2,q_3,q_4,q_5,q_6,q_7,q_8,q_9\}$

Table 5: Distribution (in %) of Hamming distances to boundaries in the soldier data

Upper Bound Statistics					Lower Bound Statistics				
$\overline{H}(x)$	Γ_G	cum	Γ_3	cum	$\underline{H}(x)$	Γ_G	cum	Γ_3	cum
0	51	51	70	70	0	51	51	70	70
1	20	71	12	82	1	24	75	19	89
2	13	84	9	91	2	15	90	4	93
3	9	93	6	97	3	2	92	3	96
4	3	96	0	97	4	7	99	3	99
5	2	98	2	99	5	1	100	1	100
6	2	100	1	100	6				
Mean	1.07		0.63			0.93		0.53	
REP_{H^*}	0.88		0.93			0.90		0.94	

Table 6: Comparison of two groups within an empirical knowledge structure

Total	"$q_5 = 0$" with group "$q_5 = 1$"						
	"0"$\succ$"1	"0"$\prec$"1"	r	$=$	non comp.	sig($r > 0$)	sig($r \neq E[r]$)
2500	181	1721	0.81	41	557	0.007	0.010

Table 7: Comparison of two groups based on theoretical terms given by Γ_3

	Upper bound comparison			Mean and 2-σ bounds of the mean					
Skill	"$q_5 = 0$"	"$q_5 = 1$"	Chi-Square$_1$	Intensity("$q_5 = 0$")			Intensity("$q_5 = 1$")		
B1	84.0%	96.0%	4.00	20.1	25.4	30.8	52.7	60.3	67.8
B2	72.0%	98.0%	13.25	17.0	22.9	28.7	52.3	59.7	67.1
B3	40.0%	86.0%	22.69	8.6	14.5	20.4	37.8	46.5	55.2
B4	38.0%	90.0%	29.34	7.2	12.0	16.8	38.2	46.5	54.8
B5	20.0%	64.0%	19.87	3.1	8.0	12.9	25.7	35.3	44.9
B6	36.0%	78.0%	17.99	6.4	11.0	15.6	33.1	42.5	51.9
B7	16.0%	54.0%	15.87	5.5	16.0	26.5	39.8	54.0	68.2
B8	6.0%	36.0%	13.56	0.0	6.0	12.8	22.3	36.0	49.7
B9	2.0%	16.0%	5.98	0.0	2.0	6.0	5.5	16.0	26.5