Skills and Knowledge Structures

Ivo Düntsch*
School of Information and Software Engineering
University of Ulster
Newtownabbey, BT 37 0QB, N.Ireland
I.Duentsch@ulst.ac.uk

Günther Gediga*
FB Psychologie / Methodenlehre
Universität Osnabrück
49069 Osnabrück, Germany
gg@Luce.Psycho.Uni-Osnabrueck.DE

Abstract

Suppose that Q is a set of problems and S is a set of skills. A *skill function* assigns to each problem q – i.e. to each element of Q – those sets of skills which are minimally sufficient to solve q; a *problem function* assigns to each set X of skills the set of problems which can be solved with these skills (a *knowledge state*). We explore the natural properties of such functions and show that these concepts are basically the same. Furthermore, we show that for every family K of subsets of Q which includes the empty set and Q, there are a set S of (abstract) skills and a problem function whose range is just K. We also give a bound for the number of skills needed to generate a specific set of knowledge states, and discuss various ways to supply a set of knowledge states with an underlying skill theory. Finally, a procedure is described to determine a skill function using coverings in partial orders which is applied to set A of the Coloured Progressive Matrices test Raven (1965).

Key words: knowledge structures, ability testing, dichotomous data, discrete modelling

1 Introduction and notation

Suppose we are given a set of problems Q. The more skills one has to solve the problems, the more problems one will solve. This simple idea is at the basis of all common theories of ability tests. Apart from this idea, the theories diverge in their assumptions about ”latent traits” or ”ability parameters” or other constructs that ”represent” the data as good as possible.

A more direct connection between theory and data is the theory of knowledge spaces described in Falmagne et al. (1990): A *knowledge structure* $\mathcal{K} = \langle Q, K \rangle$ consists of a nonempty finite set Q of problems, and a collection K of subsets of Q (the *knowledge states*) which includes $\emptyset$ and Q. A knowledge structure describes the sets of problems which subjects are capable of solving. If K is closed under union, it is called a *knowledge space*.

An even more direct connection between theory and data would be that the researcher formulates the skills a subject needs in order to be able to solve a problem in Q. Alternatively, given a set S of skills, an expert could indicate for each subset X of S which problems can be solved with the skills in X. With such a theoretical frame (we later on call a skill function, resp. a problem function) it is straightforward to

*The ordering of authors is alphabetical, and equal authorship is implied.

compute all possible knowledge states. In Falmagne et al. (1990), p.208, it is argued that there is a link between the theory of knowledge spaces and a skill based theory, and they ask for a description of this link. The simple skill function in Tab. 1 indicates that closure under union may be too strong a condition: The table implies that subjects can solve nothing, or problem 1, or problem 2, or all 3 problems. Clearly, the

Table 1: A simple skill function

Problem	Necessary skills
1	A
2	B
3	both A and B

resulting collection of subsets of Q is not a knowledge space, because it is not closed under union: There is no way a subject can solve exactly problems 1 and 2.

In this article we explore the structures that can be built using skill and problem functions, and how empirical data can be used to explore an underlying skill or problem function. In contrast to Falmagne et al. (1990), we shall not require that the family of knowledge states under consideration is closed under union. Thus, our structures are more general. The price to be paid is that we no longer have the nice structural properties of knowledge spaces: Being finite join semilattices, these are always generated by their join irreducible elements. Thus, the alternative descriptions given in Falmagne et al. (1990) - e.g. as surmise systems - are no longer possible.

Our notation is as follows: A reflexive and transitive relation R on a set M is a *quasiorder*. If such R is in addition antisymmetric, it is called a *partial order*, and the structure $\langle M, R\rangle$ a partially ordered set. A partial order is usually denoted by $\leq$. For $x, y \in M$, we set $[y, x] \stackrel{\text{def}}{=} \{z \in M : y \leq z \leq x\}$, and $(x] \stackrel{\text{def}}{=} \{z \in M : z \leq x\}$ If $P \subseteq M$, and $x \in M$, we write $P \leq x$ if $y \leq x$ for all $y \in P$. For $p, q \in M$, we say that *p covers q* if q is a minimal element of the set $\{r \in M : p < r\}$. More generally, we say that *q covers T* $\subseteq$ *M* if q is a minimal element of $\{p \in M : T \leq p\}$. A *chain* , resp. *antichain* is a set of pairwise comparable, resp. incomparable elements of M.

If $\langle L, \leq_L\rangle$ and $\langle M, \leq_M\rangle$ are partially ordered sets, then M is called a *retract* of L, if there are order preserving maps $f : L \to M$ and $g : M \to L$ such that f is onto, g is one - one, and $g \circ f : M \to M$ is the identity on M.

For a real number x, $\lceil x\rceil$ denotes the smallest integer above x, and, for a natural number m, $m \; div \; 2$ is the largest integer $\leq \frac{m}{2}$. The fact that a mapping is surjective is indicated by the $\twoheadrightarrow$ arrow.

For concepts and notation not defined here, in particular for lattices and related structures, the reader is invited to consult Grätzer (1978) or Birkhoff (1967).

2 Skill functions, problem functions and knowledge structures

Throughout, let Q be a fixed finite nonempty set of problems and S be a fixed finite nonempty set of skills.

A function $\gamma : Q \to 2^{2^S}$ is called a *skill function*, if $\gamma(q)$ is a nonempty set of nonempty, pairwise incomparable subsets of S for each $q \in Q$. This coincides with the definition given in Falmagne et al. (1990) and the skill multiassignments in Doignon (????) [1].

Table 2: Raven's skill theory of CPM set A

Problem	Skills
1	$\{\{B\}\}$
2	$\{\{B\}\}$
3	$\{\{B\}\}$
4	$\{\{B\}\}$
5	$\{\{B\}\}$
6	$\{\{B, O\}\}$
7	$\{\{B, G\}\}$
8	$\{\{B, G\}\}$
9	$\{\{B, 1, O\}\}$
10	$\{\{B, 1, O\}\}$
11	$\{\{B, 1, 2, G\}\}$
12	$\{\{B, 1, 2, O, C\}\}$

B: Basic skills and simple, continous pattern completion
1: Basic skills and simple, continous pattern completion with change in 1 dimension
2: Basic skills and simple, continous pattern completion with change in 2 dimensions
O: Orientation of the missing part
G: Gestalt formation and completion
C: Correlate creation

Intuitively, the elements of $\gamma(q)$ – which are called *competencies* in Falmagne et al. (1990) – are exactly those sets of skills which are minimally sufficient to solve problem q; in other words, each $X \in \gamma(q)$ is a set of skills sufficient to solve q, while each proper subset of X is not. If $\gamma(q)$ contains only one element X, the researcher states that each strategy to solve problem q must contain the skills which are specified by X. If $\gamma(q)$ contains $p > 1$ subsets of S, there are p essentially different strategies to solve problem q. Tab. 2 shows an example of an explicit theory expressed in terms of a skill function (Raven (1965)).

The skill function connects the theory of the researcher about the assumptions of the problem solving process to observable data. Given a skill function γ, there are combinations of problems which should not be observed empirically; thus, γ puts constraints on the observable subsets of Q. For example, given the skill function of Raven's set A, we should not observe any proper nonempty subset of the solution set $A = \{1, 2, 3, 4, 5\}$; neither should the union of the observable sets $\{1, \ldots, 8\}$ and $\{1, \ldots, 5, 7, 8, 11\}$ be observable. Tab. 3 shows all possible valid combinations of solved problems if the skill theory in Tab. 2 holds. One can observe that problems $1, \ldots, 5$ are equivalent in the sense that they cannot be separated by a knowledge state. In Falmagne et al. (1990), an equivalence relation θ on Q is introduced which identifies $p, q \in Q$, if they are in the same states. The resulting classes are called *notions*. In Raven's skill theory, there are the notions

$$\{1, \ldots, 5\}, \{6\}, \{7, 8\}, \{9, 10\}, \{11\}, \{12\}.$$

[1] We should like to thank one of the referees for bringing Doignon's article to our attention.

Nothing is lost in respect to the structure of K if one considers the quotient set Q/θ. In Doignon & Falmagne (1985), notions are introduced in order to turn a certain quasiorder into a partial order (namely, the one defined just before 2.2) which somewhat simplifies matters when one considers knowledge spaces.

Table 3: Valid states of set A given Raven's skill theory

State	Problems
1	$\emptyset$
2	$\{1,2,3,4,5\}$
3	$\{1,2,3,4,5,6\}$
4	$\{1,2,3,4,5,7,8\}$
5	$\{1,2,3,4,5,6,7,8\}$
6	$\{1,2,3,4,5,6,9,10\}$
7	$\{1,2,3,4,5,6,7,8,9,10\}$
8	$\{1,2,3,4,5,7,8,11\}$
9	$\{1,2,3,4,5,6,7,8,9,10,11\}$
10	$\{1,2,3,4,5,6,9,10,12\}$
11	$\{1,2,3,4,5,6,7,8,9,10,11,12\}$

Following Falmagne et al. (1990), we call a subset A of Q a γ - *knowledge state*, if there is some $X \subseteq S$ such that A is exactly the set of problems for which a member of $\gamma(q)$ is a subset of X. This induces a function $G(\gamma) : 2^S \to 2^Q$ defined by

$$G(\gamma)(X) = \{q \in Q : \text{There is some } A \in \gamma(q) \text{ with } A \subseteq X\},$$

whose range are exactly the γ - knowledge states. $G(\gamma)$ calculates for each $X \subseteq S$ the set of problems which can be solved with the skills in X.

Lemma 2.1. *Let $\gamma : Q \to 2^{2^S}$ be a skill function. Then, the function $G(\gamma)$ as defined above is monotone with respect to $\subseteq$. Furthermore, $G(\gamma)(\emptyset) = \emptyset$ and $G(\gamma)(S) = Q$.*

Proof. Let $X \subseteq Y \subseteq S$, and $q \in G(\gamma)(X)$. Then, there is some $T \in \gamma(q)$ such that $T \subseteq X \subseteq Y$, and thus $q \in G(\gamma)(Y)$.

The rest follows immediately from the definition of $G(\gamma)$. □

The function $G(\gamma)$ above is an instance of another strategy: We might ask an expert to determine for each $X \subseteq S$ the set of problems which can be solved with the skills of X. To characterize the induced function $\delta : 2^S \to 2^Q$ we state the properties which δ should have:

1. It is reasonable to assume that by aquiring more skills, an individual is able to solve at least the same problems as before. Thus, δ should be monotone with respect to $\subseteq$.
2. It is also reasonable that $\delta(\emptyset) = \emptyset$, and that $\delta(S) = Q$, so that solving a problem in Q needs at least one skill from S, and that the skills in S are sufficient to solve any problem in Q. If the latter condition is not fulfilled, then the problems are not adequate for the skills in S, i.e. we are posing a problem q for which there is no skill at hand with which q can be solved.

Consequently, we call $\delta : 2^S \to 2^Q$ a *problem function* if

1. $\delta(\emptyset) = \emptyset$
2. $\delta(S) = Q$
3. δ is monotone with respect to $\subseteq$.

We do not require that a problem function δ be a homomorphism with respect to $\cup$ since the combination of sets of skills X and Y may enable an individual to solve additional problems which are not in the union of $\delta(X)$ and $\delta(Y)$. This seems a reasonable assumption: Suppose that $Q = \{q\}$, and that q is a task that can be solved by the skills a and b, but not by a or b alone. Then, $\delta(\{a\}) = \delta(\{b\}) = \emptyset$, whereas $\delta(\{a, b\}) = Q$.

A problem function δ is said to *separate points* if for all $p, q \in Q,\ p \neq q$, there is some $X \subseteq S$ such that $|\delta(X) \cap \{p, q\}| = 1$. δ induces a quasiorder $\preceq$ on Q by

$$p \preceq q \stackrel{\text{def}}{\Longleftrightarrow} (\forall X \subseteq S)(q \in \delta(X) \to p \in \delta(X).$$

Intuitively, $p \preceq q$ if solving q implies solving p. The following is easy to check:

Proposition 2.2. *Let $\delta : 2^S \to 2^Q$ be a problem function. Then, δ separates points if and only if $\preceq$ is antisymmetric.* □

Point separating problem functions are useful when one wants to group problems which test the same skills (cf. the *notions* mentioned above).

Any skill function induces a problem function using the assigment $G : \gamma \mapsto G(\gamma)$ above. The next result shows that these two concepts are, in fact, equivalent:

Proposition 2.3. *Let $\mathbf{P}$ be the set of all problem functions, and $\mathbf{S}$ the set of all skill functions with respect to S and Q. Then, the assignment G defined above is bijective between $\mathbf{S}$ and $\mathbf{P}$.*

Proof. 1. G is injective: Let $\gamma_0,\ \gamma_1 : Q \to 2^{2^S}$ be skill functions such that $\gamma_0 \neq \gamma_1$. Then, there is some $q \in Q$ such that $\gamma_0(q) \neq \gamma_1(q)$. Suppose without loss of generality that $X \in \gamma_0(q) \setminus \gamma_1(q)$. Since $X \in \gamma_0(q)$, we have $q \in G(\gamma_0)(X)$. If $q \notin G(\gamma_1)(X)$, we are done. Otherwise, there is some $Y \in \gamma_1(q)$ with $Y \subseteq X$, and clearly $q \in G(\gamma_1)(Y)$. If $q \in G(\gamma_0)(Y)$, then there is some $Z \in \gamma_0(q)$ with $Z \subseteq Y$. Since $Y \subseteq X$, and the elements of $\gamma_0(q)$ are pairwise incomparable, this cannot happen. Hence, $G(\gamma_1)(Y) \neq G(\gamma_0)(Y)$, and therefore, G is one-one.

2. G is surjective: Fix some $\delta \in \mathbf{P}$. For $q \in Q$, let

$$E(q) \stackrel{\text{def}}{=} \{X \subseteq S : q \in \delta(X)\},$$

and define $\gamma : Q \to 2^{2^S}$ by

$$\gamma(q) \stackrel{\text{def}}{=} \{X \subseteq S : X \text{ is minimal in } E(q)\}.$$

We are going to show that $\gamma \in \mathbf{S}$, and $G(\gamma) = \delta$.

Suppose that $q \in Q$. Since $q \in \delta(S) = Q$, $E(q) \neq \emptyset$, and since $\delta(\emptyset) = \emptyset$, we have $\emptyset \notin E(q)$. It follows that $\gamma(q)$ is not empty and contains only nonempty sets. Since each of these is minimal in $E(q)$, the elements of $\gamma(q)$ are pairwise incomparable.

Let $X \subseteq S$. Then, $G(\gamma)(X) = \delta(X)$:

"$\subseteq$": Let $q \in G(\gamma)(X)$; then, there is some $Y \in \gamma(q)$ with $Y \subseteq X$. By definition of $\gamma(q)$, $Y \in E(q)$, and hence, $q \in \delta(Y)$. Since δ is monotone, and $Y \subseteq X$, it follows that $q \in \delta(X)$.

"$\supseteq$": Let $q \in \delta(X)$. Then, $X \in E(q)$. Suppose that $Y \subseteq X$ is minimal in $E(q)$ and $q \in Y$. Then, $Y \in \gamma(q)$, and by definition of $G(\gamma)$ and $Y \subseteq X$ we have $q \in G(\gamma)(X)$. □

One consequence of this theorem is that there are two equivalent theoretical formulations of knowledge structures connected to skills: One based on a skill function, the other one on its associated problem function. Let us call a triple $\langle Q, S, \delta \rangle$ a *skill - knowledge structure*, where $\delta : 2^S \to 2^Q$ is a problem function.

It was asked in Falmagne et al. (1990), p. 208, under which conditions the set of all γ-knowledge states is a knowledge space. We can give a sufficient condition in terms of δ:

Proposition 2.4. *Let $\gamma : Q \to 2^{2^S}$ be a skill function, and δ its associated problem function. If δ preserves union, then $K = ran(\delta)$ is a knowledge space.*

Proof. Suppose that δ preserves $\cup$, and let $A, B \in K$, $A = \delta(X),\ B = \delta(Y)$. Then,

$$A \cup B = \delta(X) \cup \delta(Y) = \delta(X \cup Y) \in K.$$

□

The condition, however, is not necessary, as the following example shows:

Let $Q = \{p, q, r\},\ S = \{1, 2, 3\}$, and

$$\gamma(q) = \{\{1, 3\}, \{2, 3\}\}, \gamma(p) = \{\{1\}\}, \gamma(r) = \{\{2\}\}.$$

Then, $K = \{\emptyset, \{1\}, \{2\}, \{1, 2\}, Q\}$ is a knowledge space, but

$$\{p, r\} = \delta(\{1, 2\}) \cup \delta(\{3\}) \neq \delta(\{1, 2, 3\}) = Q.$$

3 Constructing skill and problem functions

Let $\langle Q, K \rangle$ be a knowledge structure. Is it possible to find an abstract set of skills and a problem function δ such that $K = ran(\delta)$? The answer is: Definitely Yes! In this section we shall describe two strategies for finding such a set of skills, one based on antichains, the other being a more refined construction.

Let $\langle Q, K \rangle$ be a knowledge structure. To facilitate notation, we call $A \in K$ *proper*, if $A \neq \emptyset,\ Q$, and set $K^+ \stackrel{\text{def}}{=} K \setminus \{\emptyset, Q\}$.

3.1 Antichain based problem functions

Our first construction shows that any knowledge structure can be realised as a skill - knowledge structure via pairwise incomparable sets of skills. This leads to an upper bound of the number of skills needed for a problem function $\delta : 2^S \twoheadrightarrow K$ to exist.

Construct 1. If K has just two elements, we can choose any set S with just one element. Otherwise, let $B = 2^S$ be a Boolean set algebra with an antichain M of size $m = |K^+|$, say, $M = \{m_A : A \in K^+\}$. Now, define $\delta : 2^S \twoheadrightarrow K$ by

$$\delta(X) = \begin{cases} Q, & \text{if } m_A \subsetneq X \text{ for some } A \in K^+, \\ A, & \text{if } X = m_A, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Clearly, δ is onto, and we also have $\delta(S) = Q,\ \delta(\emptyset) = \emptyset$.

To show that δ preserves $\subseteq$, let $X \subseteq Y \subseteq S$. We can assume that $\emptyset \neq X \supsetneq Y \subsetneq S$. The first condition tells us that $m_A \subseteq X$ for some $A \in K \setminus \{\emptyset, Q\}$, and from $X \subsetneq Y$ we infer that $\delta(Y) = Q$. Thus, $\delta(X) \subseteq \delta(Y)$.

Thus, we have shown

Theorem 3.1. [2] *Let $\langle Q, K \rangle$ be a knowledge structure. Then, there is a set S and a problem function $\delta : 2^S \to 2^Q$ such that $ran(\delta) = K$.*

If, in the construction above, the chosen antichain M is the set of atoms of 2^S, then we call δ the *indicator problem function*.

Example: Let

$$Q = \{1, 2, 3\},\ S = \{a, b, c, d\},$$
$$K = \{\emptyset, \{1\}, \{2\}, \{1, 3\}, \{2, 3\}, Q\}.$$

Applying the indicator problem function we obtain

$$\begin{aligned} \delta(\emptyset) &= \emptyset, \\ \delta(\{a\}) &= \{1\}, \\ \delta(\{b\}) &= \{2\} \\ \delta(\{c\}) &= \{1, 3\} \\ \delta(\{d\}) &= \{2, 3\} \\ \delta(X) &= Q, \text{ otherwise.} \end{aligned}$$

The associated skill function is given by

$$\begin{aligned} \gamma(1) &= \{\{a\}, \{c\}, \{b, d\}\}, \\ \gamma(2) &= \{\{b\}, \{d\}, \{a, c\}\}, \\ \gamma(3) &= \{\{c\}, \{d\}, \{a, b\}\}, \end{aligned}$$

[2] A similar result was independently obtained in Doignon (????)

Observe that $\gamma(q)$ has more than one element for each $q \in Q$, and that each skill solves at least one problem.

It is instructive to consider another skill function which realizes the same set of states, and which shows that the using the indicator problem function does not necessarily produce the least number of skills:

Let $S = \{a, b, c\}$, and γ given by

$$\begin{aligned}\gamma(1) &= \{\{a\}\},\\ \gamma(2) &= \{\{b\}\},\\ \gamma(3) &= \{\{a,c\},\{b,c\},\{a,b\}\}.\end{aligned}$$

The corresponding δ is induced by

$$\begin{aligned}\delta(\{a\}) &= \{1\},\\ \delta(\{b\}) &= \{2\},\\ \delta(\{c\}) &= \emptyset,\\ \delta(\{a,b\}) &= Q,\\ \delta(\{a,c\}) &= \{1,3\},\\ \delta(\{b,c\}) &= \{2,3\}.\end{aligned}$$

The fact that skill c by itself does not solve any problem is reflected by $\delta(\{c\}) = \emptyset$.

In Falmagne et al. (1990), p. 208, it is stated that *"An interpretation of a knowledge space in terms of a structure of underlying skills is sometimes possible"*. The result above shows that this is in principle always possible, even for structures more general than knowledge spaces.

3.2 A bound for the number of skills

The preceding example shows that a knowledge structure $\langle K, Q\rangle$ can be made into a skill - knowledge structure in different ways. The question arises how many skills are needed for a problem function $\delta : 2^S \twoheadrightarrow K$ to exist. Clearly, $|S| \geq log_2(|K|)$, and $log_2(|K|)$ skills suffice, in case K is a Boolean algebra. The next result shows that no matter what K looks like, we never need much more than $log_2(|K^+|)$ skills.

We shall prove one preliminary Lemma:

Lemma 3.2. *If $n \geq 5$,*

$$log_2\binom{n}{n\ div\ 2} + log_2\left(\log_2\binom{n}{n\ div\ 2}\right) > n$$

Proof. Inspection shows that the statement is true for n=5. Suppose it is true for n=k. If k is odd, then

$$\binom{k+1}{(k+1)\ div\ 2} = 2\cdot\binom{k}{k\ div\ 2},$$

and one line implies the next:

$$\begin{aligned}
log_2\binom{k}{k\ div\ 2} + log_2\left(\log_2\binom{k}{k\ div\ 2}\right) &> k,\\
log_2\binom{k}{k\ div\ 2} + 1 + log_2\left(\log_2\binom{k}{k\ div\ 2}\right) &> k+1,\\
log_2\binom{k+1}{(k+1)\ div\ 2} + log_2\left(\log_2\binom{k}{k\ div\ 2}\right) &> k+1,\\
log_2\binom{k+1}{(k+1)\ div\ 2} + log_2\left(\log_2\binom{k+1}{(k+1)\ div\ 2}\right) &> k+1.
\end{aligned}$$

Now, suppose that $k = 2 \cdot m$. First, observe that

$$\binom{k+1}{(k+1)\ div\ 2} = \binom{k+1}{m} = \frac{k+1}{m+1}\binom{k}{m}$$

and thus

$$log_2\binom{k+1}{m} = log_2\left(\frac{k+1}{m+1}\right) + log_2\binom{k}{m}.$$

Using this and the fact that

$$\left[log_2\left(\frac{k+1}{m+1}\right)\right] \cdot (2m+1) \geq log_2\binom{k}{m},$$

it is straightforward, if somewhat tedious, to show that

$$\begin{aligned}
log_2\binom{k+1}{m} + log_2\left(log_2\binom{k+1}{m}\right) &\geq log_2\binom{k}{m} + log_2\left(log_2\binom{k}{m}\right) + 1\\
&> k+1,
\end{aligned}$$

which proves the assertion. □

Proposition 3.3. *Let $\langle Q, K\rangle$ be a knowledge structure with $|K^+| = m \geq 6$. Then, there are a set S with $|S| \leq \lceil log_2(m) + log_2(log_2(m))\rceil$ and a problem function $\delta : 2^S \twoheadrightarrow K$.*

Proof. Let n be the smallest number with $\binom{n}{n\ div\ 2} \geq m$. The construction of 1 shows that any set S with $|S| = n$ will give rise to the desired problem function if we use sets $A \subseteq S$ with $|A| = \binom{n}{n\ div\ 2}$ to be mapped to the proper states, since the sets of this form are an antichain of size $\geq m$.

If $\binom{n}{n\ div\ 2} = m$, the claim follows immediately from 3.2. Otherwise,

$$\binom{n-1}{(n-1)\ div\ 2} < m < \binom{n}{n\ div\ 2},$$

and, again by 3.2,

$$\begin{array}{ccccccc}
log_2(m) & + & log_2(log_2(m)) & + & 1 & > & \\
log_2\binom{n-1}{(n-1)\ div\ 2} & + & log_2\left(log_2\binom{n-1}{(n-1)\ div\ 2}\right) & + & 1 & > & n
\end{array}$$

whence the claim follows. □

Let K^+ be an antichain of size m, and $\delta : 2^S \twoheadrightarrow K$ a problem function. Then, 2^S must have an antichain of size m. Using Sperner's Lemma[3] (see Grätzer (1978), p. 188), it is straightforward to show that for any finite Boolean algebra B with n atoms, B has an antichain of size m if and only if $\binom{n}{n\ div 2} \geq m$. If e.g. $m = 21$, then the smallest such n is 7, and it is easily checked that $\lceil log_2(21) + log_2(log_2(21))\rceil = 7$.

[3]We should like to thank one of the referees for reminding us of Sperner's Lemma

3.3 The order structure of states and competencies

The preceding section shows that any knowledge structure can be obtained by associating its states with incomparable subsets of an abstract skill set. The practical value of this procedure, however, is rather limited for two reasons:

1. Although the state identifying mapping can be formulated parsimoniously in terms of the number of abstract skills, it is not economic in the number of different strategies or competencies to solve a problem. Since the minimization of disjunctive terms within a theory will be advantageous to the researcher, it would be helpful to have a construction which minimizes the number of strategies.

2. The method of Construction I does not reflect any relationship among the states.

These two points are, indeed, related: Let $\langle Q, K \rangle$ be a knowledge structure, and $q \in Q$. Following Falmagne et al. (1990), we call $A \in K$ a *clause* for q, if A is a minimal element of the set $\{B \in K : q \in B\}$. The clauses induce a mapping $\eta : Q \to 2^{2^Q}$ by letting $\eta(q)$ be the set of clauses for q, which in Falmagne et al. (1990) is called a *surmise mapping*. Different clauses for q will require different strategies for solving q, which are incomparable with respect to $\subseteq$. This gives us a lower bound for the required number of strategies: Whenever $\gamma : Q \to 2^{2^Q}$ is a skill function, then $|\gamma(q)| \geq |\eta(q)|$.

If K is a knowledge space, then it can be recovered from η, since in a lattice, every element is the join of irreducible elements. Furthermore, an underlying skill structure can be found in which the order structure of the clauses completely reflects the order structure of the competencies:

Proposition 3.4. *Let $\langle Q, K \rangle$ be a knowledge space, and η its surmise function. Then, regarding Q as a set of skills (i.e. identifying each problem q with a unique skill which solves q), η is a skill function, and, if $\delta : 2^Q \to 2^Q$ is its associated problem function, then δ is a retraction onto K.*

Proof. Clearly, the elements of $\eta(q)$ are pairwise incomparable nonempty subsets of Q. It remains to show that $ran(\delta) \subseteq K$, and that $\delta \restriction K = id_K$:

"$\subseteq$": Let $X \subseteq Q$. If $X \in \{\emptyset, Q\}$, there is nothing to show. Otherwise,

$$\delta(X) = \{q \in Q : (\exists A \in \eta(q))(A \subseteq X)\}.$$

If $q \in \delta(X)$, $A \in \eta(q)$, $A \subseteq X$, and $p \in A$, then there is a clause B for p with $B \subseteq A \subseteq X$. Hence, $p \in \delta(X)$, and thus, $\delta(X)$ is the union of all clauses for q below X. Since K is closed under union, $\delta(X) \in K$.

Let $A \in K$, and w.l.o.g. $A \neq \emptyset, Q$. Then,

$$q \in A \longleftrightarrow (\exists B \in \eta(q))(B \subseteq A) \longleftrightarrow a \in \delta(A),$$

which shows that δ is a retraction. □

This answers a question in Falmagne et al. (1990), p. 208.

It is no accident that in the preceding Proposition, K was a join semilattice. Indeed the next result shows, that this is necessary (and sufficient) for K to be a retract of 2^S, and thus for the order structure of K to be reflected in 2^S:

Proposition 3.5. *Let $\langle M, \leq \rangle$ be a partially ordered set. Then, there are some set S and an order retraction $f : 2^S \twoheadrightarrow M$ if and only if $\langle M, \leq \rangle$ is a lattice.*

Proof. First, note that it is enough to prove the claim for semilattices, since a finite semilattice is, in fact, a lattice, cf. Birkhoff (1967).

"$\rightarrow$": Let $f : 2^S \twoheadrightarrow M$ be a retraction; we can suppose w.l.o.g. that $M \subseteq 2^S$, and that $f \restriction M = id_M$. For $a, b \in M$, let $a \vee b \stackrel{\text{def}}{=} f(a \cup b)$. Since $a, b \subseteq a \cup b$, we have

$$a = f(a) \subseteq f(a \cup b) = a \vee b,$$

and a similar statement holds for b. If $c \in M,\ a, b \subseteq c$, then $a \cup b \subseteq c$, and hence,

$$a \vee b = f(a \cup b) \subseteq f(c) = c.$$

Hence, $\langle M, \leq \rangle$ is a join semilattice.

"$\Leftarrow$": Suppose that M is a join semilattice with smallest element 0 and largest element 1, and set $S \stackrel{\text{def}}{=} M \setminus \{0\}$. Then, the mapping $g :\ M \rightarrow 2^S$, defined by

$$x \mapsto (x] \setminus \{0\}$$

is an order preserving embedding with $g(0) = \emptyset$ and $g(1) = S$. Define $f :\ 2^S \rightarrow M$ by

$$A \mapsto \bigvee \{m \in M : g(m) \subseteq A\}.$$

If $A \subseteq B \subseteq S$, then

$$f(A) = \bigvee \{m \in M : g(m) \subseteq A\} \leq \bigvee \{m \in M : g(m) \subseteq B\} = f(B),$$

and thus, f preserves order. Furthermore, for $x \in M$,

$$f(g(x)) = \bigvee \{m \in M : g(m) \subseteq g(x)\} = x,$$

since g is an embedding. It follows, that M is a retract of 2^S. $\square$

Our second construction aims to refine the indicator problem function in the sense that it takes into account a possible non - lattice structure of K, and allows non singleton subsets of S to be mapped to proper states:

Construct 2. Define a relation $M \subseteq Q \times 2^K$ by setting $(q, \{A_0, ..., A_k\}) \in M$ iff

1. $k = 0$, and $q \in A_0$, or
2. (a) $k > 0$,
   (b) The sets A_i are pairwise incomparable with respect to $\subseteq$, and $\bigcup \{A_i : i \leq k\} \notin K$,
   (c) There exists some $A \in K$ such that
   i) A is a cover of $\{A_i : i \leq k\}$,
   ii) $q \in A - \bigcup \{A_i : i \leq k\}$.

Such an A will be called a witness for $(q, \{A_0, ..., A_k\}) \in M$. Observe that for a knowledge space K, the second condition will never arise. Indeed, this condition is to ensure, that the combination of skills needed to solve the problems in $A_0 \cup \ldots \cup A_k$ enables the subject to solve q: Since the union of the states $A_0, \ldots, A_k$ is not a state, if a subject is able to solve all problems in $A_0, \ldots, A_k$, he or she will be able to solve additional problems.

Our set S of skills will consist of elements labeled with the nonempty knowledge states, so that for each such state there is exactly one skill: Set $S = \{s_A : A \in K, A \neq \emptyset\}$, and, for each $q \in Q$, set $M(q) = \{T \subseteq K : (q,T) \in M\}$. $M(q)$ is just the range of q in the relation M.

The next step is to use the elements of each $T \in M(q)$ as labels and then build $\gamma(q)$ from the resulting sets of skills: For each $T \in M(q)$ let

$$D(q,T) = \{s_B : B \in T\} \subseteq S.$$

Observe that $D(q,T)$ is independent of q in the sense that, if $T \in M(p)$, then $D(p,T) = D(q,T)$.

Finally, set

$$\gamma(q) = \{X \subseteq S : X \textit{ is minimal in} \{D(q,T) : T \in M(q)\}\}.$$

The next result shows that γ provides $\langle Q, K\rangle$ with the desired skill structure:

Proposition 3.6. *1. γ is a skill function.*

2. Let δ be the problem function corresponding to γ. Then, $ran(\delta) = K$.

Proof. 1. This follows immediately from the definition of γ.

2. Recall that for each $X \subseteq S$,

$$\delta(X) \stackrel{\text{def}}{=} \{q \in Q : (\exists Y \in \gamma(q))(Y \subseteq X)\},$$

and hence,

$$\delta(X) = \{q \in Q : (\exists D(q,T) \in \gamma(q))(D(q,T) \subseteq X)\}.$$

Clearly, it suffices to look at nonempty subsets of S. Suppose we have shown that for any nonempty $X \subseteq S$,

(3.1) $$\delta(X) = \bigcup\{A \in K : s_A \in X\} \cup \bigcup\{A \in K : A \text{ witnesses some } D(p,T) \subseteq X\}.$$

Then, we can prove the claim as follows:

"$\subseteq$": Let $X \subseteq S$ be nonempty, and

$$T \stackrel{\text{def}}{=} \{A \in K : s_A \in X\} \cup \{A \in K : A \text{ witnesses some } D(p,T) \subseteq X\},$$

so that $\delta(X) = \bigcup(T)$ by (3.1).

Assume that $\delta(X) \notin K$. Let B be minimal in $\{C \in K : \delta(X) \subseteq C\}$, and $p \in B - \delta(X)$. Then, B covers T, and $p \notin \bigcup T$ by (3.1). It follows that $(p,T) \in M$, and hence there is some $Y \in \gamma(p)$ with $Y \subseteq D(p,T) \subseteq X$. Hence, $p \in \delta(X)$ - a contradiction.

"$\supseteq$": Let $A \in K, A \neq \emptyset$, and set $X \stackrel{\text{def}}{=} \{s_A\}$. Then, $X \in \gamma(q)$ for all $q \in A$, and thus $A \subseteq \delta(X)$.

Conversely, let $q \in \delta(X)$; then $\{s_A\} \in \gamma(q)$, and there exists some $T \in M(q)$ such that $D(q,T) = \{s_A\}$. Since $D(q,T)$ is a singleton, T is a singleton, and, by definition of $D(q,T)$, we have $T = A$. Hence, $q \in A$.

It remains to show $(^*)$:

"$\subseteq$": Let $p \in \delta(X)$. Then, there is some $D(p,T) \in \gamma(p)$ such that $D(p,T) \subseteq X$. Note that $T \in M(p)$ and $(p,T) \in M$. If $T = \{A\}$ for some $A \in K$, then $s_A \in X$, and $p \in A$ by the definition of M. Otherwise, suppose that $T = \{A_0, A_1, ..., A_k\}$. Let $A \in K$ be a witness for $(p,T) \in M$, and $q \in A$. We need to show that $q \in \delta(X)$: If $q \in A_i$ for some $i \leq k$, then $s_{A_i} \in D(p,T)$, and thus $q \in \delta(X)$. Otherwise, $(q,T) \in M$, and consequently $D(p,T) = D(q,T) \subseteq X$, which shows $q \in \delta(x)$.

"$\supseteq$": Let $q \in A \in K$ and $s_A \in X$. Then, $\{s_A\} \in \gamma(q)$, and hence $q \in \delta(X)$. Next, let $A \in K$ be a witness for $D(p,T) \subseteq X$ and $T = \{A_0, A_1, ..., A_k\}$. If $q \in A_i$ for some $i \leq k$, then $s_{A_i} \in D(p,T) \subseteq X$ implies that $q \in \delta(X)$. Otherwise, $(q,T) \in M$ and hence $D(q,T) \subseteq X$, shows $q \in \delta(X)$. $\square$

Examples:

1. $Q = \{1,2,3,4\}, K = \{\emptyset, \{1\}, \{2\}, \{1,2,3\}, \{1,2,4\}, Q\}$. We first determine the sets $M(q)$:

$$\begin{aligned} M(1) &= \{\{1\}, \{1,2,3\}, \{1,2,4\}\}, \\ M(2) &= \{\{2\}, \{1,2,3\}, \{1,2,4\}\}, \\ M(3) &= \{\{1,2,3\}, \{\{1\}, \{2\}\}\}, \\ M(4) &= \{\{1,2,4\}, \{\{1\}, \{2\}\}\}. \end{aligned}$$

   Let $a = s_{\{1\}}$, $b = s_{\{2\}}$, $c = s_{\{1,2,3\}}$, $d = s_{\{1,2,4\}}$. Then,

$$\begin{aligned} \gamma(1) &= \{\{a\}, \{c\}, \{d\}\}, \\ \gamma(2) &= \{\{b\}, \{c\}, \{d\}\}, \\ \gamma(3) &= \{\{c\}, \{a,b\}\}, \\ \gamma(4) &= \{\{d\}, \{a,b\}\}. \end{aligned}$$

2. $Q = \{1,2,3\}, K = \{\emptyset, \{1\}, \{2\}, \{1,3\}, \{2,3\}, Q\}$. $M(q)$ is found to be

$$\begin{aligned} M(1) &= \{\{1\}, \{1,3\}\}, \\ M(2) &= \{\{2\}, \{2,3\}\}, \\ M(3) &= \{\{1,2\}, \{2,3\}, \{\{1\}, \{2\}\}\}. \end{aligned}$$

   Note that $(3, \{\{1\}, \{2\}\}) \in M$ is witnessed by Q.

   Let $a = s_{\{1\}}$, $b = s_{\{2\}}$, $c = s_{\{1,3\}}$, $d = s_{\{2,3\}}$. Then,

$$\begin{aligned} \gamma(1) &= \{\{a\}, \{c\}\}, \\ \gamma(2) &= \{\{b\}, \{c\}\}, \\ \gamma(3) &= \{\{c\}, \{d\}, \{a,b\}\}. \end{aligned}$$

Table 4: Empirical density of the theoretical set A states

State	Problems	n(pretest)	n(posttest)
1	$\emptyset$	0	0
2	$\{1, 2, 3, 4, 5\}$	0	0
3	$\{1, 2, 3, 4, 5, 6\}$	2	1
4	$\{1, 2, 3, 4, 5, 7, 8\}$	2	0
5	$\{1, 2, 3, 4, 5, 6, 7, 8\}$	4	2
6	$\{1, 2, 3, 4, 5, 6, 9, 10\}$	0	4
7	$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$	45	29
8	$\{1, 2, 3, 4, 5, 7, 8, 11\}$	0	0
9	$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$	28	40
10	$\{1, 2, 3, 4, 5, 6, 9, 10, 12\}$	1	1
11	$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$	24	79
	number of hits using all states (base N=241):	106	156

In the construction of the generic skill sets above, we started by assigning a set of skills to each knowledge state different from $\emptyset$ and Q, and these were in fact all the skills. This seems reasonable as a starting point: If a particular state is present, it is safe to suppose that there is some set of skills which solves exactly the problems in that state. Yet, this assignment may be too crude, and later one may want to split up these "initial" skills into sets of more refined skills. Given a knowledge structure $\langle Q, K\rangle$, it may be worthwhile to look at the set of all skill - knowledge structures $\langle S, Q, \delta\rangle$ with $ran\ \delta = K$. There is a natural order on the set of these structures by setting $\langle S_0, Q, \delta_0\rangle \leq \langle S_1, Q, \delta_1\rangle$ iff there is an onto mapping $\theta : S_1 \to S_0$, and for each $X \subseteq S_1$ we have $\delta_0(\theta[X]) \subseteq \delta_1(X)$. In this note, we shall not pursue this further.

4 An illustrative example

4.1 The data set

The data used were orginally published by Wiedl & Bethge (1983). They used the Coloured Progressive Matrices test (Raven (1965)) and applied this test twice. The pretest was presented using the standard test frame. The presentation of the posttest was modified using a dynamic version of the CPM (Carlson & Wiedl (1992)). The subjects were forced to think aloud during the problem solving process, and they had to give a reason for the choiced reaction category. There was no reinforcement. The sample consisted of 241 children, and the age range was between 8.4 and 9.3 years.

4.2 Description of the method

Applying discrete models to empirical data faces a problem: Empirical deviations from the predicted state can be a consequence of wrong model assumptions or a consequence of a probabilistic mechanism. Subjects may solve some problems by chance although they do not have the skills to solve these problems ("false alarm"), or subjects miss a correct solution of a problem although they have the skills to solve that problem.

In the sequel we assume that both $p[misses]$ and $p[false\ alarm]$ are very small. In this case, the observed empirical frequencies reflect the theoretical structure. If both probabilities are substantial, we require a probabilistic model, but, up to now, only primary steps towards such a probabilistic model have been taken (Falmagne (????a)).

Based on this rather crude assumption we can directly measure the success of a given theory of skills comparing the subjects that are correctly classified with the number of all subjects.

If there is the possibility to cross-validate a given model, we can search for a collection of states using the sorted empirical frequencies of observed (possible) states in sample A and collect the most frequent states until a satisfactory result is obtained. Sample B can be used to validate the constructed set of states. Using the constructed set of possible states K, we use $K' = K \cup \{\emptyset, Q\}$ to construct a theoretical skill function. The observed states and/or the constructed skills can be used to update an unsuccessful theory.

4.3 Results

Tab. 4 shows the predicted states and their empirical frequencies using Raven's skill assumptions. The number of subjects fitting the theory is rather disappointing. The problem verbalization instruction seems to push subjects towards the model assumptions. One might argue that the bad model representation is due to random fluctuations within the data. But Tab. 5 shows that we need only a few states to represent the data at least as well and we can do much better, if we use better constructed states.

Table 5: States sorted by pretest frequencies

	Construction		validation	
Problems	n(pretest)	cum(pre)	n(posttest)	cum(post)
$\{1,2,3,4,5,6,7,8,9,10\}$	45	45	29	29
$\{1,2,3,4,5,6,7,8,9,10,11\}$	28	73	40	69
$\{1,2,3,4,5,6,7,8,9,10,11,12\}$	24	97	79	148
$\{1,2,3,4,5,6,8,9,10\}$	17	114	4	152
$\{1,2,3,4,5,6,7,8,9\}$	9		4	
$\{1,2,3,4,5,6,7,8,10,12\}$	9	132	21	177
$\{1,2,3,4,5,6,7\}$	6		1	
$\{1,2,3,4,5,6,7,9,10\}$	6		7	
$\{1,2,3,4,5,6,7,9,10,11\}$	6	150	5	190

To explain the data as well as the skill theory does, we need only 5 states ($\emptyset$ included) in both data sets. Tab. 6 shows possible skill theories deduced from these states whose empirical frequencies are greater than a fixed value.

A very simple and empirically relative successful skill theory is offered by the states at $Cut > 6$: All problems need a basic skill, problems 7 and 10 both need an additional skill each, and problems 11 and 12 need these three skills and something extra. Problems 11 and 12 can be solved in two different ways.

The interpretation of the resulting skill function is only valid in the population of subjects under investigation. There might be more states that are not realized in that population (e.g the state $\emptyset$). Nevertheless, the states that are realized must be a substantial part of any skill theory.

Table 6: CPM set A: Some skill theories

Problem	Raven	$Cut > 9$	$Cut > 6$	$Cut > 5$
1	B	a	a	a|b
2	B	a	a	a|b
3	B	a	a	a|b
4	B	a	a	a|b
5	B	a	a	a|b
6	BO	a	a	a|b
7	BG	a	ab	a
8	BG	a	a	b|ac
9	B1O	a	a	b|ac|ad
10	B1O	ab	ac	b|ad|ace
11	B12G	abc	abcd|abcef	adh|acej|adfj|bgj
12	B12OC	abcd	abce|abcdg	acek|adfk|bgk|adhil
no. of states	11	5	7	10

5 Discussion

The main aim of this paper is the characterization of knowledge structures based on skill functions. We have shown that any knowledge structure can be supplied with an underlying skill set S and a problem function $\delta : 2^S \twoheadrightarrow K$. Only approximately $log_2(|K|)$ abstract skills are required to obtain a proper skill representation of any knowledge structure.

We present two constructions to obtain a skill representation of a given knowledge structure:

The first one, using the indicator problem function, is a theoretical tool. Using that tool it was possible to calculate the lower bound of the number of necessary skills. The second construction is an exploratory tool for obtaining a sound skill representation which might give the researcher some hints to build a better theory.

Up to now, the theory is faced with some more or less severe problems:

1. The number of possible skill representations is infinite, and therefore we need a description what the skill representations look like, given a knowledge structure $\langle Q, K\rangle$ (see the remarks at the end of Section 3).

2. There is a problem of handling noise in data. Only if we have a model in which there are mechanisms to describe the results of the problem solving behaviour of a subject v which is in state A and tries to solve a problem $q \notin A$, can we safely distinguish between model error and random fluctuation. The ideas in Falmagne (????b) of probabilistic versions of knowledge spaces may possibly be transferred to knowledge structures.

3. Using the theory we shall always get an empirical result, because the proposed constructions recode the data in another structure of sets, namely, that of abstract skills. The consequence is that a skill structure constructed with a data set can never be invalidated using the data itself. Nevertheless, any proposed skill structure can be evaluated using an empirical data set. Note, that e.g. the constructed

skill structure in our illustrative example was validated by results of a second data set. We hold the view that describing data and testing theories are different problems. The proposed algorithms serve as exploratory data analysis tools. If we want to test theories we shall need a statistical framework (see Problem 2).

4. There exist effective algorithms to generate a knowledge space based theory by querying experts Koppen & Doignon (1990). A fourth problem, the investigation of which is the scope of Düntsch & Gediga (1996), is to find query procedures to build knowledge structures.

References

G. BIRKHOFF, *Lattice Theory*, volume 25 of *Colloquium Publications*, Amer. Math. Soc., Providence, 3rd edition (1967).

J. S. CARLSON & K.-H. WIEDL, Principles of dynamic assessment: The application of a specific model, *Advance in Cognition and Educational Practice*, 1, 235–252 (1992).

J.-P. DOIGNON, Knowledge spaces and skill assignments (????).

J.-P. DOIGNON & J.-C. FALMAGNE, Spaces for the assessment of knowledge, *Internat. J. Man–Mach. Stud.*, 23, 283–303 (1985).

I. DÜNTSCH & G. GEDIGA, On query procedures to build knowledge structures, *J. Math. Psych.*, 40(2), 160–168 (1996).

J.-C. FALMAGNE, Finite Markov learning models for knowledge structures (????a).

J.-C. FALMAGNE, Probabilistic knowledge spaces: A review (????b).

J.-C. FALMAGNE, M. KOPPEN, M. VILLANO, J.-P. DOIGNON, & J. JOHANNESEN, Introduction to knowledge spaces: How to build, test and search them, *Psychological Review*, 97, 202–234 (1990).

G. GRÄTZER, *General Lattice Theory*, Birkhäuser, Basel (1978).

M. KOPPEN & J.-P. DOIGNON, How to build a knowledge space by querying an expert, *Journal of Mathematical Psychology*, 34, 311–331 (1990).

J. C. RAVEN, *Guide to using the couloured progressive matrices sets A, Ab, B (Revised version)*, Gieve, London (1965).

K.-H. WIEDL & H. J. BETHGE, Der CPM-PV-Test: Materialien zur Entwicklung und Anwendung einer dynamischen Version des farbigen Matrizentests von Raven, *Forschungsberichte aus dem Fachbereich Psychologie der Universität Osnabrück*, 30 (1983).

erweitert um DOFA85