8

Learning Spaces: A Mathematical Compendium

Jean-Paul Doignon[1]
Jean-Claude Falmagne[2]
Eric Cosyn[3]

The core of an educational software based on learning space theory, such as the ALEKS system, is a combinatoric structure representing the cognitive organization of a particular curriculum, like beginning algebra or 4th grade arithmetic[4]. This structure consists in a family $\mathcal{K}$ of subsets of a basic set Q. The elements of Q are the *types* of problems to be mastered by a student learning the curriculum. An example of a problem type in beginning algebra is:

[P] *Express the roots of the equation* $\alpha x^2 + \beta x + \gamma = 0$ *in terms of* α, β *and* γ.

When this type of problem is proposed to a student, either in an assessment or in the course of learning, ALEKS chooses an *instance* of [P], which may be for example

[I] *What are the roots of the equation* $4x^2 + 6x - 7 = 0$?

Typically, there are thousands of instances for a particular problem type. These instances are features of the implementation and do not play any role in the theory summarized in this chapter. The set Q is called the 'domain' of the structure, and the elements of Q, the problem types, are referred to as 'items[5]' or 'questions.'

The family $\mathcal{K}$ is the 'learning space.' Its elements are called '(knowledge) states.' Every knowledge state is a set K of items that a student in that state has mastered. To wit, a student in state K cannot solve any instance of an item in $Q \setminus K$ (the instances are constructed so that lucky guesses are

[1] Dept. of Mathematics, University of Brussels.
[2] Dept. of Cognitive Sciences, University of California, Irvine, and ALEKS Corp.
[3] ALEKS Corp.
[4] Or geometry, precalculus, basic chemistry, etc. Parts of the text are excerpts of "Learning Spaces", a monograph by Falmagne and Doignon (2011).
[5] Note that 'item' is used here in a sense different from that used in standardized testing, where 'item' means what we call an 'instance.'

impossible or very rare), and moreover the student can solve any instance of any item in K, barring careless errors. In principle, the family $\mathcal{K}$ contains all such feasible knowledge states in the population of students considered. We assume that both Q and $\varnothing$ are in $\mathcal{K}$: the student may have mastered all the items in Q, or none of them; accordingly, $Q = \cup\mathcal{K}$.

By design, there are no educational gaps in the set Q or in the learning space $\mathcal{K}$. This means, for example, that a student capable of solving all the items in the domain Q of beginning algebra can be regarded as having mastered this curriculum as it is specified in the US schools. It also means that there is no gap in the learning sequence: whatever the student's knowledge state in beginning algebra at any moment, he or she can in principle learn the rest of the curriculum by gradually mastering the remaining items one by one.

Two kinds of arguments support the last statement. For one, the axioms [L1] and [L2] constraining the family $\mathcal{K}$ of knowledge states, which are specified in the next section, are consistent with the idea that, from a theoretical standpoint, gradual learning is feasible in all cases. The second argument is empirical: extensive data on student learning, based on millions of assessments, indicate that when a student is deemed by `ALEKS` ready to learn an item, then the estimated probability of successful mastery of that item is extremely high. Note that the "ready to learn" in the above sentence is given a mathematical meaning as the 'outer fringe' of a student's state (see Definition 8.3.1).

The number of items in a typical domain of school mathematics satisfying the educational standards of a U.S. state is around 650. The number of knowledge states in the learning space for such a domain is quite large, maybe on the order of 10^8. Despite this large number of states, it is nevertheless possible to assess the knowledge state of a student, accurately, in the span of 25-35 questions.

The next few sections give a concise set-theoretical presentation of the concepts and results. No proofs of results are given here[6].

8.1 Axioms for Learning Spaces

8.1.1 Definition. A *partial knowledge structure* is a pair $(Q, \mathcal{K})$ in which Q is a nonempty set, and $\mathcal{K}$ is a family of subsets of Q containing at least Q. The set Q is called the *domain* of the knowledge structure. The subsets in the family $\mathcal{K}$ are labeled *(knowledge) states*. The elements of Q are called *items* or *questions*. A partial knowledge structure is a *knowledge structure* if $\mathcal{K}$ also

[6] For a comprehensive presentation of the theory, see Falmagne and Doignon (2011). An evolutive and searchable database of references on this topic is maintained by Cord Hockemeyer at the University of Graz, Austria: `http://wundt.kfunigraz.ac.at/hockemeyer/bibliography.html`.

contains the empty set. The knowledge structure $(Q, \mathcal{K})$ is *finite* if Q is a finite set. Note that since $Q = \cup\mathcal{K}$ we can without ambiguity refer to the family $\mathcal{K}$ itself as a knowledge structure.

The knowledge structure $(Q, \mathcal{K})$ is *discriminative* if for all items $q, p \in Q$ we have

$$(\forall K \in \mathcal{K})(q \in K \iff p \in K) \quad \implies \quad q = p.$$

Axioms. A knowledge structure $(Q, \mathcal{K})$ is called a *learning space* if it satisfies the two following conditions.

[L1] Learning smoothness. If $K \subset L$ are two states, there is a finite chain of states

$$K_0 = K \subset K_1 \subset \cdots \subset K_n = L \tag{8.1}$$

with $K_i = K_{i-1} + \{q_i\}$ and $q_i \in Q$ for $1 \leq i \leq n$. We have thus $|L \setminus K| = n$.

[L2] Learning consistency. If $K \subset L$ are two states, with $q \notin K$ and $K + \{q\} \in \mathcal{K}$, then $L \cup \{q\} \in \mathcal{K}$.

Axiom [L1] implements the assumption that gradual, item by item, learning is always possible. Note that, by Axiom [L1], any learning space is finite. Axiom [L2] formalizes the idea that if some item q is learnable by a student in some state K which is included in some state L, then either q is in L or it is learnable by a student in state L.

An example of a learning space $\mathcal{H}$ on the domain $\{a, b, c, d\}$ is given by the equation

$$\begin{aligned} \mathcal{H} = \{\varnothing, \{a\}, \{c\}, \{a,b\}, \{a,c\}, &\{c,d\}, \{a,b,c\}, \{a,b,d\}, \\ &\{a,c,d\}, \{a,b,c,d\}\}, \end{aligned} \tag{8.2}$$

which is represented in Figure 8.1 by the Hasse diagram of its inclusion relation.

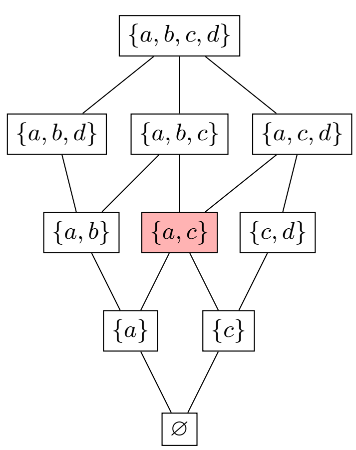

Figure 8.1. Hasse diagram of the inclusion relation of the learning space $\mathcal{H}$ defined by Eq. (8.2) on the domain $\{a, b, c, d\}$. Ignore the red shading for the moment (see Definition 8.3.1).

Note that the family $\mathcal{H}$ is closed under union. This is one of several important properties implied by Axioms [L1] and [L2]. We introduce them in the next definition.

8.1.2 Definition. Let $(Q, \mathcal{K})$ be a knowledge structure. When $\mathcal{K}$ is *closed under union*, that is, when $\cup\mathcal{A} \in \mathcal{K}$ whenever $\mathcal{A} \subseteq \mathcal{K}$, we say that $(Q, \mathcal{K})$ is a *knowledge space*, or equivalently, that $\mathcal{K}$ is a *knowledge space (on Q)*. Note that when a family is closed under union, we sometimes say for short that it is *$\cup$-closed*. The *dual* of a knowledge structure $\mathcal{K}$ on Q is the knowledge structure $\overline{\mathcal{K}}$ containing all the complements of the states of $\mathcal{K}$, that is,

$$\overline{\mathcal{K}} = \{K \in 2^Q \,|\, Q \setminus K \in \mathcal{K}\}.$$

Thus, $\mathcal{K}$ and $\overline{\mathcal{K}}$ have the same domain.

We denote by $K \triangle L = (K \setminus L) \cup (L \setminus K)$ the *symmetric difference* between two sets K and L, and by $d(K, L) = |(K \setminus L) \cup (L \setminus K)|$ the *symmetric difference distance* between those sets.

A family of sets $\mathcal{K}$ is *well-graded* if, for any two distinct sets K, L in $\mathcal{K}$, there exists a finite sequence $K_0 = K, K_1, \ldots, K_n = L$ of sets in $\mathcal{K}$ such that $d(K_{i-1}, K_i) = 1$ for $1 \leq i \leq n$ and $n = d(K, L)$. We call the sequence of sets (K_i) a *tight path* from K to L. It is clear that a well-graded knowledge structure is discriminative. It is also necessarily finite since we can take $K = \varnothing$ and $L = Q$.

A family $\mathcal{K}$ of subsets of a finite set $Q = \cup\mathcal{K}$ is an *antimatroid*[7] if it is closed under union and moreover satisfies the following axiom.

[MA] If K is a nonempty subset of the family $\mathcal{K}$, then there is some $q \in K$ such that $K \setminus \{q\} \in K$.

We may also say then that the pair $(Q, \mathcal{K})$ is an antimatroid. An antimatroid $(Q, \mathcal{K})$ is *finite* if Q is finite. In such a case, $(Q, \mathcal{K})$ is a discriminative knowledge structure.

The next theorem specifies the relationship between these various concepts.

8.1.3 Theorem. *For any knowledge structure $(Q, \mathcal{K})$, the following three conditions are equivalent.*

(i) $(Q, \mathcal{K})$ *is a learning space.*
(ii) $(Q, \mathcal{K})$ *is an antimatroid.*
(iii) $(Q, \mathcal{K})$ *is a well-graded knowledge space.*

The equivalence of (i) and (iii) was established by Cosyn and Uzun (2009). The proof that (ii) is equivalent to (i) is straightforward and is contained in Falmagne and Doignon (2011). It is clear that, under each of the three conditions, the knowledge structure $(Q, \mathcal{K})$ is discriminative.

[7] Cf. Welsh (1995).

The large number of states in empirical learning spaces may create practical problems of manipulation and storage in a computer's memory. The fact that any learning space is closed under union established by Theorem 8.1.3 plays an important role in enabling an economical representation of any learning space in the form of its 'base.'

8.2 The Base and the Atoms

8.2.1 Definition. The *span* of a family of sets $\mathcal{G}$ is the family $\mathcal{G}'$ containing any set which is the union of some subfamily of $\mathcal{G}$. In such a case, we say that $\mathcal{G}$ *spans* $\mathcal{G}'$. By definition, $\mathcal{G}'$ is then a $\cup$-closed family. A *base* of a $\cup$-closed family $\mathcal{K}$ is a minimal subfamily $\mathcal{B}$ of $\mathcal{K}$ spanning $\mathcal{K}$ (where 'minimal' is meant with respect to set inclusion: if $\mathcal{J}$ spans $\mathcal{K}$ for some $\mathcal{J} \subseteq \mathcal{B}$, then $\mathcal{J} = \mathcal{B}$). By a standard convention, the empty set is the union of the empty subfamily of $\mathcal{B}$. Thus, the empty set never belongs to a base. It is also clear that an element K of some base $\mathcal{B}$ of $\mathcal{K}$ cannot be the union of other elements of $\mathcal{B}$.

8.2.2 Theorem. *Let $\mathcal{B}$ be a base for a knowledge space $(Q, \mathcal{K})$. Then $\mathcal{B} \subseteq \mathcal{F}$ for any subfamily $\mathcal{F}$ of states spanning $\mathcal{K}$. Consequently, a knowledge space admits at most one base. Any finite knowledge space has a base.*

Some knowledge spaces have no base, as for instance, the collection of all the open subsets of the real line.

The base of the learning space $\mathcal{H}$ of Eq. (8.2) and Figure 8.1 is the subcollection

$$\{\{a\}, \{c\}, \{a, b\}, \{c, d\}, \{a, b, d\}\}.$$

In the cases of learning spaces encountered in education, the cardinality of the base of a learning space $\mathcal{L}$ is typically much smaller than the cardinality of $\mathcal{L}$. The example of the family 2^A for any finite set A, in which the base is the collection $\{\{x\} \,|\, x \in A\}$ of all the singleton subsets of A, is suggestive in that regard.

Several efficient algorithms are available for the construction of the base of a knowledge space and for generating a knowledge space from its base (see in particular Dowling, 1993b; Falmagne and Doignon, 2011, Section 3.5, pages 49-50).

The states of the base have an important property.

8.2.3 Definition. Let $\mathcal{F}$ be a nonempty family of sets. For any $q \in \cup\mathcal{F}$, an *atom at q* is a minimal set in $\mathcal{F}$ containing q, where 'minimal' refers to the inclusion relation. A set X in $\mathcal{F}$ is an *atom* if it is an atom at q for some $q \in \cup\mathcal{F}$.

8.2.4 Theorem. *Suppose that a knowledge space has a base. Then this base is formed by the collection of all the atoms.*

This property will play an essential role in the construction of an assessment algorithm for very large learning spaces. It will allow us to manufacture a state from any set of items by forming the union of some atoms of these items (see Step 9 in 8.8.1).

8.3 The Fringe Theorem

In the case of standardized tests the result of an assessment is a number regarded as measuring some aptitude. By contrast, the outcome of an assessment by a system such as `ALEKS` is a knowledge state, which may contain hundreds of items[8]. Displaying such an outcome by a possibly very long list of these items is awkward and not particularly useful. Fortunately, a considerably more concise representation of a knowledge state is available, which is meaningful to a student or a teacher. It relies on the twin concepts of the 'inner fringe' and the 'outer fringe' of a knowledge state.

8.3.1 Definition. The *inner fringe* of a state K in a knowledge structure $(Q, \mathcal{K})$ is the subset of items

$$K^{\mathcal{I}} = \{q \in K \,|\, K \setminus \{q\} \in \mathcal{K}\}.$$

The *outer fringe* of a state K is the subset

$$K^{\mathcal{O}} = \{q \in Q \setminus K \,|\, K \cup \{q\} \in \mathcal{K}\}.$$

For example, the inner fringe and the outer fringe of the state $\{a, c\}$ in the learning space $\mathcal{H}$, which is shaded red in Figure 8.1, are $\{a, c\}^{\mathcal{I}} = \{a, c\}$ and $\{a, c\}^{\mathcal{O}} = \{b, d\}$. The *fringe* of K is the union of the inner fringe and the outer fringe. We write

$$K^{\mathcal{F}} = K^{\mathcal{I}} \cup K^{\mathcal{O}}.$$

Let $\mathcal{N}(K, n)$ be the set of all states whose distance from K is at most n, thus:

$$\mathcal{N}(K, n) = \{L \in \mathcal{K} \,|\, d(K, L) \leq n\}. \tag{8.3}$$

We have then $K^{\mathcal{F}} = \big(\cup\mathcal{N}(K, 1)\big) \setminus \big(\cap\mathcal{N}(K, 1)\big)$. We refer to $\mathcal{N}(K, n)$ as the *n-neighborhood* of the state K. The importance of these concepts lies in the following result.

8.3.2 Theorem. *In a learning space, any state is defined by its two fringes; that is, there is only one state having these fringes.*

In fact, a stronger result holds: a finite knowledge structure is well-graded if and only if any state is defined by its two fringes (Falmagne and Doignon, 2011, Theorem 4.1.7).

[8] C.f. Chapter 1.

The fringes of the states play a major role in the ALEKS system. The fringes can be displayed at the end of an assessment to specify the knowledge state exactly. This is marked progress over the numerical score provided by a standardized test. Indeed, the importance of such a representation lies in the interpretation of the fringes. The inner fringe may be taken as containing the items representing the 'high points' of the student's competence in the topic. The outer fringe is even more important because its items may be regarded as those that the student is ready to learn. That feature plays an essential role in the 'learning module' of the ALEKS system. When an assessment is run as a prelude to learning and returns for a student a knowledge state K, the computer screen displays a window listing all the items in the outer fringe of K. The student may then choose one item in the list and begin to study it. A large set of learning data from the ALEKS system shows that the probability that a student successfully masters an item selected in the outer fringe of his or her state is very high. In beginning algebra, the estimated median probability, based on a very large sample of students, is .92.

8.4 Projections of a Knowledge Structure

The concept of projection for learning spaces is closely related to the concept bearing the same name for media introduced by Cavagnaro (2008) (cf. also Theorem 2.11.6 in Eppstein et al., 2007; **?**). This concept has been encountered in Chapter 1 where an example has been given (see page 19). A more precise discussion requires some construction.

Let $(Q, \mathcal{K})$ be a partial knowledge structure—thus, $\varnothing$ is not necessarily in $\mathcal{K}$—and let Q' be any proper subset of Q. Define a relation $\sim$ on $\mathcal{K}$ by

$$\begin{aligned} K \sim L \quad &\iff \quad K \cap Q' = L \cap Q' \\ &\iff \quad K \triangle L \subseteq Q \setminus Q'. \end{aligned}$$

Thus, $\sim$ is an equivalence relation on $\mathcal{K}$. The equivalence between the two right hand sides is easily checked.

We denote by $[K]$ the equivalence class of $\sim$ containing K, and by $\mathcal{K}_\sim = \{[K] \,|\, K \in \mathcal{K}\}$ the partition of $\mathcal{K}$ induced by $\sim$.

Let $(Q, \mathcal{K})$ be a knowledge structure and take any nonempty $Q' \subset Q$. We say that the family

$$\mathcal{K}_{|Q'} = \{W \subseteq Q \,|\, W = K \cap Q' \text{ for some } K \in \mathcal{K}\}$$

is the *projection* of $\mathcal{K}$ on Q'. We have thus $\mathcal{K}_{|Q'} \subseteq 2^{Q'}$. Note that the sets in $\mathcal{K}_{|Q'}$ may not be states of $\mathcal{K}$. For example the state $\{c, g, j\}$ of the projection pictured in Graph B of Figure 1.5 on page 19 is not a state of the original learning space (see Graph B on that figure). For any $\varnothing \not\sim K \in \mathcal{K}$ and with $[K]$ as above, define

$$\mathcal{K}_{[K]} = \{M \,|\, M = \varnothing \text{ or } M = L \setminus \cap[K] \text{ for some } L \sim K\}.$$

So, if $\varnothing \in \mathcal{K}$, we have $\mathcal{K}_{[\varnothing]} = [\varnothing]$. The families $\mathcal{K}_{[K]}$ are called the *Q'-children* of $\mathcal{K}$, or simply the *children* of $\mathcal{K}$ when the set Q' is obvious from the context. We refer to $\mathcal{K}$ as the *parent* structure. Notice that we may have $\mathcal{K}_{[K]} = \mathcal{K}_{[L]}$ even when $K \not\sim L$.

Here is the key result.

8.4.1 Theorem. *Let $\mathcal{K}$ be a learning space (resp. a well-graded $\cup$-closed family) on a domain Q with $|Q| \geq 2$. The following two properties hold for any proper nonempty subset Q' of Q:*

(i) *the projection $\mathcal{K}_{|Q'}$ of $\mathcal{K}$ on Q' is a learning space (resp. a well-graded $\cup$-closed family);*
(ii) *in either case, the children of $\mathcal{K}$ are well-graded and $\cup$-closed families.*

Note that the singleton $\{\varnothing\}$ is vacuously a partial knowledge structure which is, also vacuously, well–graded and $\cup$-closed.

For a proof, see Falmagne (2008) or Falmagne and Doignon (2011, Theorem 2.4.8).

8.5 Building a Learning Space

At this time, the construction of a learning space is still a demanding enterprise extending over several months. It is based partly on the expertise of competent teachers of the topic. Their input provides a first draft of the learning space. Ideally, if the teachers were omniscient, we could ask them questions such as:

[Q] *Suppose that a student has failed to solve items $p_1, \ldots, p_n$. Do you believe this student would also fail to solve item q? You may assume that chance factors, such as lucky guesses and careless errors, do not interfere in the student's performance.*

Such a query is summarized by the nonempty set $\{p_1, p_2, \ldots, p_n\}$ of items, plus the single item q. Thus, all positive answers to the queries form a relation $\mathcal{P}$ from $2^Q \setminus \{\varnothing\}$ to Q. The expert is consistent with the (unknown) knowledge space $(Q, \mathcal{K})$ exactly when the following equivalence is satisfied for all $A \in 2^Q \setminus \{\varnothing\}$ and $q \in Q$:

$$A \,\mathcal{P}\, q \quad \Longleftrightarrow \quad (\forall K \in \mathcal{K} : A \cap K = \varnothing \;\Rightarrow\; q \notin K). \tag{8.4}$$

The relation $\mathcal{P}$ could thus be used to remove from the family of potential states all the sets K falsifying the implication in the r.h.s. of (8.4). In practice, however, it is only when $|A| = 1$ that a human expert can provide reliable responses to such queries. In this case, $\mathcal{P}$ is essentially a binary relation which, if (8.4) is satisfied, defines a quasi order on Q (reflexive, transitive). The resulting learning space is then closed under both union and intersection[9]. It contains all the 'true' states, but also possibly many fictitious states which must be eliminated by further analysis by other means. This initial space is nevertheless acceptable and can be used with students since it contains all the 'true' states. A statistical analysis of student data allows then to refine the initial learning space and remove the fictitious states (for details, see Chapter 15 and 16 in Falmagne and Doignon, 2011).

The two theorems below play a key role.

8.5.1 Theorem. *Let $(Q, \mathcal{K})$ be a knowledge structure, and suppose that $\mathcal{P}$ is the relation from 2^Q to Q defined by Equation (8.4). Then, necessarily:*

(i) *$\mathcal{P}$ extends the reverse membership relation, that is:*

$$\textit{if}\ \ p \in A \subseteq Q, \ \ \textit{then}\ \ A\mathcal{P}p;$$

(ii) *for any A, $B \in 2^Q \setminus \{\varnothing\}$ and $p \in Q$:*

$$\textit{if}\ A\mathcal{P}b \ \ \textit{and}\ \ B\mathcal{P}p \ \ \textit{for all}\ b \in B, \ \ \textit{then}\ \ A\mathcal{P}p.$$

8.5.2 Definition. An *entailment for* the nonempty domain Q (which may be infinite) is a relation $\mathcal{P}$ from $2^Q \setminus \{\varnothing\}$ to Q that satisfies Conditions (i) and (ii) in Theorem 8.5.1.

8.5.3 Theorem. *There is a one-to-one correspondence between the family of all knowledge spaces $\mathcal{K}$ on the same domain Q, and the family of all entailments $\mathcal{P}$ for Q. This correspondence is defined by the two equivalences*

$$A\,\mathcal{P}\,q \iff (\forall K \in \mathcal{K} : A \cap K = \varnothing \Rightarrow q \notin K), \tag{8.5}$$

$$K \in \mathcal{K} \iff \big(\forall (A, p) \in \mathcal{P} : A \cap K = \varnothing \Rightarrow p \notin K\big). \tag{8.6}$$

For the proof see Koppen and Doignon (1990) and Falmagne and Doignon (2011, Theorem 7.1.5).

A computer algorithm called QUERY has been designed to perform the actual construction of the space on the basis of an entailment (Koppen, 1993, 1994; Falmagne and Doignon, 2011, Chapters 15 and 16). However, as made clear by Theorem 8.5.3, the resulting structure is a knowledge space but not necessarily a learning space. The QUERY algorithm has been amended so as to allow the elimination of a set only when such an elimination does not invalidate Axiom [MA] of an antimatroid (cf. page 134 and Theorem 8.1.3). For details

[9] This results from a classical result from Birkhoff (1937).

about such a construction, see Falmagne and Doignon (2011, Chapter 15 and 16).

A different possibility is to use the QUERY algorithm as such, which delivers a knowledge space, and then judiciously add some of the critically missing states so as to obtain a learning space. Part of Chapter 13 is devoted to this technique.

8.6 Probabilistic Extension

The concept of a learning space is a deterministic one. As such, it cannot provide realistic predictions of students' responses to the problems of a test. There are two ways in which probabilities must enter in a realistic model. For one, the knowledge states will certainly occur with different frequencies in the population of reference. It is thus reasonable to postulate the existence of a probability distribution on the collection of states. For another, a student's knowledge state does not necessarily specify the observed responses. A student having mastered an item may be careless in responding, and make an error. Also, a student may guess the correct response to an item which is not in her state. This may happen, for example, when a multiple choice paradigm is used.

Accordingly, it makes sense to introduce conditional probabilities of responses, given the states. A number of simple probabilistic models are described in Chapter 13 of Falmagne and Doignon (2011). They illustrate how probabilistic concepts can be introduced within knowledge space theory. One of these models provides the context for the assessment algorithm outlined in the next section.

8.7 The Assessment Algorithm

An informal description of an assessment algorithm has been given in Section 1.3 on page 13. The general scheme sketched there is consistent with several formal interpretations. One of them is especially important and is currently used in many schools and colleges[10]. We give below its basic components and axioms[11].

8.7.1 Concepts and notation. Given a learning space $\mathcal{K}$ on a domain Q, any assessment is a realization of a Markovian stochastic process. Our notation is as follows. We use r.v. as an abbreviation for 'random variable.'

---

[10] As part of the `ALEKS` system.
[11] For details, see Falmagne and Doignon (2011, Chapter 13).

n	the step number, or *trial* number, $n = 1, 2, \dots$;
$\mathcal{K}_q$	the subfamily of all the states containing q ;
Λ_+	the set of all positive probability distributions on $\mathcal{K}$;
$\mathbf{L}_n$	a random probability distribution on $\mathcal{K}$; we have $\mathbf{L}_n = L_n \in \Lambda_+$ (so $L_n > 0$) if L_n is the probability distribution on $\mathcal{K}$ at the beginning of trial n ;
$\mathbf{L}_n(K)$	a r.v. measuring the probability of state K on trial n ;
$\mathbf{Q}_n$	a r.v. representing the question asked on trial n; we have $\mathbf{Q}_n = q \in Q$ if q is the question asked on trial n ;
$\mathbf{R}_n$	a r.v. coding the response on trial n: $\mathbf{R}_n = \begin{cases} 1 & \text{if the response is correct} \\ 0 & \text{otherwise.} \end{cases}$
$\mathbf{W}_n$	the random history of the process from trial 1 to trial n ;
ι_A	the indicator function of a set A : $\iota_A(q) = \begin{cases} 1 \text{ if } q \in A \\ 0 \text{ if } q \notin A \end{cases}$;
$\zeta_{q,r}$	with $1 < \zeta_{q,r}$ for $q \in Q, r = 0, 1$, a family of parameters specifying the updating operator (see Axiom [U] below).

The process begins, on trial 1, by setting $\mathbf{L}_1 = L$, for some particular $L \in \Lambda_+$. So, the initial probability distribution is the same for any realization. Any further trial $n > 1$ begins with a value $L_n \in \Lambda_+$ of the random distribution $\mathbf{L}_n$ updated as a function of the event on trial $n-1$. We write for any $\mathcal{F} \subseteq \mathcal{K}$,

$$L_n(\mathcal{F}) = \sum_{K \in \mathcal{F}} L_n(K). \tag{8.7}$$

Three general axioms specify the stochastic process $(\mathbf{L}_n, \mathbf{Q}_n, \mathbf{R}_n)$.

The version of these axioms given below is an important special case.

8.7.2 Axioms.

[U] **Updating Rule.** We have $\mathbb{P}(\mathbf{L}_1 = L) = 1$, and for any positive integer n, with $\mathbf{L}_n = L_n$, $\mathbf{Q}_n = q$, $\mathbf{R}_n = r$, and

$$\zeta_{q,r}^K = \begin{cases} 1 & \text{if } \iota_K(q) \neq r, \\ \zeta_{q,r} & \text{if } \iota_K(q) = r, \end{cases} \tag{8.8}$$

we have

$$L_{n+1}(K) = \frac{\zeta_{q,r}^K L_n(K)}{\sum_{K' \in \mathcal{K}} \zeta_{q,r}^{K'} L_n(K')}. \tag{8.9}$$

This updating rule is called *multiplicative with parameters* $\zeta_{q,r}$.

[Q] **Questioning Rule.** For all $q \in Q$ and all integers $n > 0$,

$$\mathbb{P}(\mathbf{Q}_n = q \,|\, \mathbf{L}_n, \mathbf{W}_{n-1}) = \frac{\iota_{S(L_n)}(q)}{|S(L_n)|} \tag{8.10}$$

where $S(L_n)$ is the subset of Q containing all those items q minimizing

$$|2L_n(\mathcal{K}_q) - 1|.$$

Under this questioning rule, which is called *half-split*, we must have $\mathbf{Q}_n \in S(L_n)$ with a probability equal to one. The questions in the set $S(L_n)$ are then chosen with equal probability.

[R] **Response Rule.** For all positive integers n,

$$\mathbb{P}(\mathbf{R}_n = \iota_{K_0}(q) \,|\, \mathbf{Q}_n = q, \mathbf{L}_n, \mathbf{W}_{n-1}) = 1$$

where K_0 is the *latent state* representing the set of all the items currently mastered by the student, that is, the state that must be uncovered by the Markovian procedure.

So, if the item selected by the process belongs to the latent state K_0, the probability of a correct response is equal to 1. In the more realistic versions of this axiom used in practice, one additional parameter is used which specifies the probability of a *careless error*. It may also be necessary to introduce a *lucky guess* parameter, for example if a multiple choice paradigm is used.

8.7.3 Some Key Results. These results follow from Axioms [U], [Q] and [R]. (For proofs, see Chapter 13 in Falmagne and Doignon, 2011).

1. The updating operator specified by Equation (8.9) in Axiom [U] is essentially a Bayesian operator. This can be shown by an appropriate transformation of the equation (see 13.4.5 on page 251 in Falmagne and Doignon, 2011).

2. This updating operator is *permutable*. This term is used in the functional equations literature (Aczél, 1966) to designate a function F satisfying the equation
$$F(F(x,y),z) = F(F(x,z),y).$$
This property is essential because it means that the order of the questions has no import on the result of the assessment.
3. The stochastic process $(\mathbf{L}_n)$ is Markovian.
4. The stochastic process $(\mathbf{L}_n)$ converges to the latent state K_0 in the sense that
$$\mathbf{L}_n(K_0) \xrightarrow{\text{a.s.}} 1$$
(in which 'a.s.' means 'almost surely').

Remark. Another Markovian assessment procedure was also developed, which is based on a different principle (Falmagne and Doignon, 1988b and Chapter 14 in Falmagne and Doignon, 2011). The stochastic process is a finite Markov chain, the states of which are subsets of states of the learning space.

8.8 About Practical Implementations

The three axioms [U], [Q] and [R] are the foundation pieces of the assessment mechanism used by the `ALEKS` system. As mentioned in Chapter 1, however, a direct implementation of these axioms as an assessment software is not possible for two reasons.

One is that students commit careless errors. This means that Axiom [R] has to be modified by the introduction of a 'careless error parameter.' An obvious possibility is the axiom:

[R'] **Modified Response Rule.** For all positive integers n,

$$\mathbb{P}(\mathbf{R}_n = \iota_{K_0}(q) \,|\, \mathbf{Q}_n = q, \mathbf{L}_n, \mathbf{W}_{n-1}) = \begin{cases} 1 - \beta_q & \text{if } q \in K_0 \\ 0 & \text{if } q \notin K_0\,, \end{cases}$$

in which β_q is the probability of making an error in responding to the item q which is in the latent state K_0.

In some cases, 'lucky guess' parameters must also be used, for example to deal with the case of the multiple choice paradigm.

The second difficulty is that learning spaces formalizing actual curricula are always very large, with domains typically counting several hundred items. Such learning spaces may have many million states and cannot be searched in a straightforward manner. The adopted solution is to partition the domain Q of the basic learning space $\mathcal{K}$—the *parent learning space*—into some number N of subdomains $Q_1, \ldots, Q_N$. These subdomains are similar in that they contain approximately the same number of items and are representative of the domain, for example in such a way that each would be suitable as a placement test. Via Theorem 8.4.1, these N subdomains determine N projection learning spaces $\mathcal{K}^1, \ldots, \mathcal{K}^N$ of the parent learning space $\mathcal{K}$.

The general idea of the algorithm is to assess the student simultaneously on all the N projection learning spaces, with mutual updating of the probability distributions, according to the scheme outlined below.

8.8.1 The updating steps on trial n. We write $\mathcal{K}^j_q$ for the subfamily of $\mathcal{K}^j$ containing q, with $1 \leq j \leq N$; so, $q \in Q_j$. For $K \in \mathcal{K}^j$, we denote by $L^j_n(K)$ the probability of the state K in the learning space $\mathcal{K}^j$ on trial n.

1. On trial n, pick an item q minimizing $|2L^j_n(\mathcal{K}^j_q) - 1|$, for all $1 \leq j \leq N$ and $q \in Q_j$. If more than one item achieves such a minimization, pick randomly between them[12].
2. Suppose that the chosen item q belongs to Q_j. Record the student's response to item q. Update the probability distribution L^j according to Axiom [U] and Equation (8.9), (with $L_n = L^j_n$ and $L_{n+1} = L^j_{n+1}$).
3. Add item q to all the $N-1$ subdomains $Q_i \neq Q_j$, and write $Q^\star_i = Q_i \cup \{q\}$.
4. Build the $N-1$ projections $\mathcal{K}^{i\star}$ of $\mathcal{K}$ on the subdomains $Q^\star_i$. Note that $\mathcal{K}^{i\star}$ can in turn be projected on Q_i. For any state K in $\mathcal{K}^{i\star}$, let $[K]$ denote its equivalence class with respect to the projection on Q_i, and let $K_{|Q_i}$ denote its projection on Q_i.
   That is, $M \in [K]$ if $M \cap Q_i = K \cap Q_i$, and
   $$K_{|Q_i} = \begin{cases} K & \text{if } q \notin K, \\ K \setminus \{q\} & \text{if } q \in K\,. \end{cases}$$
5. For $1 \leq i \leq N$ and $i \neq j$, define the probability distributions $L^{i\star}_n$ on $\mathcal{K}^\star_i$ by the equation:
   $$L^{i\star}_n(K) = \frac{L^i_n(K_{|Q_i})}{|[K]|}. \tag{8.11}$$
   So, the rule $L^{i\star}_n(K)$ splits equally the probability $L^i_n(K)$ among the states created by the addition of q to Q_i.
6. Now, update all the $N-1$ probabilities $L^{i\star}_n$ according to Axiom [U] and Equation (8.9). For each learning space $\mathcal{K}^\star_i$, $1 \leq i \leq N$ and $i \neq j$, this results into a probability distribution $L^{i\star}_{n+1}$.
7. Remove item q from all the domains $Q_i \cup \{q\}$, $1 \leq i \leq N$ and $i \neq j$, and normalize all the $L^{i\star}_{n+1}$ probability distributions. That means: compute the probability distribution L^i_{n+1} on $\mathcal{K}^i$ from $L^{i\star}_{n+1}$ by the formula
   $$L^i_{n+1}(K) = \sum_{M \text{ s.t. } M_{|Q_i} = K} L^{i\star}_{n+1}(M).$$

[12] This rule is consistent with Equation (8.10).

The above scheme can be altered in two major ways. One concerns Step 3, where more than one question can be added to the subdomains. For instance, on trial n, one can add the n questions answered so far by the student to each subdomain, and then 'replay' the assessment on such extended subdomains. Another way concerns Step 5, where the rule $L_n^{i\star}(K)$ can split the probability $L_n^i(K)$ non-equally among the states created by the addition of q to Q_i.

8.8.2 The construction of the final state. The procedure described below for the construction of the final state is only one of several possibilities. Suppose that the assessment ended at trial n. For each item q, we define its likelihood $\omega(q)$ as

$$\omega(q) = \sum_{K \in \mathcal{K}_q^j} L_n^{i\star}(K), \qquad (q \in Q_j).$$

Let K_f denote the final knowledge state to be assigned to the student. This final state is built recursively. With $1 \leq i \leq |Q|$, let (q_i) denote the sequence of items ordered by decreasing value of $\omega(q_i)$. The procedure starts by setting $K_f = \varnothing$ and proceeds recursively along (q_i). On step i, the procedure examines each atom at q_i and add it to K_f if it passes the following simple criteria: for atom A at q_i

$$\frac{1}{|A \setminus K_{f,i-1}|} \sum_{q \in A \setminus K_{f,i-1}} w(q) \geq \delta,$$

where $K_{f,i-1}$ is the state of the final knowledge state after item q_{i-1} and $\delta, 0 < \delta < 1$, is a threshold parameter.

The procedure terminates when it reaches the end of the sequence (q_i). Since K_f is at all times either empty or the union of atoms, the procedure returns a knowledge state.