Journal of Mathematical Psychology 114 (2023) 102773

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Editorial

Special issue on knowledge structures: Theoretical developments and applications

In an seminal paper, Doignon and Falmagne (1985) introduced knowledge structures as a highly flexible set-theoretic framework for representing the organization of knowledge elements. Knowledge structures offer a precise non-numerical representation of the individual strengths and weaknesses in a particular domain. They provide the theoretical basis for implementing efficient algorithms for the assessment of knowledge, and for personalized learning. Many empirical applications of Knowledge Structure Theory (KST) demonstrate its usefulness, among them the commercially successful educational software ALEKS.[1] The literature on knowledge structures has grown massively in the past 37 years, and it is still expanding rapidly. This development is documented by a series of textbooks (Albert & Lukas, 1999; Doignon & Falmagne, 1999; Falmagne, Albert, Doble, Eppstein, & Hu, 2013; Falmagne & Doignon, 2011) which, however, leave a gap of almost ten years.

This special issue collects current contributions that elaborate and generalize the theory and demonstrate today's spectrum of topics that are tackled from a KST perspective. The compilation of nine papers covers work that extends the classical deterministic theory of knowledge structures as well as the class of probabilistic models that may be defined on it. Moreover, it addresses practical aspects of implementing KST on a large scale and its validation, and it paves the way to innovative applications beyond the educational setting.

Knowledge structures

The appeal of KST lies in the simplicity and generality of its basic notions. A domain of knowledge is characterized by a set of test items, and the knowledge state of an individual consists of the subset of items that the individual in principle masters. The collection of all the possible states then forms a knowledge structure. Due to dependencies between items, not all subsets of items will occur as knowledge states. Notice that mastering an item is to be distinguished from actually solving it, so that knowledge states are to be conceived as latent constructs. Emphasizing this distinction indicates that KST originally focused on item-related behavior, and cognitive concepts entered the stage only later on within a so-called skill- or competence-based extension: to each item the subset(s) of skills sufficient for mastering the item are assigned.

Knowledge structures satisfying additional properties have received particular attention. These structures include knowledge spaces (closed under union), and quasi ordinal knowledge spaces (closed under union and intersection), which were shown to be in one-to-one correspondence to quasi orders defined on the domain Doignon and Falmagne (1985). These so-called precedence relations describe dependencies between items. Another important class is formed by the learning spaces, which are knowledge spaces where learning (i.e., moving from one state in direction to the full domain) proceeds in steps of adding single items to the current state.

The so-called basic local independence model (BLIM; Doignon & Falmagne, 1999) is the standard probabilistic model in KST. It predicts the probability of any possible subset of correct responses based on a probability distribution on the states, and parameters capturing the probability of a lucky guess and a careless error, respectively, for each of the items. The characterization of the identifiability of the parameters in the BLIM has received considerable attention (e.g. Heller, 2017; Spoto, Stefanutti, & Vidotto, 2012; Stefanutti, Heller, Anselmi, & Robusto, 2012), showing that there are identifiability issues depending on structural properties of the underlying knowledge structure. For an overview see Doignon, Heller, and Stefanutti (2018). Besides the BLIM, various discrete- as well as continuous-time stochastic process models were suggested (see Falmagne & Doignon, 2011) in order to represent learning within the KST framework, and for uncovering the latent state by appropriate questioning in knowledge assessment.

Although quite concise, this introduction nevertheless supplies the interested reader with the core concepts and results from which the subsequently outlined contributions depart.

Deterministic theory

A series of four papers is devoted to further developing the deterministic theory of knowledge structures.

In his contribution, Suck (2021) inverts the traditional KST perspective of going from items to skills, and starts out from a situation where a partially ordered set of skills is given. The mathematical tool of a set representation of a partial order is then used to construct a knowledge space or a learning space on a set of items. This approach is based on the notion of the basis of a knowledge space, conceived as the collection of minimal states containing an item, from which the whole knowledge space can be generated by taking unions. The elements of the basis may

[1] Assessment and LEarning in Knowledge Spaces (https://www.aleks.com/).

https://doi.org/10.1016/j.jmp.2023.102773

quite naturally be interpreted as the skills necessary for mastering the respective item (Doignon, 1994). In the construction the given skills are then identified with the basis of an appropriately chosen set representation. Within this innovative perspective particular items (called multiple items) and partial orders (called fans) may serve as building blocks of the set representation.

Anselmi, Heller, Stefanutti, and Robusto (2022) also start from a set of skills in a certain domain. They intend to develop a test by collecting items such that the resulting test is as informative as possible concerning the assessment of the skills an individual has available. Different scenarios are covered, including the construction of a test from scratch as well as the improvement and shortening of an existing test. In this approach items are identified with subsets of skills under a conjunctive (all skills necessary for mastering the item) or a disjunctive (each of the skills sufficient for mastering the item) model. This lets one identify items that are redundant or missing in order to allow for a unique skill assessment, and thus may provide guidelines for constructing or revising tests.

The original conception of a knowledge structure as outlined above is formulated for dichotomous items which are either mastered or not mastered. Heller (2021) shows that the construction of Doignon and Falmagne (1985) establishing a one-to-one correspondence between precedence relations on (dichotomous) items and the quasi ordinal knowledge spaces can be generalized to apply to polytomous response formats, where response values are partially ordered so that they form a lattice. In this respect, and because the approach allows for item-specific response scales, it generalizes previous approaches (Schrepp, 1997; Stefanutti, Anselmi, de Chiusole, & Spoto, 2020), which can be characterized as special cases showing a kind of factorial structure. This generalization takes an important step for applying KST to psychological testing, since it not only allows for treating situations where partial credit is granted, but also admit application to personality tests which usually are compiled of Likert type polytomous items. The KST approach opens a new perspective to psychological testing as it avoids aggregating information across items (as, for example, in a sum score), and thus provides a detailed picture of the profile of a testee.

Putting these theoretical developments to practical use, however, requires to also generalize methods for building knowledge structures from data available for the dichotomous case. Schrepp (1999) introduced item tree analysis to KST to uncover a precedence relation between items given the observed responses, a method originally established for Boolean analysis of questionnaires (van Leeuwe, 1974), and refined later to inductive item tree analysis (Schrepp, 2003). In their paper, Ünlü and Schrepp (2021) extend this approach to polytomous items with nominal and ordinal response scales (which may also be item-specific), and demonstrate its empirical applicability by analyzing survey data.

Probabilistic models

The remaining five contributions to the special issue consider probabilistic models defined on knowledge structures.

Based on the same parameter space as the BLIM, Doignon (2021) considers what he calls the Correct Response Model, which predicts the probability of a correct response to any single item. The paper investigates this model with respect to testability, identifiability and characterizability. Mainly drawing upon the theory of polytopes (Grünbaum, 2003; Ziegler, 1998) it either provides explicit results or points out serious problems preventing a definitive answer.

In general, probabilistic knowledge structures impose no constraints on the probability distribution defined on the knowledge states. The number of parameters in these probabilistic models thus tends to be large, and may be drastically reduced if the distribution can be built from parameters that are linked to items rather than states. This is realized in the so-called Simple Learning Model (Falmagne, 1994), where the parameters refer to the probability of "learning an item". Its application, however, is limited to learning spaces. Noventa, Heller, and Stefanutti (2021) generalize this model by suggesting a method to build the state probabilities as products of the probabilities of single (or groups of) items on a much wider class of regular knowledge structures. This result allows for establishing more parsimonious probabilistic models to handle situations that may otherwise be intractable due to the combinatorial explosion of the number of states on increasingly large domains.

Anselmi, Stefanutti, de Chiusole, and Robusto (2021) model learning in knowledge structures through a bivariate Markov process (Ephraim & Mark, 2012), consisting of a pair of continuous-time stochastic processes that are jointly Markov. One of the processes is observable and one latent, which in the considered application corresponds to processes that refer to the navigation behavior in a web-based tutoring system and the associated learning process, respectively. Learning is represented by transitions among states in a competence structure (the analog of a knowledge structure with skills replacing items). This approach adds KST to the areas of successful application of bivariate Markov processes, and goes beyond previous modeling efforts by explicitly linking observable navigation behavior in a learning environment to the latent learning path.

The contribution of Stefanutti, de Chiusole, and Brancaccio (2021) bridges between the investigation of human problem solving and knowledge assessment, extending the application of KST beyond the educational setting. It builds upon the derivation of a learning space from a problem space (Newell & Simon, 1972) as suggested by Stefanutti (2019), and proposes a discrete-time Markov process modeling the solution process in a problem-solving task based on the underlying knowledge state of the problem solver. For empirical validation the developed theory is applied to data from experimental studies on the Tower of London test.

Cosyn, Uzun, Doble, and Matayoshi (2021) add a practical perspective to the special issue by considering the ALEKS educational software system, a heavily used large-scale implementation of KST. After highlighting the key theoretical concepts the system builds upon, they evaluate its probabilistic assessment and learning mode using standard and KST-based measures. The analysis relies on the data of millions of users. While the findings in general are viewed as validating the KST approach, they are interpreted as questioning the BLIM assumptions that the careless error and lucky guess probabilities do not depend on the knowledge state. The presented analysis reveals systematic deviations from constancy as a function of the so-called layer, which captures the difficulty of an item relative to a state (Doble, Matayoshi, Cosyn, Uzun, & Karami, 2019). This observation clearly stimulates further research in probabilistic KST models generalizing the BLIM.

Conclusions

The present special issue demonstrates that KST provides a general framework that, even after a history of almost four decades, offers plenty of potential for new theoretical developments as well as innovative applications. It paints a picture of KST as a many-faceted research strand drawing upon tools from different areas of mathematics (e.g., order theory, polyhedral combinatorics, probability theory and stochastic processes). The contributions document the current state of the art, showing the diversity of theory building in KST and highlighting some of its successful applications. In spite of the associated combinatorial complexity, these applications include large-scale implementations and cover areas beyond the originally intended representation and assessment of knowledge. KST can offer an alternative view in realms such as psychological testing and cognitive psychology, and may help opening new avenues for future research.

References

Albert, D., & Lukas, J. (1999). *Knowledge spaces: theories, empirical research, applications*. Mahwah, NJ: Lawrence Erlbaum Associates,

Anselmi, P., Heller, J., Stefanutti, L., & Robusto, E. (2022). Constructing, improving, and shortening tests for skill assessment. *Journal of Mathematical Psychology*, *106*, Article 102621.

Anselmi, P., Stefanutti, L., de Chiusole, D., & Robusto, E. (2021). Modeling learning in knowledge space theory through bivariate Markov processes. *Journal of Mathematical Psychology*, *103*, Article 102549.

Cosyn, E., Uzun, H., Doble, C., & Matayoshi, J. (2021). A practical perspective on knowledge space theory: ALEKS and its data. *Journal of Mathematical Psychology*, *101*, Article 102512.

Doble, C., Matayoshi, J., Cosyn, E., Uzun, H., & Karami, A. (2019). A data-based simulation study of reliability for an adaptive assessment based on knowledge space theory. *International Journal of Artificial Intelligence in Education*, *29*, 258–282.

Doignon, J.-P. (1994). Knowledge spaces and skill assignments. In G. Fischer, & D. Laming (Eds.), *Contributions to mathematical psychology, psychometrics and methodology* (pp. 111–121). New York: Springer-Verlag.

Doignon, J.-P. (2021). A correct response model in knowledge structure theory. *Journal of Mathematical Psychology*, *102*, Article 102519.

Doignon, J.-P., & Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies*, *23*, 175–196.

Doignon, J.-P., & Falmagne, J.-C. (1999). *Knowledge spaces*. Berlin, Heidelberg, and New York: Springer-Verlag.

Doignon, J.-P., Heller, J., & Stefanutti, L. (2018). Identifiability of probabilistic models, with examples from knowledge structure theory. In W. H. Batchelder, H. Colonius, & E. N. Dzhafarov (Eds.), *New handbook of mathematical psychology, volume 2, modeling and measurement* (pp. 128–184). Cambridge, UK: Cambridge University Press.

Ephraim, Y., & Mark, B. L. (2012). Bivariate Markov processes and their estimation. *Foundations and Trends in Signal Processing*, *6*(1), 1–95.

Falmagne, J.-C. (1994). Finite Markov learning models for knowledge structures. In G. Fischer, & D. Laming (Eds.), *Contributions to mathematical psychology, psychometrics, and methodology* (pp. 75–89). New York: Springer.

Falmagne, J.-C., Albert, D., Doble, C., Eppstein, D., & Hu, X. (Eds.), (2013). *Knowledge spaces: applications in education*. Berlin, Heidelberg: Springer.

Falmagne, J.-C., & Doignon, J.-P. (2011). *Learning spaces: interdisciplinary applied mathematics*. Berlin, Heidelberg: Springer.

Grünbaum, B. (2003). Polytopes. In V. Kaibel, V. Klee, & G. M. Ziegler (Eds.), *Volume 221 of graduate texts in mathematics*, *Convex polytopes* (2nd ed.). (pp. 35–60). New York, NY: Springer.

Heller, J. (2017). Identifiability in probabilistic knowledge structures. *Journal of Mathematical Psychology*, *77*, 46–57.

Heller, J. (2021). Generalizing quasi-ordinal knowledge spaces to polytomous items. *Journal of Mathematical Psychology*, *101*, Article 102515.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Noventa, S., Heller, J., & Stefanutti, L. (2021). Some considerations on the factorization of state probabilities in knowledge structures. *Journal of Mathematical Psychology*, *102*, Article 102542.

Schrepp, M. (1997). A generalization of knowledge space theory to problems with more than two answer alternatives. *Journal of Mathematical Psychology*, *41*, 237–243.

Schrepp, M. (1999). On the empirical construction of implications between bi-valued test items. *Mathematical Social Sciences*, *38*, 361–375.

Schrepp, M. (2003). A method for the analysis of hierarchical dependencies between items of a questionnaire. *Methods of Psychological Research-Online*, *19*, 43–79.

Spoto, A., Stefanutti, L., & Vidotto, G. (2012). On the unidentifiability of a certain class of skill multi map based probabilistic knowledge structures. *Journal of Mathematical Psychology*, *56*, 248–255.

Stefanutti, L. (2019). On the assessment of procedural knowledge: From problem spaces to knowledge spaces. *British Journal of Mathematical and Statistical Psychology*, *72*(2), 185–218.

Stefanutti, L., Anselmi, P., de Chiusole, D., & Spoto, A. (2020). On the polytomous generalization of knowledge space theory. *Journal of Mathematical Psychology*, *94*, Article 102306.

Stefanutti, L., de Chiusole, D., & Brancaccio, A. (2021). Markov solution processes: Modeling human problem solving with procedural knowledge space theory. *Journal of Mathematical Psychology*, *103*, Article 102552.

Stefanutti, L., Heller, J., Anselmi, P., & Robusto, E. (2012). Assessing local identifiability of probabilistic knowledge structures. *Behavior Research Methods*, *44*, 1197–1211.

Suck, R. (2021). Skills first — An alternative approach to construct knowledge spaces. *Journal of Mathematical Psychology*, *101*, Article 102517.

Ünlü, A., & Schrepp, M. (2021). Generalized inductive item tree analysis. *Journal of Mathematical Psychology*, *103*, Article 102547.

van Leeuwe, J. (1974). Item tree analysis. *Nederlands Tijdschrift Voor de Psychologie*, *29*, 475–484.

Ziegler, G. M. (1998). *Lectures on polytopes (revised ed.)*. Berlin: Springer.

Jürgen Heller
Department of Psychology, University of Tübingen, Schleichstr. 4, 72076 Tübingen, Germany
E-mail address: juergen.heller@uni-tuebingen.de.

Available online 3 April 2023