Techniques for Sampling Quasi-orders

Ali Ünlü and Martin Schrepp

Abstract In educational theories, e.g., learning spaces, mastery dependencies between test items are represented as reflexive and transitive binary relations, i.e., quasi-orders, on the item set of a knowledge domain. Item dependencies can be used for efficient adaptive knowledge assessment and derived through exploratory data analysis, for example by algorithms of item tree analysis. To compare item tree analysis methods, typically large-scale simulation studies are employed, with samples of randomly generated quasi-orders at their basis and assumed to underlie the data. In this context, a serious problem is the fact that all of the algorithms are sensitive to the underlying quasi-order structure. Thus, it is crucial to base any simulation study that aims at comparing the algorithms in a reliable manner on representative samples, meaning that each quasi-order in the population is equally likely to be selected as part of a sample. Suboptimal sampling strategies were considered in previous studies leading to biased conclusions. In this paper, we discuss sampling techniques that allow us to generate representative, or close to representative, random quasi-orders. The

Ali Ünlü
Technical University of Munich (TUM), Centre for International Student Assessment (ZIB), Arcisstrasse 21, 80333 Munich, Germany
✉ ali.uenlue@tum.de

Martin Schrepp
SAP AG, Walldorf, Germany
✉ martin.schrepp@sap.com

ARCHIVES OF DATA SCIENCE (ONLINE FIRST) DOI 10.5445/KSP/1000058749/03
KIT SCIENTIFIC PUBLISHING ISSN 2363-9881
Vol. 2, No. 1, 2017

item tree analysis methods are compared on ten items with a representative, large sample of quasi-orders, thereby supporting their invariant ordering.

1 Introduction

Hierarchies among items represented by reflexive, transitive binary relations, i.e., *quasi-orders*, on item sets are an important concept of *knowledge* or *learning space theory* (Doignon and Falmagne, 1999; Falmagne and Doignon, 2011). As a model of the mastery implications between the questions or problems of a psychological or educational test, quasi-orders can be used to design efficient adaptive, computer-based knowledge assessment and training procedures. The *Assessment and Learning in Knowledge Spaces* (ALEKS) system[1] is a fully automated math tutor on the internet (e.g., see Falmagne et al, 2013).

The basic idea of this approach is that knowledge in many domains, e.g., mathematics, is organized in a hierarchical structure. Some pieces of knowledge can only be learned successfully, if other more basic pieces of knowledge are already mastered. In this sense, pieces of knowledge, and hence the items or problems used to measure them, may imply one another and be organized in a reflexive and transitive relation, a quasi-order, which is also called a *surmise relation* in knowledge or learning space theory. The following example with a test of six elementary algebra problems shown in Fig. 1 may help to illustrate the main idea.

Item tree analysis (ITA) is a data mining technique that tries to extract logical implications between binary variables (Schrepp, 1999, 2003; Sargin and Ünlü, 2009). The goal of ITA is to reconstruct, by exploratory analysis of a collection of observed noisy response patterns due to careless errors or lucky guesses, the underlying or true quasi-order structure of the between-item implications in a knowledge domain under reference. The general scheme of ITA-type methods is as follows (for details, see Schrepp, 1999; Sargin and Ünlü, 2009):

1. Generating the selection set of competing quasi-orders.
2. Using some criterion for assessing the fit of each candidate quasi-order to the data.
3. Choosing the "best" fitting quasi-order.

[1] `www.aleks.com`

a. A car travels on the freeway at an average speed of 52 miles per hour. How many miles does it travel in 5 hours 30 minutes?
b. Using the pencil, mark the point at the coordinates $(1,3)$.
c. Perform the multiplication $4x^4y^4 \cdot 2x \cdot 5y^2$ and simplify as much as possible.
d. Find the greatest common factor of $14t^6y$ and $4tu^5y^8$. Simplify as much as possible.
e. Graph the line with slope -7 passing through $(-3,-2)$.
f. Write an equation for the line that passes through the point $(-5,3)$ and is perpendicular to the line $8x+5y=11$.

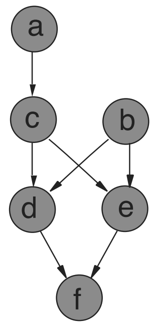

Fig. 1 A plausible surmise relation on the domain $\{a,\ldots,f\}$. For instance, the mastery of problem e implies the mastery of problem b.

Inductive ITA (Schrepp, 1999; Sargin and Ünlü, 2009) was developed as an improvement of an initial algorithm by Van Leeuwe (1974). Software for the ITA methods are the Windows program `ITA 2.0`[2] (Schrepp, 2006) and the platform independent `R` package `DAKS`[3] (Ünlü and Sargin, 2010).

Why are representative random quasi-orders important? Large-scale simulation studies have been employed to evaluate and compare the quality of ITA-type item hierarchy mining techniques for recovering the posited true quasi-orders. In a recent work by Ünlü and Schrepp (2015b), it was seen that, in order to design principled simulation studies for reliable and sound comparisons, the essential and difficult task is to draw unbiased samples of quasi-orders for the simulations. The samples taken must be *representative* for the population or the set of all possible reflexive and transitive binary relations, in the sense that *each drawn quasi-order is contained in a sample with the same probability*. In particular, Ünlü and Schrepp (2015b) found that using non-representative quasi-order samples yielded biased or erroneous simulation results regarding the recovery and coverage qualities with the ITA algorithms.

This can be explained by the fact that there is the dependency of the ITA algorithms on the discrete structure of the underlying quasi-order (Ünlü and Schrepp, 2015a,b). Such dependencies result from a complex interplay between

[2] `www.jstatsoft.org/v16/i10`
[3] `www.jstatsoft.org/v37/i02`

the structure of the quasi-order and the response error probabilities. For some structural types of orders it may be easier to detect the correct implications than for others. Representative quasi-order or surmise relation samples will produce the least biased results when generalizing the findings obtained from simulation studies to the population of all quasi-orders.

2 Random Quasi-order Generation: The Problem

The problem addressed can be described as follows. Presuppose a set Q of n test items. We aim at randomly creating, using a random process, a sample of quasi-orders, which is *representative*, in the sense that each quasi-order on Q has the same probability to be included in this sample.

Definition 1. Let the set of all quasi-orders on Q be denoted by $\mathscr{QO}(n)$. Let $X \subset \mathscr{QO}(n)$ be a non-empty, randomly generated subset of quasi-orders on Q. We call X a *representative sample* if $\mathbb{P}(\leq_1 \in X) = \mathbb{P}(\leq_2 \in X)$ for any two quasi-orders $\leq_1$ and $\leq_2$ in $\mathscr{QO}(n)$.

In particular, such a sample X is "proportionally representative" in terms of the distributions of quasi-order size, quasi-order width, and quasi-order height in the population. These are the evaluation criteria that we will use in this paper to assess representativeness.

The following remarks are pertinent to the discussion and were motivated by critical comments made by a reviewer.

1. The problem of sampling quasi-orders can be put in the broader context of the random generation of complex combinatorial or discrete-mathematical structures (e.g., Beckenbach, 1964; Harary and Palmer, 1973; Nijenhuis and Wilf, 1978; Roberts and Tesman, 2009). There, the primary focus has been on the enumeration or counting, than actually the representative construction, for example, of various kinds of graphs. Literature in the field that were addressing the generation of combinatorial structures uniformly at random studied this for structures other than the quasi-orders, and for the unlabeled case. For example, Dixon and Wilf (1983) and Kerber et al (1990) introduced an algorithm for the uniform generation or selection of unlabeled general graphs. This algorithm cannot be directly applied in the present context. In fact, it seems that we have contributed with this paper to the theory of graphs. Quasi-orders correspond to transitive directed graphs (or finite topologies),

and it is their counts for different numbers of points that has been determined in the literature only, rather than developing an algorithm for uniformly, or representatively, constructing them, as we have done with the present work. The publications probably closest to the subject of quasi-orders are the works by Brinkmann and McKay (2002, 2005) and Pfeiffer (2004), who "merely" counted quasi-orders or deterministically constructed, with no randomization considered, unlabeled partial orders. This is not directly applicable, too.

However, from the foregoing we can conclude with an interesting direction for future research. Labeled quasi-order sampling appears to be much easier to accomplish than the corresponding unlabeled problem. In the unlabeled case, there is more symmetry and the isomorphism problem involved. Thus, research is needed to explore to what extent the published approaches in the context of the random generation of combinatorial or discrete-mathematical structures can help in developing techniques for the representative sampling of quasi-orders in the unlabeled case.

2. Representative samples are essential in statistics. Based on the well-known statistical theory of survey sampling (e.g., Cochran, 1977; Pedigo and Buntin, 1994; De Leeuw et al, 2008; Gideon, 2012; Thompson, 2012), myriads of theoretical and applied results have been published in numerous problem areas, including genetic programming, resampling methods, experimental design, social surveys, and educational large-scale assessment studies.

   In spite of this repertoire of widely-used statistical sampling techniques, it is worth noting that the problem of constructing representative samples of axiomatically-defined combinatorial objects cannot be solved by simple application of these tools. A general framework for a principled sampling theory for mathematical structures such as the quasi-orders appears to be different from the common statistical theory of survey sampling. Sampling mathematical structures typically includes preparatory combinatorial work and thus necessitates the provision of this paper's type of algorithm. For example, approaches similar to the simple random and stratified sampling, the two basic statistical techniques used in surveys, have been practically not feasible or lacking in the context of sampling quasi-orders. In this paper, we will discuss one possible algorithm for the uniform selection of labeled quasi-orders (Sect. 3). The idea underlying this technique may also be applied when samples of other ordered structures are needed. In such cases, many "random" generation processes generally exhibit non-linear biases.

2.1 Representative, But Infeasible, Direct Sampling Methods

Though theoretically implying representative samples, from a practical point of view, the following methods are not feasible in realistic settings for higher item numbers.

Census-like Uniform Sampling: The direct method for drawing a representative sample of quasi-orders on a set of n items is by creating all quasi-orders, storing them, and then uniformly sampling from them.
This only works for small n, since the number of quasi-orders increases very rapidly with n. On $n = 6$ items, we have $209,527$ (labeled) quasi-orders. This number increases to $9,535,241$ / $642,779,354$ / $63,260,289,423$ / $8,977,053,873,043$ for $n = 7$ / 8 / 9 / 10 items, respectively.

Entry-wise Uniform Sampling: All off-diagonal entries of the *relational matrix* are randomly filled, and the diagonal entries are set $r_{ii} := 1$ for all i. Exemplified with $n = 3$ items, the respective entries are marked:

$$\begin{bmatrix} r_{11} := 1 & r_{12} & r_{13} \\ r_{21} & r_{22} := 1 & r_{23} \\ r_{31} & r_{32} & r_{33} := 1 \end{bmatrix}.$$

Here $r_{12}, r_{13}, r_{21}, r_{23}, r_{31}, r_{32} \sim_{iid} Bernoulli(1/2)$, where $Bernoulli(1/2)$ is the Bernoulli distribution with success (i.e., $r_{ij} = 1$) probability $p = 1/2$, and *iid* stands for "independent and identically distributed." The resulting random reflexive binary relation is retained if it satisfies transitivity.
This procedure as well becomes infeasible in n. The proportions of quasi-orders among all reflexive relations very rapidly decrease with increasing item numbers n: there are $2^{n(n-1)}$ reflexive relations, and for $6 \leq n \leq 10$, the proportions of transitive and reflexive relations are $1.95 \cdot 10^{-4}$, $2.17 \cdot 10^{-6}$, $8.92 \cdot 10^{-9}$, $1.34 \cdot 10^{-11}$, and $7.25 \cdot 10^{-15}$, respectively.

2.2 Flexible, But Non-representative, Ad-hoc Sampling Strategies

Previous studies tried to avoid this problem by implementing *ad-hoc* procedures. These procedures are flexible and fast to compute. However, they generally lack representativeness of the created sets of random quasi-orders.

Normal Parametric Sampling: In Sargin and Ünlü (2009) the following procedure was used:

1. The process starts with the diagonal that contains exactly the reflexive item pairs (i,i) for $i = 1,\dots,n$.
2. All other pairs are added to the diagonal with a probability δ, thereby yielding a relation $\mathscr{R}$. The probability δ itself is drawn randomly from a normal distribution with $\mu = 0.16$ and $\sigma = 0.06$. Values less than 0 or greater than 0.3 are set to 0 or 0.3, respectively.
3. The transitive closure of $\mathscr{R}$ is the resulting random quasi-order.

Uniform Parametric Sampling: The above random process is an improvement of a previous procedure that drew δ based on a uniform distribution from the interval 0 to 0.4, or 0 to 1, which yielded non-representative samples with overly represented large quasi-orders (see Sargin and Ünlü, 2009).

An appreciation of the difference in the representativeness of the quasi-order samples can be obtained from Fig. 2. Histogram and kernel densities show the lack of representativeness of the normal parametric sampling procedure, in contrast to a true simple random sample drawn according to the census-like uniform sampling method.

2.3 Biased Conclusions Due To Suboptimal Sampling

This difference in the representativeness of the sampling methods has a negative impact in the comparison of the ITA algorithms (Ünlü and Schrepp, 2015b). It biases the findings. For example, Figs. 3 and 4 show the *dist* values (symmetric differences between the derived and true quasi-orders, i.e., the numbers of item pairs that differ) used to compare and evaluate the ITA algorithms concerning their qualities to recover the true quasi-orders underlying the simulations.

If a representative mechanism for sampling the quasi-orders such as the census-like uniform sampling is employed, consistently the same ranking of the ITA algorithms can be observed. This ranking of the ITA algorithms may be deemed invariant, where in any of the panels in Fig. 4, the three curves are ordered increasingly from green, to red, and to blue, along the entire x-axis of error probabilities. Thus, the minimized corrected algorithm performs slightly better than the corrected, and the corrected algorithm significantly improves on the original procedure. That ranking among the ITA methods may also be

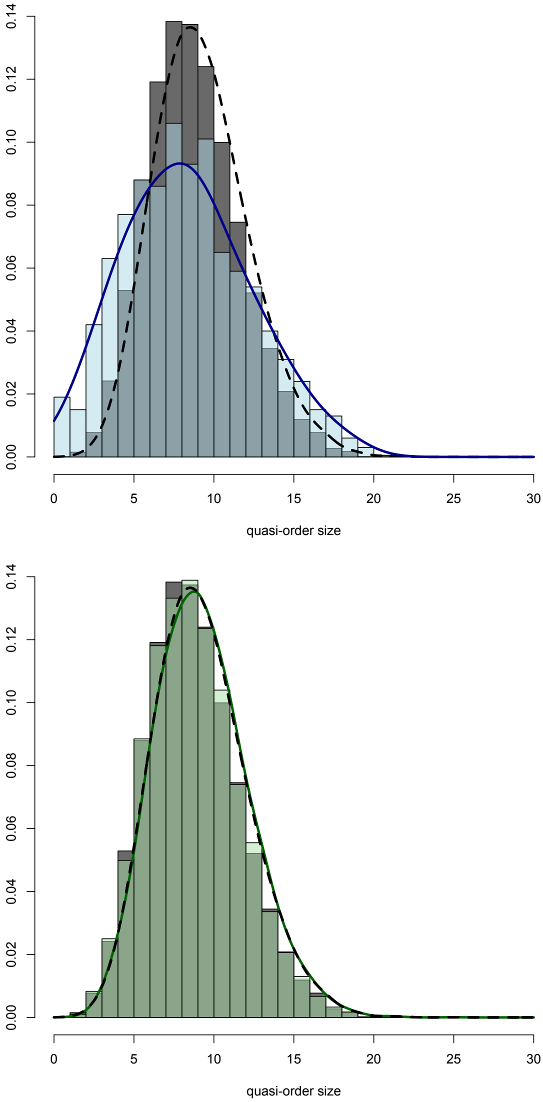

Fig. 2 *Top panel:* For six items, histogram and kernel density estimate (in blue) of quasi-order sizes for $1,000$ random δ values sampled according to the normal parametric method. In gray, the true distribution is shown. Averages over 100 quasi-orders were generated for each of the $1,000$ δ values and are plotted. Obtained quasi-order sizes deviate from the sizes expected for a representative sample. *Bottom panel:* Histogram and kernel density estimate (in green) of quasi-order sizes for $10,000$ quasi-orders drawn without replacement and with equal probability from the set of all $209,527$ quasi-orders based on six items. This random sample taken according to the census-like uniform sampling matches the true distribution (in gray). However, it is obtained with extra cost of computation and storage, which becomes infeasible for larger number of items.

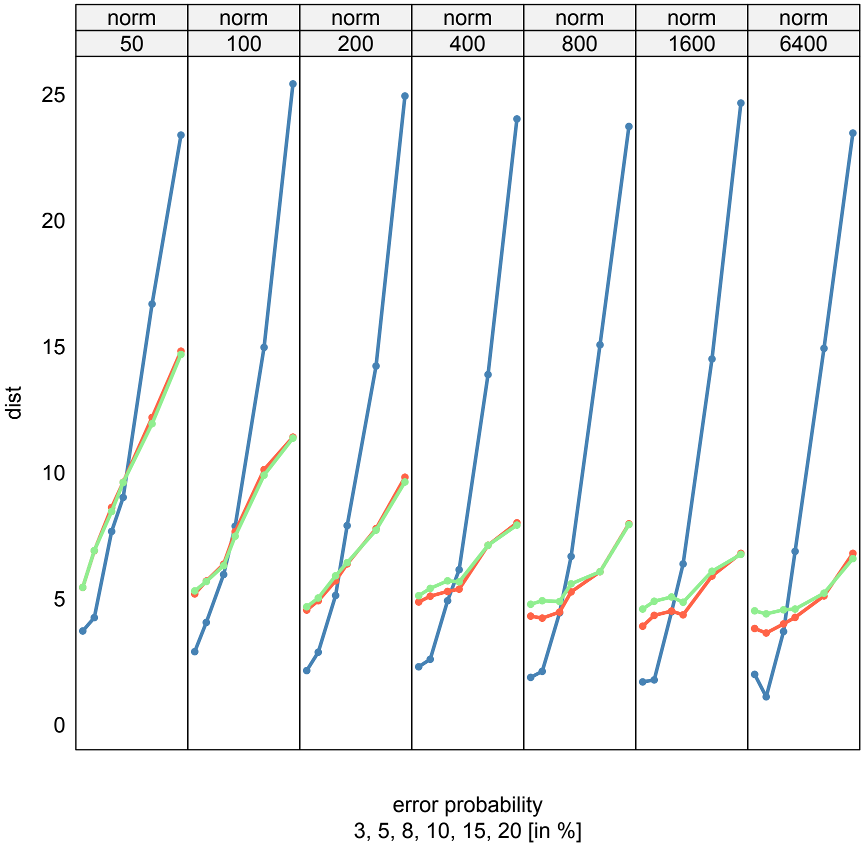

Fig. 3 Average *dist* values under the three existing original (blue), corrected (red), and minimized corrected (green) ITA algorithms, as a function of error probability $\{0.03; 0.05; 0.08; 0.10; 0.15; 0.20\}$ conditioned on sample size $\{50; 100; 200; 400; 800; 1,600; 6,400\}$. Panels "norm" are for nine items with normal parametric sampling. For each combination of error probability and sample size, averages over $1,000$ random quasi-orders drawn according to the normal parametric method are reported.

considered to be plausible theoretically. According to Sargin and Ünlü (2009), the corrected algorithm was introduced as estimation-correcting improvement on the original, and the minimized corrected algorithm as a further-optimized variant based on the corrected estimators.

In contrast, with a suboptimal sampling strategy, like the normal parametric technique, ordering reversals or discrepancies between the ITA methods prevail,

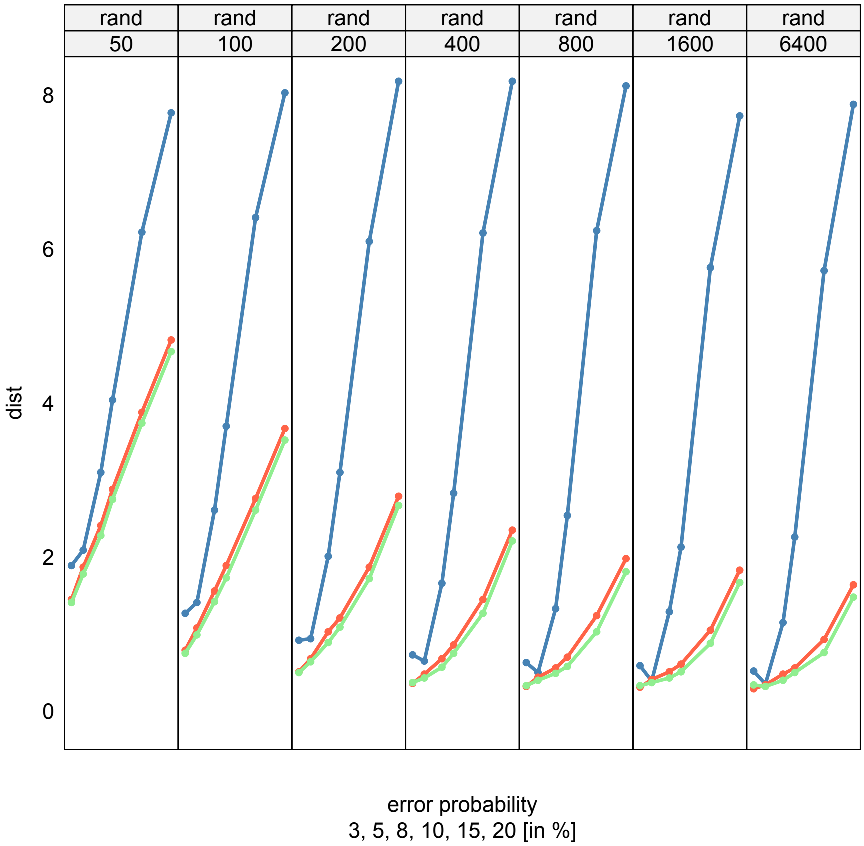

Fig. 4 Average *dist* values under the three existing original (blue), corrected (red), and minimized corrected (green) ITA algorithms, as a function of error probability $\{0.03; 0.05; 0.08; 0.10; 0.15; 0.20\}$ conditioned on sample size $\{50; 100; 200; 400; 800; 1,600; 6,400\}$. Panels "rand" are for six items with census-like uniform sampling. For each combination of error probability and sample size, averages over $1,000$ random quasi-orders drawn according to the census-like uniform method are reported.

thereby biasing the expected invariant algorithm ordering. For example, based on non-representative quasi-order samples, the curves intersect in the "norm" panels in Fig. 3, where with small error probabilities (e.g., 3% and 5%) the blue curve for the original algorithm lies consistently below the red and green curves for the corrected and minimized corrected algorithms, respectively. In addition, the "norm" panels for the large sample sizes (e.g., $1,600$ and $6,400$)

show that the corrected algorithm yielded lower *dist* values compared with the minimized variant.

3 Inductive Uniform Extension Technique

The sampling technique discussed in this section is *inductive* (Schrepp and Ünlü, 2015). It describes a procedure for generating a representative set of quasi-orders on $n+1$ items, on the basis of prior constructed quasi-orders on n items. The technique relies on the notion of a random reflexive extension.

Given a *trace quasi-order* R on the items $1,\ldots,l$, we construct a *random reflexive extension* of it, on the items $1,\ldots,l,l+1$, which extends the relational matrix r_R of R with a new $(l+1)$th row and $(l+1)$th column, retaining the original values of r_R otherwise, and which fills these new entries randomly. Exemplified with $l=2$, for a *trace adjacency matrix* r_R of dimension 2×2, the randomly filled new entries are marked:

$$\begin{bmatrix} \multicolumn{2}{c}{\multirow{2}{*}{r_R}} & \colorbox{lightgray}{r'_{13}} \\ & & \colorbox{lightgray}{r'_{23}} \\ \colorbox{lightgray}{r'_{31}} & \colorbox{lightgray}{r'_{32}} & r'_{33} := 1 \end{bmatrix}_{(l+1)\times(l+1)}$$

and $r'_{13}, r'_{23}, r'_{31}, r'_{32} \sim_{iid} Bernoulli(1/2)$.

Definition 2. The uniform extension approach is defined inductively.

Anchoring: The inductive procedure starts with a representative sample of quasi-orders, including the case of the complete inventory of all possible quasi-orders, on a sufficiently small number of items, say l.

Inductive Step: Presuppose a representative sample of quasi-orders on $n\geq l$ items, $QO(n)\subset \mathcal{QO}(n)$. For each quasi-order in $QO(n)$, we compute a pre-specified number z of random reflexive extensions. Non-transitive random extensions are excluded, and the transitive extensions are added to a new collection of quasi-orders on $n+1$ items, $QO(n+1)\subset \mathcal{QO}(n+1)$.

For example, the inductive procedure can be anchored by using the set of all 355 quasi-orders on four items, or with a simple random sample of $1,000$ quasi-orders drawn from all of the $209,527$ quasi-orders for six items. Optional modifications are possible. The duplicates can be removed from $QO(n+1)$

depending on the envisaged application, or $QO(n+1)$ may be reduced to a simple random sample of feasibly limited size, when this inductive construction is repeated several times—e.g., starting with the set of all four quasi-orders on two items, to create a representative sample consisting of $1,000$ quasi-orders on ten items.

It can be proved (Schrepp and Ünlü, 2015) that this method yields representative samples of random quasi-orders according to Definition 1. More precisely, if $\leq_1$ and $\leq_2$ in $\mathscr{QO}(n+1)$ are any two quasi-orders on $Q = \{1,\ldots,n+1\}$, and $X = QO(n+1)$ is generated according to the proposed inductive procedure, then $\mathbb{P}(\leq_1 \in X) = \mathbb{P}(\leq_2 \in X)$.

4 Results

To evaluate the quality of the inductive sampling technique, we investigated how representative the created quasi-order samples are on a set of six items. We also compared the ITA procedures based on a representative sample of $10,000$ random quasi-orders on ten items drawn utilizing this sampling technique.

4.1 Quality of Representativeness

Throughout the simulation study six items were used.

1. We generated 100 random samples (or trials) each of sample sizes of 100, $1,000$, and $5,000$ quasi-orders for six items, which we compared with the set of all $209,527$ quasi-orders possible on six items.
2. Beside the *size* or *cardinality* (i.e., the number of item pairs in relation) of a quasi-order $\leq$, the following properties were also used as the criteria for representativeness of the samples:

   width, the size of a longest anti-chain in $\leq$, and
   height, the size of a longest chain in $\leq$.

3. We started sampling with the set of all four quasi-orders on two items, i.e., the anchoring was with $l = 2$. The inductive step was employed four times, thereby yielding random quasi-orders on $n = 6$ items.
4. The computations were made with a C program on a computer with an Intel Core i5 2.50 GHz processor.

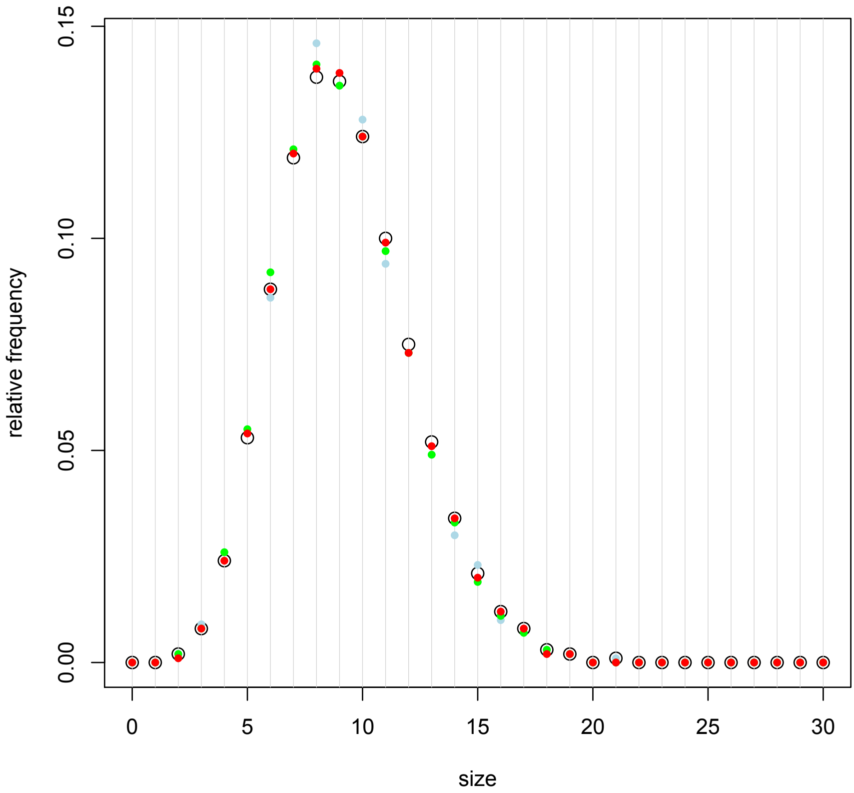

Fig. 5 The proportions of size computed in the population of all quasi-orders (unfilled black circle) compared with the means of the proportions of size computed over 100 trials in each of the samples of 100, $1,000$, and $5,000$ quasi-orders (filled blue, green, and red circles, respectively).

The estimated and true distributions of the sizes, widths, and heights are shown in Figs. 5, 6, and 7, respectively. The Kolmogorov distances of the estimated and true distributions are summarized in Table 1. We see that the simulation results approximate the true values very well, especially with larger sample sizes.

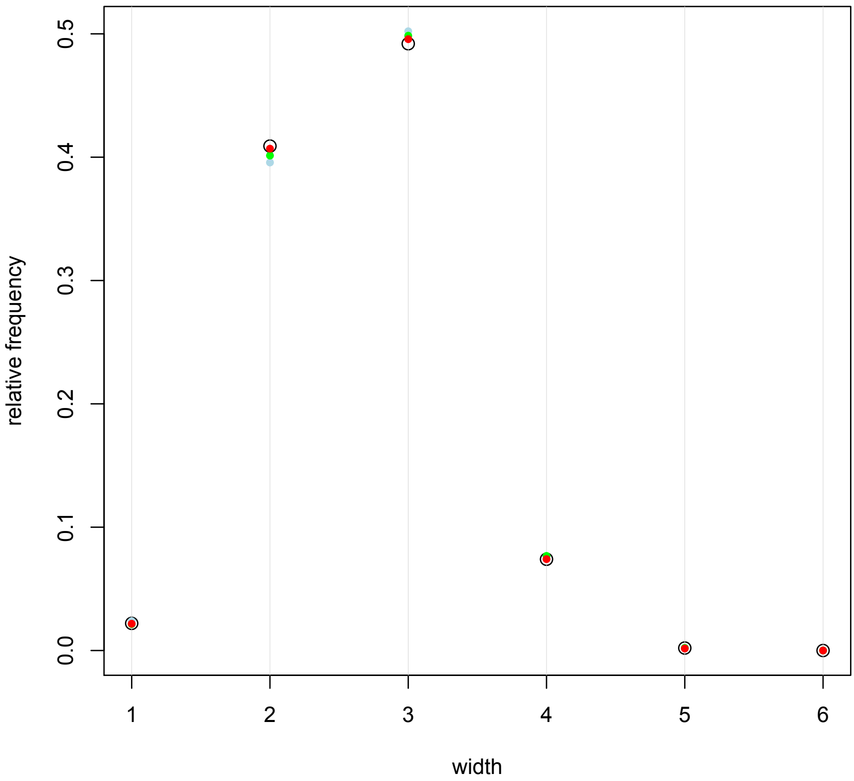

Fig. 6 The proportions of width computed in the population of all quasi-orders (unfilled black circle) compared with the means of the proportions of width computed over 100 trials in each of the samples of 100, 1,000, and 5,000 quasi-orders (filled blue, green, and red circles, respectively).

Table 1 Kolmogorov distances of the estimated and true distributions

	Sample size (quasi-orders)		
Criterion	100	1,000	5,000
Size	0.015	0.014	0.006
Width	0.012	0.009	0.003
Height	0.013	0.011	0.005

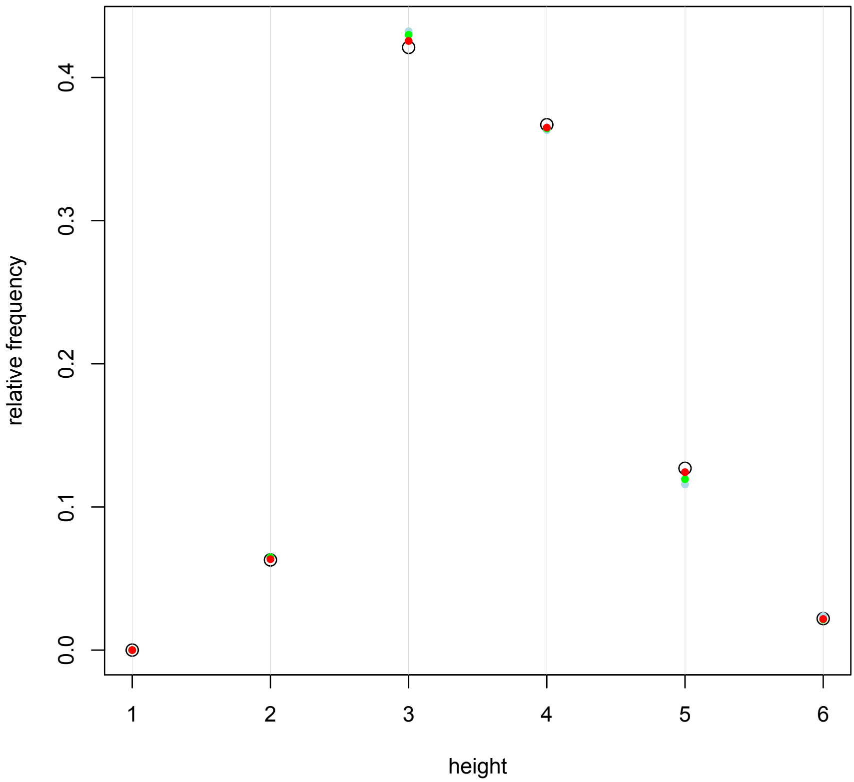

Fig. 7 The proportions of height computed in the population of all quasi-orders (unfilled black circle) compared with the means of the proportions of height computed over 100 trials in each of the samples of 100, $1,000$, and $5,000$ quasi-orders (filled blue, green, and red circles, respectively).

4.2 Representative Comparison of ITA Algorithms

Running the inductive uniform extension technique, a representative sample consisting of $10,000$ random quasi-orders was drawn from the population of all possible $8,977,053,873,043$ quasi-orders on ten items. Based on this sample of postulated true mastery hierarchies among the ten questions (cf. Sect. 1), the original and corrected ITA algorithms (Schrepp, 1999; Sargin and Ünlü,

Table 2 Average *dist* values obtained under the original (first line) and corrected (second line) ITA algorithms, based on a representative sample of $10,000$ quasi-orders for ten items with inductive uniform extension sampling

	Sample size (response patterns)		
τ	100	$1,000$	$5,000$
0.01	6.09	5.12	5.20
	3.22	1.43	1.30
0.05	7.51	5.82	5.46
	5.26	2.09	1.81
0.10	17.29	20.31	20.17
	8.24	4.13	3.61
0.15	21.58	22.80	22.88
	11.32	7.69	7.27

2009) were compared regarding their performance when reconstructing the underlying relational dependencies between the items from simulated noisy response data. Currently, for such a high number of items, the ITA algorithms have not been compared based on a representative sample, in a reliable manner.

Table 2 reports the average symmetric differences, i.e., the means of the $10,000$ *dist* values between the underlying quasi-orders of the representative sample and the ITA data analysis solutions (Sect. 2.3). The response error probability τ of the data generating latent variable model was specified with values of 0.01, 0.05, 0.10, and 0.15. The sample size varied among 100, $1,000$, and $5,000$ response patterns.

We can see from Table 2 the invariant ordering of the ITA methods that was mentioned in Sect. 2.3. Based on a representative mechanism for sampling the quasi-orders, consistently over all simulation settings the same ranking of the ITA algorithms can be observed. The corrected algorithm significantly improves on the original procedure. As compared to a suboptimal sampling strategy, in this study using representative random quasi-orders on ten items no ordering reversals between the two ITA methods do prevail.

Furthermore, in contrast to the original ITA algorithm, the modifications made with the corrected procedure yield solutions that improve constantly with increasing sample size. If we compute the quantity $(dist_{\tau,n_1} - dist_{\tau,n_2})/dist_{\tau,n_1}$, where $n_1 < n_2$, as the percentage change in *dist* with increasing sample size, for the corrected ITA method, this measure attains solely positive values, with the median of 0.43. For the original algorithm, the measure does take negative values too, and the median of the changes with increasing sample sizes is 0.00. Thus, with the corrected ITA method, there is a *relative reduction* of the mean *dist* values, approximately at an average of more than 40%, within a range of response error probabilities as more data become available.

The mean *dist* value computed over the twelve simulation conditions may be used as a gross location measure for the average number of misclassified item pairs, which is 4.78 for the corrected algorithm, and 13.35 for the original. In fact, the simulation results indicate "asymptotic convergence" and suggest that the corrected algorithm may reconstruct the true quasi-orders sufficiently well, "asymptotically" in increasing sample size or with decreasing random errors.

5 Conclusion

Representativeness of random quasi-orders drawn and postulated to underlie item hierarchy mining related simulation studies is an important requirement for reliable comparison of ITA-type data analysis methods. In this paper, we have recapitulated techniques for sampling quasi-orders on psychological or educational item sets. In particular, we have discussed the inductive uniform extension approach that allows creating representative samples of quasi-orders even for higher item numbers. The invariant ordering of the ITA algorithms has been supported on ten items with a representative, large sample of quasi-orders.

In further research, representative quasi-order samples can be employed to investigate the properties of ITA, independent of specific sample characteristics. An important research direction is to develop procedures for the construction of all quasi-orders on an item set, which could be made probabilistic to yield a process for sampling representative, or close to representative, quasi-orders.

References

Beckenbach EF (ed) (1964) Applied combinatorial mathematics. Wiley, New York

Brinkmann G, McKay BD (2002) Posets on up to 16 points. Order 19:147–179, DOI 10.1023/A:1016543307592

Brinkmann G, McKay BD (2005) Counting unlabelled topologies and transitive relations. Journal of Integer Sequences 8(Article 05.2.1):1–7

Cochran WG (1977) Sampling techniques. Wiley, New York

De Leeuw ED, Hox JJ, Dillman DA (eds) (2008) International handbook of survey methodology. Lawrence Erlbaum, New York

Dixon JD, Wilf HS (1983) The random selection of unlabeled graphs. Journal of Algorithms 4:205–213, DOI 10.1016/0196-6774(83)90021-4

Doignon JP, Falmagne JC (1999) Knowledge spaces. Springer, Berlin, DOI 10.1007/978-3-642-58625-5

Falmagne JC, Doignon JP (2011) Learning spaces. Springer, Berlin, DOI 10.1007/978-3-642-01039-2

Falmagne JC, Albert D, Doble C, Eppstein D, Hu X (eds) (2013) Knowledge spaces: Applications in education. Springer, Heidelberg, DOI 10.1007/978-3-642-35329-1

Gideon L (ed) (2012) Handbook of survey methodology for the social sciences. Springer, New York, DOI 10.1007/978-1-4614-3876-2

Harary F, Palmer EM (1973) Graphical enumeration. Academic Press, New York

Kerber A, Laue R, Hager R, Weber W (1990) Cataloging graphs by generating them uniformly at random. Journal of Graph Theory 14:559–563, DOI 10.1002/jgt.3190140507

Nijenhuis A, Wilf HS (1978) Combinatorial algorithms. Academic Press, New York

Pedigo LP, Buntin GD (eds) (1994) Handbook of sampling methods for arthropods in agriculture. CRC Press, Boca Raton

Pfeiffer G (2004) Counting transitive relations. Journal of Integer Sequences 7(Article 04.3.2):1–11

Roberts FS, Tesman B (2009) Applied combinatorics. Chapman & Hall/CRC, Boca Raton

Sargin A, Ünlü A (2009) Inductive item tree analysis: Corrections, improvements, and comparisons. Mathematical Social Sciences 58:376–392, DOI 10.1016/j.mathsocsci.2009.06.001

Schrepp M (1999) On the empirical construction of implications between bi-valued test items. Mathematical Social Sciences 38:361–375, DOI 10.1016/S0165-4896(99)00025-6

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

Schrepp M (2006) ITA 2.0: A program for classical and inductive item tree analysis. Journal of Statistical Software 16(10):1–14, DOI 10.18637/jss.v016.i10

Schrepp M, Ünlü A (2015) On the creation of representative samples of random quasi-orders. Frontiers in Psychology 6(1791):1–8, DOI 10.3389/fpsyg.2015.01791

Thompson SK (2012) Sampling. Wiley, Hoboken, DOI 10.1002/9781118162934

Ünlü A, Sargin A (2010) DAKS: An R package for data analysis methods in knowledge space theory. Journal of Statistical Software 37(2):1–31, DOI 10.18637/jss.v037.i02

Ünlü A, Schrepp M (2015a) Biasing effects of non-representative samples of quasi-orders in the assessment of recovery quality of IITA-type item hierarchy mining. In: Wilhelm AFX, Kestler HA (eds) Analysis of large and complex data, Springer, Heidelberg, Studies in Classification, Data Analysis, and Knowledge Organization, DOI 10.1007/978-3-319-25226-1_48

Ünlü A, Schrepp M (2015b) Untangling comparison bias in inductive item tree analysis based on representative random quasi-orders. Mathematical Social Sciences 76:31–43, DOI 10.1016/j.mathsocsci.2015.03.005

Van Leeuwe JFJ (1974) Item tree analysis. Nederlands Tijdschrift voor de Psychologie 29:475–484