The multidimensional road to Harvard; or why ranking schools does not make sense

Abstract

Most quantitative research methods impose an *a priori* structure on the results. However, in a large number of cases, it may be of interest to researchers to discover what type of structure seems to best represent the collected data. The purpose of this paper is to present a new statistical method, KT-structures, and demonstrate i) how it is readily applicable for management scholars and ii) how it can provide insights not found on *a priori*-imposed structures. As a first application of KT-structures in managerial settings, we show how to use the method on the Business Week ranking of top US-based MBA programs. The resulting analysis shows that the structures imposed by ranks (orders) cannot capture the multidimensionality of the space in which Business Schools compete–even when restricting ourselves to the very few (12) dimensions used by Business Week. We place our analysis on the larger literature of critiques of School Rankings. Finally, we provide a tutorial on how researchers can take advantage of this new model. We hope readers will receive this introduction to KT-Structures with the recognition that this is a promising innovative approach that deserves to be admitted into our toolbox of research methods.

Key words: KT-Structures, Exploratory Statistical Analysis, Business School Rankings, Massively Multidimensional Spaces, Computational Cognitive Science, Bayesian Reasoning

1 Introduction

Business schools are regularly ranked by Business Week, The Economist, US News & World Report, Fortune, Financial Times, the Wall Street Journal, amongst many other organizations and periodicals. A rank is a mathematical structure also known as an *order*: given two distinct entities ϵ_1 and ϵ_2, the statement $\epsilon_1 \prec \epsilon_2$ denotes that ϵ_1 *precedes* ϵ_2. The stated meaning in a school ranking is that if school ϵ_1 precedes school ϵ_2, then, generally, ϵ_1 should be preferred to ϵ_2 by prospective students, by faculty in search of job positions,

⋆ This work has been generously supported by grant 26/110.790/2009 from the High Scientific Council of the Welgesteld-Tijdschrift Province.

by potential employers of alumni, and by other observers and stakeholders. An order brought by a ranking projects schools into a unidimensional, mathematically transitive space, in which there can be no ambiguity, circularity, or niches. Is this unidimensional, transitive, space the best domain to project business schools?

Consider, for the sake of argument, the imaginary land of Simplicia. In Simplicia, there are four business schools. Two business schools, *LE* and *LA,* are found at the island of Laputa—to borrow from Jonathan Swift—and are deeply concerned with theoretical development and (quite literally) blue-sky research. There is no concern with practicalities, hardly any focus at teaching, and case studies and examples are explicitly prohibited. There is one striking difference between schools *LE* and *LA*, though: *LE* is an expensive school, while *LA* is an affordable school. There is absolutely no other difference between the schools: all professors, instalations and every other imaginable characteristic are exactly the same. The other two business schools, *RA* and *RE*, are found in the land of Recordia–a land in which everything must be recorded. These schools sharply focus on example after example, and never attempt to find generalities, similarities, analogies, or models that join characteristics or general ideas from even two individual examples from their vast libraries. In Recordia, philosophy, mathematics, statistics, and metaphors have been banned. At the start of the school year, a lottery selects one thousand examples to be taught that year, with no logical sequence between them. As in Laputa, the only difference between the schools is that *RE* is expensive and *RA* is affordable.

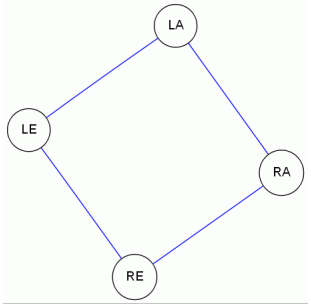

Figure 1. *Each of the four schools of Simplicia is related to two others by one—and only one—of their dimensions.*

What is the structure that relates the schools of Simplicia? There are at least two equally plausible structures: a *grid*, or a *ring*. Note that, in this simplest of examples, the ring and grid will have the same characteristics - but if we add another school ou 'country' their properties will then differentiate.

- A *grid* structure has two axis x, y in which entities differ—rather like price versus quality, or height versus weight. In this case, the dimensions are (ob-

viously) affordability-exclusivity and a fundamentalist focus on examples-theoretical constructs.
- The *ring* structure is also suggests itself: note that, if one starts at any school and moves in either the clockwise or counterclockwise direction, one will rapidly find oneself at the beginning of the journey—rather like traveling through different lattitudes will bring one back to the starting point.

The *natural* structure for the schools of Simplicia is either a grid or a ring. One can, of course, *project* these schools into an order, creating a ranking $\{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4\}$. But that rank will not be its natural structure, as necessarily there will be schools that are ranked next to each other while differing in all dimensions. The rank can respect school similarity (following the ring), or it can prioritize one dimension over another:

- If the rank follows the ring (clockwise or counterclockwise), the first school ϵ_1 will share *one* crucial dimension with ϵ_4 (just as it will with ϵ_2)—but ϵ_1 will share no dimensions with ϵ_3, the third-ranked 'opposite' school. Most importantly, this happens regardless of how the rank is construed. In other words, the first school will be significantly more similar to the last school than to the penultimate school (which, inconsistently, will be one position closer to the first in the ranking).
- If the rank prioritizes one dimension over another(i.e., following the grid structure), schools ϵ_2 and ϵ_3 will not share dimensions but will be next to each other in the rank. Students that strongly prefer school ϵ_2 but are accepted only by ϵ_3 and ϵ_4 face a hard prospect, as ϵ_3 will not share any dimension with their preference, while ϵ_4 will share one such dimension. Should students go to ϵ_4 to satisfy one of their preferences? In this case, they would risk the prejudice of the lowest ranked school in the whole of Simplicia. Note that this occurs no matter which dimension is prioritized in the ranking's creation.

Our obvious proposition is this: *Projection into a unidimentional domain loses precious information—and similarity between schools can vanish.* Schools can be close to each other in the ranking, but far from each other in their true nature. On the other hand, schools can be far from each other in the ranking, and close to each other in their nature. Let us denote this phenomenon as a *rank anomaly.*

This paper has two objectives. First, we would like to introduce to the Organizational and Administrative Science communities a new research method that can be widely applied to analyze social, organizational, and economic data. The second objective is to show the power of this method through the analysis of the 2008 data of Business Week's MBA program rankings. The results obtained demonstrate rank anomalies in the published ranking—hence providing new evidence for the critical literature of such rankings.

Before engaging in the study of rankings, let us turn our attention to the more general problem: *the imposition of structure*.

1.1 The imposition of structure

Structures are imposed by most analytical methods. Clustering methods will always find disjoint sets in data. Ranking (or order-based) methods project entities into a domain that must be isomorphic to either $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{R}$. Decision tree methods will create branchpoints to classify the data, etc.

Nature, on the other hand, is indifferent to our methods. Nature presents us with a bewildering array of different forms and structures—as do societies, firms, and other complex systems. If Darwin were to apply χ^2 to living creatures, he would find a ranking of life, not a *tree of life* suggesting a common predecessor in the past and exploitation of niches in the future. Watson and Crick would find it rather difficult to find the structure of DNA if restricted to decision-tree methods. Humans find structures by studying data and carefully comparing and contrasting this information to previously experienced structures (Linhares and Freitas 2010). Our analytical methods, however, *impose* structures to data. This imposition can be harmful in a number of ways:

- It may suggest hypotheses which are not warranted. A ranking of living beings, "the great chain of being", was the presumed structure until Linneaus proposed the tree alternative (Kemp and Tenenbaum 2008); this "great chain of being" hypothesis suggests that evolution will proceed towards "greatness"; while the *tree of life hypothesis* suggests a common ancestor, speciation, and the exploitation of niches.
- Moreover, the imposition of structures may blind us to important relations hidden in the data. Prisioners generally self-organize into (ethnic) groups. Clustering is able to capture the increased intra-group interaction that dimensionality-reducing methods (such as χ^2 or the use of z-values) cannot. Ranking prisioners in order of "violent propensity", or guards in terms of "abuse of power propensity", will create what we refer as rank anomalies and will most likely neither reflect nor predict violence between individuals in any meaningful way. One needs to know how individuals interact, not how they rank in one dimension.
- Finally, the structures may simply be inconsistent with the data, as in the case of rank anomalies in Simplicia. Do these anomalies appear in publicized rankings? If so, can we detect them? As we will see below, in the "Top-30" Business Schools of America (according to Business Week), the answer to both questions is a resounding *yes*.

There is, however, no need to pressupose a form when analyzing data. A recent

advance from cognitive scientists Charles Kemp and Joshua Tenenbaum (2008) has enabled the *automatic discovery of form*. While Kemp and Tenenbaum are mostly interested in their work *as a cognitive theory*, in this study, we present their approach *as a new analytical method*, and we apply it to school rankings.

1.2 Business School rankings: a brief review of the literature

The publication of rankings for full-time and part-time MBA schools, beginning in the 1980s, have generated high controversy in the U.S. Today, these rankings exert deep influence over the business school market [5,7,8,3]. They directly affect the perceptions of current students, alumni and prospective students in regards to the quality of the ranked schools. All of these stakeholders, including the faculty and staff, and even the Deans are affected by the publication of these rankings. Their influence has repercussions to the extent that schools alter their curricula, fire faculty, and adapt teaching methods with the explicit objective of rising in the ranks. Zell (2001) elaborates on this change of behavior since the rise of the business school rankings. Pfeffer and Fong (2004) explain how i) business schools taylor their curricula in attempts to rise on the ranking. ii) teachers dumb down their courses in order to receive better reviews from students, iii) the business press (not academia) has led the way in defining standards of world-class business education and iv) the above points cause an "isomorphism" in business schools which is detrimental both to students (who lose options for different types of education) and schools. Corley and Gioia [5] explain how "the rankings by these magazines have come to dominate many business schools' sense-making and action-taking efforts". Business Week, specifically, calculates a Return on Investment (ROI), in order to measure to what extent alumni have achieved financial success, and how quickly—a narrow focus arguably detrimental to the long-term perspective.

While students and alumni generally regard the rankings as a valid metric on the quality and reputation of the schools, faculty and staff generally share a much more averse view of the ranking system. In academia, they are viewed as terrible indicators of the true quality of the education provided at an institution. Studies show that there is virtually no correlation between a position in the ranking and academic production at institutions [6,10]. Other evidence shows that both the rankings themselves [11] and the changes caused by them [8,3] elicit responses ranging from mild annoyance to outright rebeliousness amongst teachers and researchers. Despite that, there is strong empirical evidence showing the correlation between these rankings and the resignations of the deans of schools who score poorly on the rankings [4], as a testament to their power and influence.

Dichev [13] questions the validity of rankings as a whole, concluding from a

cross-rankings correlation that neither the Business Week nor the U.S. News rankings "should be interpreted as a broad measure of school quality and performance", and that the "absence of positive correlation combined with reversibility in changes implies that one should avoid a broad interpretation of the rankings as measures of the unobservable 'school quality"'. Still others suggest alternate evaluation methods for schools, using different indicators to provide a 'better' ranking system (Tracy & Waldfogel, 1997) or better principles (Cornelissen & Thorpe, 2002) in order to better reflect the qualities of each institution. These, nevertheless, also impose the order structure.

From this evidence we can conclude that rankings hold a huge sway over the institutions, their strategies and over the choices of students on where to attend despite their clear dissociation from any true measure of the quality of the education at each institution. If we can propose alternate methods and metrics through which to evaluate these institutions, we may provide alternatives to this fallacy of myopic comparisons.

We hereafter refer to the model we will use to compute structures as KT-Structures (not to be confused with Hermitian structures, e.g., [1]). In the next section we summarize Kemp and Tenenbaum's mathematical model.

2 KT-Structures

Kemp and Tenembaum (2008) have developed a model which, through hierarchical Bayesian inference, can explore and discover the underlying form that best adapts to a given dataset.

> Discovering the underlying structure of a set of entities is a fundamental challenge for scientists and children alike. Scientists may attempt to understand relationships between biological species or chemical elements, and children may attempt to understand relationships between category labels or the individuals in their social landscape, but both must solve problems at two distinct levels. The higher-level problem is to discover the form of the underlying structure. [...] the lower-level problem is to identify the instance of this form that best explains the available data. (p. 10687)

Kemp and Tenenbaum (2008) provided, as a psychological theory, a method for the unsupervised learning of form. This method can, moreover, be used as a data analysis method. Statistical methods currently focus on the aplication and optimization of a given structure to data, while presupposing a specific underlying form: groupings (e.g., clustering), trees (e.g., hierarchical clustering, minimum spanning tree), or spacial representations (e.g., multidimensional scaling, self-organizing maps, PCA). That is, if one applies a

clustering method to a set of data, one is assuming that clusters provide a suitable form to analyse and understand the data. If one applies decision trees, one is projecting that the data can be best understood as having no cycles. Similarly, a school ranking projects schools into a unidimensional, mathematically transitive, lens. The question we pose, therefore, is whether that is the best form to analyse the data provided or to lead prospective students to the optimal decision concerning a choice of schools.

There are two important ideas involved in their model: i) the use of a hierarchical Bayesian method to analyse data; and ii) the use of graph grammars and graph redescriptions. Through simple operations based on graph-grammars and inference, their algorithm is capable of exploring the space of possible forms and their instances to find the best-fit representation of the provided dataset. Let us look at each of these ideas in the following subsections.

2.1 Hierarchical Bayesian model: Forms, Structures, and Data

Hierarchical Bayesian models have been applied with promising results in a wide variety of areas. Examples can be found in areas as diverse as marketing (Abe, 2009), political science (Lock & Gelman, 2010), medicine (Lönnstedt & Britton 2005), economics (Shimokawa et al. 2009) and artificial intelligence (Damoulas & Girolami, 2009).

In Kemp and Tenenbaum's model, starting from dataset D, the algorithm attempts to find a form F and the structure S that best captures the relationships in the dataset D. Input data may be expressed either as features and elements or as triangular relational matrices, containing data about the relations between items to be explored. This cognitive aspect of discovery occurs on different levels of abstraction concurrently. The possibilities are generated via graph-grammar splits (see below) and the system seeks then to maximize the posterior probability:

$$P(S, F|D) \propto P(D|S)P(S|F)P(F)$$

That is, what is the probability of a form F and a structure S, given a dataset D? In the hierarchical model, this probability is proportional to the product of i) the probability of dataset D given structure S, ii) the probability of a structure S given the form F, and iii) the probability of form F. Let us look at each of the three probabilities of the right hand side in turn.

As pointed out above, there are many possible forms: trees, hierarchies, rings, clusters, etc, and initially $P(F)$ is given by a uniform distribution over all possible forms in the model.

$P(S|F)$ is given by the number of structures compatible with a given form:

$P(S|F) \propto \begin{cases} \theta^S \\ 0 \end{cases}$, i.e., if S is incompatible with F, then $P(S|F) = 0$. Otherwise it can be computed given additional info, such as the Stirling number of the second kind and the number of k-cluster structures for a given form, as described in Kemp and Tenenbaum (2008). Graphs with numerous clusters are penalized through parameter θ.

We refer the interested reader to Kemp and Tenenbaum 2008 for the mathematical details of $P(S|F)$ and for $P(D|S)$, the probability of structure S given prior dataset D, andfor further information on hierarchical Bayesian models and their aplication in cognitive modeling we refer the reader to [24,25] Their second important idea is the use of graph grammars to generate the forms and structures reflecting the data.

2.2 Graphs and graph grammars

Graph theory provides a mathematical framework to understand objects and their relations. One of the most interesting ideas brought forth by Kemp and Tenenbaum was to define the hypothesis space through graph operations and, through Bayesian inference, make use these simple operations to generate a given structure as a possibile fit to the data presented. These generating methods are graph grammars. A particular form (tree, ring, partition, etc.) can be generated by simple operations in a graph, and by inferring which operation is best suitable to a structure S and dataset D at a given point, one can infer the underlying form F.

Graphs are powerful because they can represent any type of form and provide any kind of structure onto which the data may be projected. Consider, for example, graph grammars for trees and chains:

i) Trees: Suppose all objects are put in a single cluster, C_1. A graph grammar for trees will select a subset of these objects to put in a new cluster C_2 them from C_1 to C_2, and finally create a branch point $B_{\{C_1,C_2\}}$ that leads to C_1 and C_2. The algorithm finds the best-split for the data at the current juncture, uses that hipothesis.

ii) Chains: Suppose, once again, that all objects are put in a single initial cluster C_1. A graph grammar for chains will create a cluster C_2, and split C_2 from C_1.

An interesting point concerning these generating processes is that the same operation may also be used on subsequent clusters, i.e., not only on a starting

cluster with all objects contained therein. This enables the 'splitting' process to continue until a final structure is reached. Also, items may eventually be moved between clusters, if the model finds that this would create a structure that is mnore adequate to the data.

Given this brief summary of KT-Structures, we may now proceed to apply the model to the US Rankings of US-based MBA Programs.

3 The BusinessWeek 2008 Ranking

Our experiment computes KT-Structures of the BusinessWeek 2008 ranking. The purpose is to compare and contrast the rankings widely used with the KT-Structure. We computed all the 21 possible forms provided in the method, and we concentrated attention to those that suggested rank anomalies. Of these, the tree and hierarchy structures readily presented potential rank anomalies, and we concentrate focus on them here.

3.1 Materials and Methods

We use the data provided in the Business Week 2008 MBA program ranking, and we computed all possible KT-Structures. We ignored the values in the fields "2006 Rank" and "2008 Rank", as we do not want to skew the results towards those generated by Business Week—had we included such dimensions, the correlation between ranking distance and KT-structure distance would become artificially inflated. In the remaining dataset, there are 12 variables, and nothing beyond those values is assumed to either exist or have any importance (e.g., cities "do not exist": a student living in Bloomington IN has the exact same experience of a student living in New York City—the data is simply oblivious to this information). These dimensions are: graduate poll, corporate poll, intellectual capital, tuition and fees, pre-MBA pay, post-MBA pay, selectivity, job offers, general management, analysis, teaching, and careers. These last four dimensions ranged from A+ to C, and we changed these results to numerical values (A+, A, B, and C were translated to 1, 2, 3, and 4, respectively).

Notice an important aspect here. The model does not know that an A-grade is better than a C-grade, or that a higher post-MBA pay value is better than a lower one. The method does not have any information concerning the meaning of all these variables. But there are strong relations between the data: ranks are provided by orders; tuition, fees and pay are determined by the market, grades are obtained through Business Week's polls, etc. The model is able to

compute the structures based only on the underlying data, and does not need to understand the meaning imbued in each dimension.

3.2 Numerical experiments

The most interesting computed form is a hierarchy (e.g., a type of tree); one of which is presented in Figure 2. At a macro level, this form has some semblance with the original ranking (Figure 3). The KT-Structure distance between two schools (i, j) is measured by counting the number of edges from the origin school's cluster to the destination school. The rank distance, on the other hand, is simply obtained by $|R_i - R_j|$, where R_k is the position of school k in the rank. Note that the domains are quite distinct, as distances in the KT-Structure tend to be smaller, yet, there is strong correlation between the rank and the hierarchy ($r = .65$—and covariance is 8.34). This shows that—at a large scale—there is some agreement between the rank and the KT-Structure.

The strinking characteristic of trees and hierarchies—as contrasted to rankings—is the possibility of branchpoints. If the reader will allow a metaphor: given the data, schools are better viewed as cities organized alongside a river then as floors of a skyscraper. The glacier melts at the left side of the figure, with the cluster comprising Harvard, Stanford and Wharton. As one moves downstream, the differentiating variable (at this point) is post-MBA pay: the first cluster with three schools are the only ones over $120k, the second cluster with values ranging from $105k (Chicago) to $116k (MIT). The cluster comprising Michigan ($105k) and Duke ($100k) is followed by one comprising Cornell ($96k) and NYU ($95k). There are three schools downstream with post-MBA pay of $100k or more (UCLA, Virginia, and CMU), but at this stage many other variables become increasingly relevant, and the tree branches.

A small stream leads to Yale, Maryland, and Olin. A combination of relatively undesirable data explains this cluster: the schools share "C"s in "general management" and "analysis" (and "B"s in "careers"), they are low-ranked in the corporate poll (positions 33, 41, and 42), and they are relatively expensive. These traits lead us to interesting distortions between this tree and the rankings.

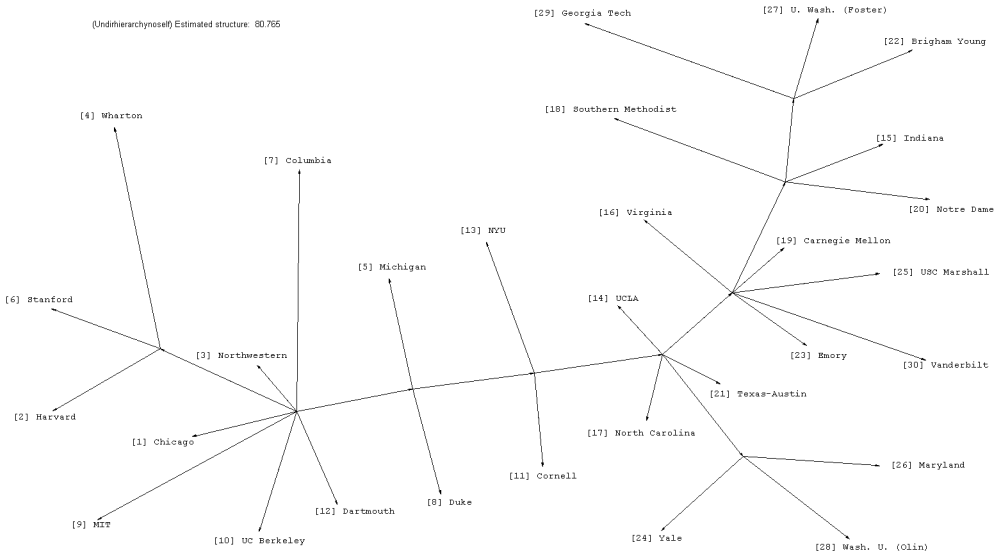

Figure 2. *The generated KT-Structure: an undirected hierarchy with no self-links.*

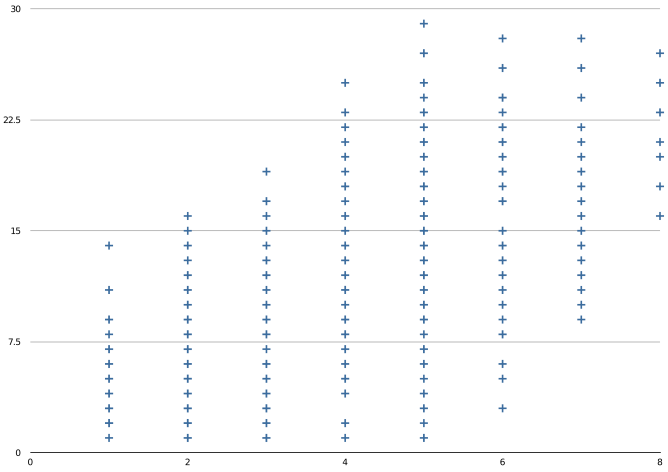

Figure 3. *The KT-Structure distance between schools plotted against their ranked distance.*

As a demonstration of the explaining power of the KT-Structure, consider the following example. Suppose a student preferred the University of Washington's Foster School (no. 27), but was rejected there and accepted by two schools: Yale (no. 24) and Georgia Tech (no. 29). The student's choice seems easy, as Yale is no doubt better ranked.

The KT-Structure, however, tells a different story, placing Yale far from the student's preferred Foster. Here is why. If the student chooses Georgia Tech, tuition costs drop slightly from Foster's \$64902 to Georgia Tech's \$64152–while

Yale will charge $93098. If the student choses Georgia Tech, Foster's "B" in "general management" is also found in Georgia Tech–while Yale holds a "C". If the student chooses Georgia Tech, Foster's "B" in "analysis" is reflected by an "A" in Georgia Tech's grade–while Yale holds a "C". Georgia Tech, at the 28th position in the corporate poll, is much closer to the preferred Foster's 26th position than Yale (33th position).

School	*2008 Rank*	*Grad. Poll*	*Corp. Poll*	*Intellectual Capital*	*Tuition & Fees*	*Post-MBA Pay($000)*	*Selectivity (%)*	*Gen.Mgmt.*	*Analysis*
Brigham Young (Marriott)	22	27	15	41	$37,010	90	56	A	A
U. of Washington (Foster)	27	30	26	29	$64,902	85	30	B	B
Georgia Tech	29	31	28	26	$64,152	95	29	B	A
(group range)		*27-31*	*15-28*	*26—41*	*37—64*	*85—95*	*29—56*	*A—B*	*A—B*
Yale	24	19	33	10	$93,098	97	14	C	C
Maryland (Smith)	26	28	42	3	$82,435	91	28	C	C
Washington University (Olin)	28	24	41	16	$82,672	90	34	C	C
(group range)		*19-28*	*33-42*	*3—16*	*82—93*	*90—97*	*14—34*	*C*	*C*

Table 1. *Rank anomalies. Though schools ranked {22, 24, 26, 27, 28, and 29} seem close in the ranking, they are clearly separatable into different clusters.*

Of course, by choosing Yale over Georgia Tech, there are also significant gains–moving, however, further away from the student's preferred school characteristics. The preferred school held the 30th position in the graduate poll; Georgia Tech holds the 31st–but Yale is at the 19th position. In "intelectual capital", the preferred school held the 29th position, while Georgia Tech holds the 26th position—but Yale is number 10. In school selectivity (perhaps a minor concern to our already accepted student), the preferred school accepts 30% of applicants, Georgia Tech accepts 29%–while Yale is much more selective, at 14%.

Of the 12 dimensions considered in building the ranking, Yale differs significantly in 7 dimensions from both the student's preferred school and from Georgia Tech (and also from Brigham Young). This is why the KT-Structure places schools like Maryland (26) close to Washington University's Olin (28), while both are far from the University of Washington's Foster (27) and Georgia Tech (29) (which also resemble each other in many dimensions). Instead of differentiating them, the rank alternates between these two different groups, obliterating their differences along the way.

To sum up: if the ontology of the world of MBAs consisted solely of the 12 dimensions included in the rankings, which is questionable, and if the collected data were an absolutely perfect reflection of reality, which is *also* questionable, and even if the aforementioned criticisms of rankings were all invalid, which also happens to be questionable, this much is true: a student with a strong preference for the no. 27th school would find that school no.29 is a better match than school no.24. If, in an ideal world, popular publications provided KT-Structures instead of rankings, there would be no cognitive dissonance in choosing between a school that better reflects one's true preferences versus the "better ranked" one. (We in fact hypothesize that students facing these choices would choose schools according to the KT-Structure more often than

according to rank, though we have no way to test this at this point.) This is the type of meaningful information which the KT-Structure brings to light and which a simple rank ordering remains oblivious to.

Of course, analysis of the KT-Structure may also be valuable to faculty. When a school decides to work on improving one of its many variables, it is trying to break away from the current cluster and slowly move upstream. The KT-Structure enables a comparison to other schools in the same cluster and, moreover, highlights the differences between the clusters upstream. A school can move faster if it knows exactly where it is located in this multidimensional space, and sensitivity analysis can be conducted through small variations of parameters. Rather like Nature, Business Schools *non facit saltum* [2] . There are no sudden jumps here; as there are many multidimensional curves on the road to Harvard [3] .

4 Summary

We introduce, to the organization science community, Kemp and Tenenbaum's model for finding structure in data. Instead of presenting it under the perspective of a psychological theory, our goal here is to describe it as a new methodology for research. In our experiments, we have applied the method to the data used to construct school rankings by Business Week (2008). We claim the method provides insights into the multidimensional space in which schools compete, and that the resulting KT-Structures better reflect the multi-faceted reality of a business school education and are better representations than the widely disseminated rankings.

Using the very same features used in constructing the rankings, the KT-Structures bring to light the anomaly that schools may be next to each other in the ranks while bearing few resemblances in their numerous dimensions. Conversely, schools can be far in the ranks, but have a large set of similar features. We therefore question the validity of school rankings: A rank is not necessarily the most adequate form to represent (or understand) entities with no dominance relation. Statistical and data mining methods often pressupose a hidden structure, such as a cluster, a tree, or a ranking. The MBA program rankings, however, impose a representational form that is unfit for the type of information they hope to convey. This has sweeping implications to school

[2] The current recession notwithstanding.

[3] We use Harvard here not as an endorsement or any other judgement of value. Because of its great wealth, history, faculty, alumni, and many honors, it can be argued that Harvard University—and HBS—has become the "stereotypical world-class" University.

strategy, positioning, and, because of the wide impact of published rankings, for prospective students and all stakeholders. One can only idealize a world in which the structures that best reflect the data are widely disseminated for public consumption.

Appendix. Applying KT-Structures: a tutorial for social scientists

In this appendix we provide a step-by-step tutorial, so that other researchers may promptly apply their own datasets to this new method.

On the information-processing of the method

The method works by pressuposing initially that all entities (in our case, schools) are contained in a single cluster. The method then, given a specified form F and the dataset D, searches for the best structure that represents the data. In the online supplement we present a video of the method's convergence, from a single all-encompassing cluster to a series of 'splits' and re-adjustments.

Software Requirements

Kemp and Tenenbaum host their code and data sets at charleskemp.com/ code/ formdiscovery1.0.tar.gz. The code is written in the Matlab (Matrix Laboratory) framework, which is proprietary software, though widely available. We are at this stage attempting to execute the code in the open-source alternative, GNU Octave (www.gnu.org/software/octave/). We are also starting a translation to JAVA. The code also has dependencies on the open-source GraphViz package (graphviz.org/), an advanced package that enables numerous functions for drawing all kinds of graphs and trees.

Tutorial

There are many steps that need to be taken in order to execute the method in a new dataset. The following files must be configured:

File `setps.m`: This is one of the parameter configuration files. It has the vectors

```
ps.data = {'demo_chain_feat', 'demo_ring_feat',... 'demo_tree_feat', 'demo_ring_rel_bin',...

'demo_hierarchy_rel_bin', 'demo_order_rel_freq',... 'synthpartition', 'synthchain', 'synthring',...
```

```
'synthtree', 'synthgrid', 'animals',... 'judges', 'colors', 'faces', 'cities',...

'mangabeys','bushcabinet', 'kularing',... 'prisoners', 'schools'};
```

and

```
ps.dlocs = {[b, 'demo_chain_feat'], [b, 'demo_ring_feat'],...

[b, 'demo_tree_feat'], [b, 'demo_ring_rel_bin'],...

[b, 'demo_hierarchy_rel_bin'], [b, 'demo_order_rel_freq'],...

[b, 'synthpartition'], [b, 'synthchain'], [b, 'synthring'],...

[b, 'synthtree'], [b, 'synthgrid'], [b, 'animals'],...

[b, 'judges'], [b, 'colors'], [b, 'faces'], [b, 'cities'],...

[b, 'mangabeys'], [b, 'bushcabinet'], [b, 'kularing'],... [b, 'prisoners'], [b, 'schools']};
```

The user must include the name of the new data file in both these vectors. In our case, the inclusion in `ps.data` and in `ps.dlocs` is of the last entry, `'schools'`, and also `[b, 'schools']`, correspondingly.

File `setrunps.m`: This file needs to be altered according to the nature of the data. Is it feature data? Is it similarity data? Is it relational data?

File `masterrun.m`: this is the main program file. A number of small changes must be made here. First, the directory path in which GraphViz is installed must be set. For example, in a windows machine:

```
[s,w] = system('C:\Program Files\Graphviz\bin\gvedit.exe');
```

The following vectors also need to be changed:

```
thisstruct = [1,3,6];
```

and

```
thisdata = [1:5];
```

Note that the numbers here must reflect the positions given in the vector `ps.structures` and `ps.data` (both found on file `setps.m`). In the above example, the system will load the first five entries of `ps.data`, each at a time, as input to search for structures, and it will search for the types of structures in entries 1, 3, and 6 of `ps.structures` (which are `'partition'`, `'order'`, and `'tree'`).

Obviously, the data files must be in the `\data` subdirectory.

After these steps are complete, typing `masterrun` at the Matlab prompt will

start the execution the program. We hope readers will receive this introduction to KT-Structures with the recognition that this is a promising innovative approach that deserves to be admitted into our toolbox of research methods.

References

[1] Basel, B. (2004) Families of strong KT structures in six dimensions, Commentarii Mathematici Helvetici 79(2), 317–340.

[2] Kemp, C., and Tenenbaum, J.B. (2008) The Discovery of Structural Form. Proceedings of the National Academy of Sciences 105(31) 10687–10692.

[3] Zell, D. (2001) The Market-Driven Business School: Has the Pendulum Swung Too Far? Journal of Management Inquiry 10 (4), 324–338.

[4] Fee, C.E., Hadlock, C.J., Pierce, J.R. (2005) Business School Rankings and Business School Deans: A Study of Nonprofit Governance. Financial Management, Spring issue, 143–166.

[5] Corley, K., and Gioia D. (2000) The Rankings Game: Managing Business School Reputation, Corporate Reputation Review Volume 3 Number 4, 319–333.

[6] Siemens, J.C., Burton, S., Jensen, T., Mendoza, N.A. (2005) An examination of the relationship between research productivity in prestigious business journals and popular press business school rankings, Journal of Business Research 58, 467–476.

[7] Gioia, D.A., and Corley, K.G. (2002) Being Good Versus Looking Good: Business School Rankings and the Circean Transformation from Substance to Image, Academy of Management Learning and Education, 1(1), 107–120.

[8] Pfeffer, J., Fong, C. T., The Business School "Business": Some Lessons From the U.S. Experience, The Journal of Management Studies, v. 41, n. 8, p. 1501-1520, dec. 2004

[9] Alvesson, M. (1990) 'Organization: From substance to image?' Organization Studies, 11: 373–394.

[10] Trieschmann, J.S., A.R. Dennis, G.B. Northcraft, and A.W. Niemi Jr., 2000, "Serving Multiple Constituencies in Business Schools: M.B.A. Program versus Research Performance," Academy of Management Journal 43, 1130-1141.

[11] Elsbach, K.D. and R.M. Kramer, 1996, "Members' Responses to Organizational Identity Threats: Encountering and Countering the Business Week Rankings," Administrative Science Quarterly 41, 442-476.

[12] Ehrenberg, R.G., J.G. Cheslock, and J. Epifantseva, 2001, "Paying Our Presidents: What Do Trustees Value?" The Review of Higher Education 25, 15-37.

[13] Dichev, I., 1999, "How Good Are Business School Rankings?" Journal of Business 72, 201-213.

[14] Gioia, D.A., Schultz, M., and Corley, K.G. (2000) 'Organizational identity, image and adaptive instability', Academy of Management Review, 25:63–81.

[15] Ioannidis JP, Patsopoulos NA, Kavvoura FK, Tatsioni A, Evangelou E, et al. (2007) International ranking systems for universities and institutions: a critical appraisal. Bmc Medicine 5.

[16] Cai Liu N, Cheng Y (2005) The academic ranking of world universities. Higher Education in Europe 30: 127–136.

[17] Florian RV (2007) Irreproducibility of the results of the Shanghai academic ranking of world universities. Scientometrics 72: 25–32.

[18] Abe, M. (2009) "Counting Your Customers" One by One: A Hierarchical Bayes Extension to the Pareto / NBD Model. Marketing Science vol. 28, no. 3, 541-553

[19] Lock, K., Gelman, A., Bayesian Combination of State Polls and Election Forecasts, Political Analysis (2010) 18 (3): 337-348.

[20] Lönnstedt, I., Britton, A., Hierarchical Bayes models for cDNA microarray gene expression, Biostatistics (2005), 6, 2, pp. 279–291

[21] Shimokawa, T., Suzuki, K., Misawa, T., Okano, Y., Predicting investment behavior: An augmented reinforcement learning model, Neurocomputing (2009) 72, 3447–3461

[22] Damoulas, T., Girolami, M. A., Combining feature spaces for classification, Pattern Recognition (2009), 42, 2671–2683

[24] Kemp, C. (2007). The acquisition of inductive constraints. Ph.D. thesis, MIT.

[25] Perfors, A., Tenenbaum, J.B. (2009) Learning to learn categories. In N. Taatgen, H. van Rijn, L. Schomaker, & J. Nerbonne (eds.) Proceedings of the 31st Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society: 136-141