INSEAD

The Business School for the World®

Faculty & Research Working Paper

Identifying Styles in Product Design

Tian CHAN
Jürgen MIHM
Manuel SOSA
2013/24/TOM

Identifying Styles in Product Design

Tian Chan*

Jürgen Mihm**

Manuel Sosa***

* PhD Candiate in Technology and Operations Management at INSEAD, 1 Ayer Rajah Avenue, Singapore 138676, Singapore. Email: TianHeong.chan@insead.edu

** Associate Professor of Technology and Operations Management at INSEAD, Boulevard de Constance 77305 Fontainebleau, France. Email: jurgen.mihm@insead.edu

*** Associate Professor of Technology and Operations Management at INSEAD, 1 Ayer Rajah Avenue, Singapore 138676, Singapore. Email: manuel.sosa@insead.edu

Abstract

This paper introduces an approach to identify styles in product designs. We first built a theoretical foundation on *what a style is* and *how designs are configured into styles*. On this basis we executed a graph-clustering algorithm to identify styles in a population of design patents filed in the USA between the years 1976 to 2010. The design patent is uniquely suitable for this task because it is purely focused on how designs *look*. We validated the outcome obtained from the algorithm using experiments, which suggests that the algorithm outcome is no different from how humans would configure designs into styles. By establishing styles in designs, we open up a new avenue towards design research, upon which the dynamics of styles can now be studied.

INTRODUCTION

The Bauhaus style defined a *look* for architecture and furniture design – a look that emphasizes geometric simplicity, and conveys a sense of comfort and functionality (Gardner & Kleiner, 2009). This look influenced how millions of products were designed. But what factors made the Bauhaus so influential? Many other styles never proved to be influential to any meaningful degree, let alone to the degree of the Bauhaus movement. Abstracting from the specific example of Bauhaus, which factors make a particular style in general influential and which ones prevent success? This particular question provided the initial motivation for this paper.

Take Apple as an example – it is the biggest consumer products firm in the world, and it is well-known to make products with a distinctive *look*. How did this look come to replace our existing preconceptions on the looks of electronic devices? Does being located in California, the biggest *design hub* in the US, help it to establish that look? Who are the *people* that originated the style? How are styles with origins from a firm like Apple different from styles with origins from a school like the Bauhaus? Are they?

Without a thorough understanding of the notion of styles and a systematic way of identifying styles in relevant datasets, it is impossible to address such questions. Hence, these questions have so far eluded empirical researchers because the essential ingredient in the empirical setup has been missing: styles as a unit of analysis. The goal of this paper is to *empirically find styles* in a set of designs. More specifically, we use the *design patents database* from the US Patent & Trademark Office as a basis to identify styles among design patents. Design patents have so far not received attention from the academic community, but they form the ideal backdrop for our study since they are purely focused on how things *look*.

At the heart of our analysis is a firm understanding of the concept "style". Despite our ability to visually recognize and intuitively understand what a style is, "style" as a concept is difficult to define and hard to measure. Reflecting on two simple questions reveals the inherent difficulty: "What defines the Bauhaus style?" and "How do I know which designs belong to the Bauhaus school?" Clearly, the study of style is complicated by the ambiguity surrounding style definition (what is a style) and style measurement (how can we know if a design belongs to a style). This paper aims to resolve these two issues by providing a formal definition of style, theorizing on how designs are configured into styles, and empirically obtaining and validating a style construct on the US design patent database.

We structure our research agenda for this paper in the following way. First, we discuss and formulate a *definition of style*. Next, we put forward a theory on *how designs are configured into styles*. Then, we present our theory in a formal language, which provides a solid foundation to develop a complete *graph clustering* approach to identify styles in the vast dataset of the USPTO design patents. Lastly, we *validate*, with a series of experiments, that the outcome obtained from our algorithm are styles.

WHAT IS A STYLE?

Style – a physical interpretation

Munro (1946) described a style as consisting "of a combination of traits or characteristics which tend to *recur together* in different works of art". This recurrence occurs naturally, particularly in specific time and locations, so that artworks form "*distinctive clusters of… interrelated traits*". He used the example of Gothic architecture: pointed arches, high vaults, pitched roofs, slender piers, thin walls, large stained glass windows, and flying buttresses. to illustrate the idea that a style is *a set of physical traits*. Empirically, there is some evidence that this is true – Chan (2000) showed that the frequency of appearance of a set of pre-defined forms can evoke recognition of a style in architecture.

Designs and styles, however, do not just include individual traits but also a structural consideration of how the traits are *configured together*. An improvement to the "style as a set of physical traits" view is applied to product design by Stiny (1980) using shape grammar. The *shape grammar* is a design language that uses fundamental geometric shapes (squares, circles, etc.). Designs are then composed from these basic shapes through a series of transformation rules. The shape grammar has been used successfully to derive *individual physical definitions* of styles from studying the works of architects such as Palladio (Stiny & Mitchell, 1978) and Frank Lloyd Wright (Koning & Eizenberg, 1981) , or for brands such as Harley Davidson (Pugliese & Cagan, 2002) and Buick (McCormack, Cagan, & Vogel, 2004).

Note that a physical definition of style is typically *complex*, and requires a human to *identify* the set of physical traits salient to the definition of a style. The consequence is that methods that aim to autonomously *find styles* in product designs through physical measurements – see, for example, (Jupp & Gero, 2006) and (Hanna, 2006) – have been limited to very specific areas of application and even then have typically required some level of human supervision.

Style – a perceptual interpretation

The dichotomy of styles is that, while difficult to capture in a purely physical manner, is easily perceived and colloquially understood. A different way to approach the notion of style, according to McMahon (2003), is to think of it as a concept of classification through which an artwork is *perceived, understood, and interpreted* (see the same work for a discussion of how the use of the word style shifted through art history).

Potter (1976) showed through visual experiments that humans have the phenomenal ability to visually perceive and interpret the *meaning* of an individual picture (such as a politician, a car chase, or a picnic at the beach) within a split second. Humans can do this irrespective of the amount of clutter present in the picture, suggesting an ability to get a *quick overall impression*, or the gist, of a picture (Oliva & Torralba, 2006). In the context of style recognition, this suggests that a complex physical design get *translated* into a particular *concept* in the human mind. This translation process helps us answer the question of *which* set of physical traits is salient to the definition of a style – it is the set that maps to a particular *concept*.

What then exactly is a *concept*? A concept is our *mental representation* of the style. This mental representation helps the perceiver gain a coherent understanding of what the style *means*. For example, we can understand the "geometric simplicity" of Bauhaus designs through its founders' *values* that focus on functionality and egalitarianism, a desire for designs that *feels* "clean" and "logical", and the need to create mass-produced furniture suitable for middle-class homes (Burdek, 2005). This suggests that the concept of a style is composed of complex associations with values, emotions, and the time and place where the style is encountered (Stacey, 2006).

Style as a visually recognizable concept

The notion of style involves the interaction between the *physical characteristics* of the style, and its *perception* by an observer. This interaction is governed by the *perception and recognition* of *style concepts* by the perceiver, which in turn determines the set of physical traits that defines the style.

Thus, styles are *visually recognizable concepts*, or *mental representations* that help us gain a coherent understanding of the *meaning* of designs. The concept of a style is composed of complex associations with values, emotions, and the time and place where the style is encountered. This *conceptualization* helps to simplify understanding, and leads to the dichotomy that we observe in styles – complex to define physically, but intuitive and instantaneous to recognize.

HOW ARE DESIGNS CONFIGURED INTO STYLES?

The definition of style outlined in the previous section provides the *guiding principle* for organizing a set of designs into styles. Using this as a guiding principle, our goal is to empirically obtain styles given a set of designs. To do this, however, we need to know *how designs are configured into styles*.

Styles have different *extents*. Some styles, e.g. the Bauhaus, Gothic and the Baroque, were major styles that influenced designs in many products; others, e.g. the *batik* motifs prevalent in some parts of South East Asia, were designed for particular tastes and culture. Munro (1946) observed the presence of such *major* and *minor* styles. Yet, given any arbitrary set of designs, the set in general would *not* form a mega-style. The reason is that there are limitations on the *amount of variation* allowed for people to recognize a set of designs as a style. We theorize that a set of designs need to have a *sufficiently narrow concept to be considered a style*.

Styles are also *related*. Taking a real-world example, we can see that the family of large sedan cars conveys a sense of luxuriousness, comfort, and internal sophistication. Yet within this family we can clearly identify the BMWs from the Audis, and even within the BMWs we can clearly tell apart the 7-series and the 5-series. Munro (1946) first made a distinction between extensive styles (e.g. Italian Renaissance) and restricted styles (e.g. Florentine), and argued that restricted styles are sub-styles, variants, or phases of the extensive styles. This suggests that the relationship between styles is best captured within a *hierarchical structure* where super styles encompass a set of sub-styles which may encompass yet smaller sub-styles.

Thus, a set of designs should be configured into a *hierarchy of styles*. Within this hierarchy, designs with similar concepts should be placed within the same group (what is termed technically as *convergence*), and designs with different concepts should be placed within different groups (correspondingly, *divergence*). Groups may contain sub-groups, and at the "top" of the hierarchy (where the groups are the largest and most heterogeneous) each group should still have a *sufficiently narrow concept* to be recognized as a style.

Formal Properties of a Hierarchy of Styles

To avoid any ambiguity of what we mean by a hierarchy of styles, we use set notation to formalize our ideas. The formality is useful for two other reasons. Firstly, the concepts of convergence, divergence, and hierarchy are interrelated, and we show this interrelation by proving that convergence and divergence *leads* to a hierarchical configuration. Secondly, we will leverage on the formality introduced here to identify *what is required* of an algorithm to configure designs into styles.

Starting with the basic element of a design, we assume that every design has a visually recognizable concept V. Presumably V lies in some abstract space. We assume also that V can be compared – that is, given two arbitrary designs D_A, D_B, humans can compare the two designs stylistically and decide if they are similar or dissimilar. We model this as a distance function $f(D_A, D_B) \to R^+$.

Given an arbitrary set of designs $S = \{D_1, \ldots, D_n\}$, a *configuration* Ω is a *partition* of S into *clusters* $\{C_1, \ldots, C_k\}$ so that there is (1) *mutual exclusivity*: $C_i \cap C_j = \emptyset, \forall i, j$ and (2) *completeness*: $\cup_{i=1}^{k} C_i = S$

Definition: Ω ***configures designs into styles*** if there exists a non-negative number ϵ_T such that

1) *Convergence*: all the designs from a cluster are close in *V*, that is, $f(D_A, D_B) \leq \epsilon_T, \forall D_A, D_B \in C_i, \forall i$
2) *Divergence*: designs from different clusters are different in *V*, that is, $f(D_A, D_B) > \epsilon_T, \forall D_A \in C_i, D_B \in C_j, \forall i \neq j$
3) $\epsilon_T < \epsilon_{TMax}$, where ϵ_{TMax} is the *maximum allowable stylistic variation*, beyond which designs would not be recognized as a style.

Proposition 1 (*Existence*): given any arbitrary *S*, there exists at least one *Ω* that *configures the designs into styles.*

Note that the trivial configuration where all unique designs are in their own clusters (corresponds to $\epsilon_T = 0$) always configures designs into styles.

Proposition 2 (*Uniqueness*): given any value ϵ_T, if there is an organization *Ω* that configures designs into styles, this organization is *unique*.

Proposition 3 (*Nested-ness*): given a series of threshold values $\epsilon_1 > \epsilon_2 > \cdots > \epsilon_{k-1} > \epsilon_k$, the corresponding organizations $\Omega_1, \Omega_2, \ldots, \Omega_{k-1}, \Omega_k$ that configures designs into styles are nested, i.e. Ω_2 partitions Ω_1, Ω_3 partitions Ω_2, and so on.

Now, we have all the ingredients to formally define how styles are configured:

Definition*:* A ***hierarchy of styles*** is a set of organizations $\{\Omega_1, \Omega_2, \ldots, \Omega_{k-1}, \Omega_k\}$, corresponding to threshold values $\epsilon_1 > \epsilon_2 > \cdots > \epsilon_{k-1} > \epsilon_k$, each of which *configures designs into styles*.

ALGORITHMIC APPROACH

Now that we have established a theory on the nature of styles, we move to a discussion on the *algorithmic approach* used to identify styles. We divide the discussion into the following sections – the structure is modified from Milligan's work (1996) on cluster analysis with special considerations on how to obtain and validate *a set of nested clusters* of the same data.

1) **Data**: a discussion on the features of *design patents*
2) **Similarity index**: a measure of the *stylistic similarity* between two designs is obtained
3) **Clustering method**: the means by which *clusters* are formed
4) **Clustering steps**: a description of the $select \rightarrow partition \rightarrow flag$ steps of the algorithm.
5) **Outcome** of the algorithm
6) **Conditions required of the algorithm**: we bring the theory and algorithm together to state the conditions required for the algorithm to obtain a hierarchy of styles

Data: the design patent

A designer can protect a novel design by filing a *design patent* with the US Patent & Trademark Office (USPTO). Once the filing is successful, the designer can protect the intellectual property embodied in the *visual characteristics* of a design (USPTO, 2006). Note that the design patent is *purely focused on how a design looks*. A sample of a design patent document (for the design of the iconic Coke bottle) is outlined in **Appendix A.** The patent document consists of the drawings, designer information, company information, location, date of filing, product category, and a list of references made to previous works. Unlike other classes of patents which typically require a lengthy description for the invention, in design patents the design drawings contain *all* the intellectual property to be protected, and the description section typically contains only a single line, e.g. "the ornamental design for a bottle, as shown".

A granted patent is by definition *novel*. A design is novel if *an ordinary observer* would treat the design as different from already existing designs (USPTO, 2006). Thus, from the perspective of an ordinary observer, a set of design patents is a body of *unique* designs. The uniqueness property of the design patent lends added interpretability of styles found from it – for example, *the size of a style* is now meaningful in that it captures a number of unique designs instead of multiplicities of the same design.

The focus on visual looks and the uniqueness property of the design patent makes it particularly suitable for style analysis. As of today, close to 400,000 design-patent data from January 1976 onwards is made available on the internet by USPTO. We started a large-scale empirical analysis into styles based on design patents. The sample used in the paper includes design patents from January 1976 to January 2010 for a total of 365,444 designs.

Similarity Index

A similarity index is a measure of the *stylistic similarity* between any two designs. It is the main input to the algorithm. Any information contained within the design patent could be used to construct this index. For example, designs wrought from the hands of the same designer are *likely* to be of the same style. Alternatively, we could *directly examine* the design drawings and compare them based on style (though making such a comparison starting from scratch would be very time-consuming!).

Fortuitously, such comparisons are done as part of the *patent determination process*. When determining whether a design should be granted or refused patent status, patent examiners create references based on a comparison of the *visual impression* and *overall appearance* (USPTO, 2006) of the design under examination and prior patents. These references thus embody a rich layer of information capturing both the *physical characteristics* of designs and how patent examiners perceive the *concept* of the designs. This process of creating references is also *tightly controlled* – patent examiners follow a strict code on citations, and references suggested by inventors or attorneys have to be vetted, approved and potentially complemented (Pachys, 2012).

The reference is thus a *direct* and *consistent* measure of stylistic similarity between designs. This leads to our choice of using the references as the *single information source* to construct the similarity index.

A good similarity index should have some basic properties:

1) *Symmetric*: the similarity between two designs should be symmetric. In comparison, references are *directional*.
2) *Monotonic:* two designs that are stylistically more similar should have a higher index. This condition can be violated in the set of references because of the *time-dependence* of reference-making. Design patents filed around the same time may have significant stylistic similarities but would not refer to each other. This is especially common when designers file multiple patents at the same time to protect variations of a single design. To improve monotonicity, we make the observation that similar patents filed around the same time would refer and be tested for patentability against *the same set* of prior patents. This allows us to capture the similarity across two designs by measuring *the extent of overlap* in the sets of references – this method is called *bibliographic coupling* (Kessler, 1963).
3) *Bounded & Granular:* there are upper and lower bounds to the index, and it is possible to achieve any value within the thresholds. The reference relationship between two designs takes only two values (it exists, or it does not), meaning that the reference is binary instead of granular. To capture the different degrees of stylistic similarity between designs, we make the assumption that a patent that refers to many other patents is likely to be less similar to each one of them. Using this assumption, we adopt a heuristic by which *the weight* of a reference is normalized by the number of outgoing references of a patent.

The outcome of building this similarity index is a *weighted graph* spanning across all the designs (each design is a node), with each edge a measure of *stylistic similarity*. Given these requirements, we can actually suggest a specific index.

Mathematical definition of the Similarity Index: As input, we have a list of *designs* and a list of *references* between designs. Consider a single pair of designs *a* & *b*. The pair has a reference relationship $R(a, b)$ that takes 3 possible values: *1* if *a* refers to *b*, *-1* if *b* refers to *a*, and 0 if *a* and *b* do not refer to each other.

Denote the set of outgoing citations of *a* & *b* as *A*, *B* respectively. Denote |.| as the size of a set. The index is constructed as follows:

1. Construct a *normalized* measure $N(a, b)$ of reference to improve *granularity*:

$$N(a,b) = \begin{cases} R(a,b)/|A| \ if \ R(a,b) = 1 \\ R(a,b)/|B| \ if \ R(a,b) = -1 \\ 0 \ if \ R(a,b) = 0 \end{cases}$$

2. *Symmetrize* the measure by taking its absolute: Replace $N(a, b)$ with $|N(a, b)|$
3. Construct *bibliographic coupling*, or the extent of overlaps in reference, to improve *monotonicity*: $J(a, b) = |A \cap B|/|A \cup B|$
4. The sum gives the complete similarity index: $SI(a, b) = |N(a, b)| + J(a, b)$

By using the Jaccard Index (Jaccard, 1901) in Step 3, we ensure that $SI(a, b)$ is *bounded* between 0 and 1 (proof given in **Appendix B**).

Clustering Method

We use a graph clustering method, the *Ng-Jordan-Weiss (NJW)* algorithm (Ng, Jordan, & Weiss, 2002), to identify clusters. The key reason for adopting this algorithm is that it adopts a *global optimization* approach – it takes the entire similarity graph, defines a quality function, and finds a way to group data such this function is optimized. The formality allows us to interpret the nature of the clusters obtained through the quality function (the detailed *reasoning* of the choice and the *mechanism* of the algorithm are laid out in **Appendix C**).

Taking the *similarity graph* obtained in the previous section as input, the algorithm does three things. First, it *selects* a graph to partition (in the first step there is only a single graph, so by default it is selected). Second, it *partitions* the selected graph into two sub-graphs (or two clusters of designs). Third, it checks if this new cluster configuration is potentially a good way to group designs into styles (if so, it is *flagged*). The cycle of $select \rightarrow partition \rightarrow flag$ repeats until clusters are as fine as needed. By using the algorithm in this iterative manner, we automatically ensure that we obtain a *hierarchy* of clusters (See **FIGURE 3** for an example of a *dendrogram* representation of how designs can be configured hierarchically).

High-level Pseudo-code: Input – a set of designs $\{D_1, \ldots, D_n\}$.

1) ***Select***: Let $\boldsymbol{C}$ be the set of clusters generated by the algorithm (at the first step, there is only one cluster $C_1 = \{D_1, \ldots, D_n\}$, so $\boldsymbol{C} = \{C_1\}$). Select a cluster from $\boldsymbol{C}$ and denote it C_s.
2) ***Partition***: generate sub-clusters A and $\bar{A}$ from C_s,where $A \cap \bar{A} = \emptyset$ and $A \cup \bar{A} = C_s$
3) ***Flag***: Replace $\boldsymbol{C} = \{\ldots, C_s\}$ with $\boldsymbol{C_{new}} = \{\ldots, A, \bar{A}\}$. If $\boldsymbol{C_{new}}$ is a potential solution, flag it.
4) Repeat Steps 1 to 3 using $\boldsymbol{C_{new}}$ instead of $\boldsymbol{C}$**.**

Clustering Mechanism

Conductance: The graph measure of *conductance* is central to the *selection, partition,* and *flagging* steps of the algorithm. A brief introduction of the concept here will help in understanding the nature of clusters obtained.

Intuitively, the goal of clustering is to find a set of clusters that satisfies both convergence (designs in the cluster are very similar), and divergence (designs across different clusters should be different). One indication of convergence, in graph language, is *volume*. It is the *sum of edges* of all the vertices contained in a cluster. A cluster with high volume indicates that the graph representing the cluster is *well-connected*, suggesting convergence. One indication of divergence is *cut*. Given two clusters, cut is defined as the *sum of edges* that crosses between the two clusters. Two clusters with a small cut is *well-separated*, suggesting divergence.

Thus, given a cluster C that we want to partition into sub-clusters A and $\bar{A}$, a good partition would be such that $cut(A, \bar{A})$ is small, while $volume(A)$ and $volume(\bar{A})$ are large. *Conductance* balances these two ideals in a single objective function:

$$Conductance(C) = min_{A \subset C} \frac{cut(A, \bar{A})}{volume(A) volume(\bar{A})}, \qquad A \cup \bar{A} = C$$

Given a cluster (represented by a graph), *NJW* finds its conductance approximately by solving the minimization problem for A and $\bar{A}$.

Select: Given a set of clusters with different levels of conductance, the one with the lowest conductance has the most clearly identifiable sub-clusters. Clearly, it is the most suitable candidate selected for partitioning.

Partition: If A and $\bar{A}$ is the solution that solves for conductance, then this means that it is the best way to partition C so that $cut(A, \bar{A})$ is small while $volume(A)$ and $volume(\bar{A})$ large – divergence and convergence, exactly our requirements for a good partition. Thus, A and $\bar{A}$ is the way we partition C.

Flag: We want to flag organizations that *configure designs into styles* – these organizations would constitute a hierarchy of styles, our desired style construct. There are three requirements for such organizations – *convergence*, *divergence*, with the clusters having *sufficiently narrow concepts* (refer to the definition under **Formal Properties of a Hierarchy of Styles**).

The measure of conductance, with a minor tweak, can inform us whether an organization is *convergent* and *divergent*. However, because the algorithm does not *visually perceive* designs as a human do, it cannot inform if an organization contains clusters with sufficiently narrow concepts. Thus, we should interpret the flagged organizations as *potentially* organizing designs into styles until they are validated (in **Is the Algorithm Flagging Correctly?**Error! Reference source not found.).

The approach to flagging is best described by an example - suppose we have a dataset that we know has three clusters, with the three clusters each containing two sub-clusters each. We run it through the algorithm, and it generates $o_1, o_2, o_3, \ldots$ (the subscript represents the number of clusters in the organization). Our goal is to flag o_3 (representing the coarsest set of styles) and o_6 (the most fine-grained set of styles) as possible solutions (see **FIGURE 1** for a picture).

Insert **FIGURE 1** Here

The clue lies in tracing the path of conductance as the algorithm progresses. Focusing on o_1, we can see that the single big cluster has a low conductance because it has 3 clearly seen sub-clusters. Focusing on o_2, there is still a big cluster with a low conductance because it has 2 clearly seen sub-clusters. However, the moment we reach o_3, all 3 clusters have a higher conductance compared to that seen in o_1 and o_2.

The example suggests that a good organization is indicated by a high conductance over *all its clusters*. Suppose we track the conductance of the *worst* cluster (i.e. lowest conductance) as the algorithm progresses, a *jump* in this value would correspond to finding a good organization. This jump can be more easily observed if we take differences[1] (i.e. conductance of the worst luster in o_2 minus the conductance of the worst cluster in o_1, and so on). Observing *local peaks* in this difference as the algorithm progresses is equivalent to observing *jumps*.

---

[1] This difference approach is termed the *eigengap* heuristic (von Luxburg, 2007). See the same work for a more formal exposition of the method.

Outcome of the Algorithm

Using a smoothing technique – taking the difference in conductance over 100 steps of the algorithm instead of just 1 step – we plotted these differences as the algorithm progresses (see **FIGURE 2**). We picked 5 organizations (denoted O_1 to O_5, identified by the red dots) that corresponded to the *local peaks* in the graph. They were also chosen because they are far apart and roughly *equally spaced*, as conductance roughly doubles at each subsequent configuration.

Insert **FIGURE 2** Here

Conditions for the algorithm to obtain a hierarchy of styles

In this section, we bring the theory of styles and the algorithmic approach together by asking the question: *under what conditions* does the algorithm obtain our desired construct: a hierarchy of styles? We will show formally the *necessary and sufficient* conditions required of the algorithm. These conditions would later form the basis of validation.

Intuitively, the algorithm is driven by the process of *select →partition →flag*. All three steps need to be "correct", which we will first explain using simple examples before formal exposition.

Select: Imagine we have two clusters. We know one cluster contain designs from five different styles, while the other contains designs from two different styles. In this case the algorithm is allowed to select *either* cluster. However, suppose the second cluster, instead of containing designs from two different styles, contains designs only from one style (even though the designs can be further divided into sub-styles). In this case the algorithm must select the first cluster.

Partition: Consider the cluster described above that contains designs from five different styles. The algorithm is allowed to partition it into two sub-clusters with one containing designs from four different styles and the other containing designs from one style, or one with three styles and the other two styles, or any other such permutations. The restriction is that the algorithm should not partition it into fractions (e.g. one containing 4.5 styles and the other only 0.5).

Flagging: If the algorithm selects and partitions correctly, then a hierarchy of styles lies within the output of the algorithm. All that is needed here is to flag organizations that satisfy the properties of divergence, convergence, and containing clusters with sufficiently narrow concepts.

The implication: the theory here suggests that there is *flexibility* in how an algorithm can obtain styles – e.g. there is nothing wrong with an algorithm that identifies styles one by one versus one that tries to arrange the data as a whole. The requirement is that the *hierarchy be respected*, i.e. the algorithm should *not* look for sub-styles if not all main styles are identified.

Formal Requirements of the Algorithm: Given a set of *n* designs, an iterative algorithm *A* generates a set of organizations $\{o_1, o_2, \ldots, o_n\}$ with the following properties:

1) Each successive organization has one additional cluster.
2) Each successive organization is a partition of the one before.

Thus, the subscript *i* in the organization o_i represents the number of clusters of that organization.

We say that *A contains* a hierarchy of styles if $\{\Omega_1, \Omega_2, \ldots, \Omega_{k-1}, \Omega_k\} \subseteq \{o_1, o_2, \ldots, o_n\}$.

The following proposition identifies the requirements of these two steps to ensure that the algorithm obtains a hierarchy of styles.

For ease of exposition, we define the following variable and functions:

1) $\epsilon_0 = \infty$.
2) *Heterogeneity* of a cluster: $H(C) = max_{D_A, D_B \in C} f(D_A, D_B)$
3) *Partition quality* of a cluster: for a cluster *C* that is partitioned into C_A, C_B, define the partition quality as $Q(C_A, C_B) = min_{D_A \in C_A, D_B \in C_B} f(D_A, D_B)$

4) For any organization o_i, define its *stage* s as an integer value such that $0 \leq s < k$. An organization is in stage s if the following two conditions hold:
    a. No cluster in o_i is more heterogeneous than ϵ_s
    b. There exists a non-empty set of clusters C_s such that clusters in the set are more heterogeneous than ϵ_{s+1}.

Proposition 5: *A* obtains a hierarchy of style *if and only if* the following two conditions hold:

1) *Correct selection*: for any organization o_i in stage *s*, the next cluster selected for partition must be within the set C_s.
2) *Correct partition*: for any cluster $C \in C_s$ that is partitioned, the sub-clusters C_A, C_B generated must achieve at least a quality of $Q(C_A, C_B) > \epsilon_{s+1}$.

VALIDATION WITH EXPERIMENTS

Having established the conditions required for an algorithm to obtain a hierarchy of styles, the goal of this section is to find out if the algorithm actually meets those conditions. Because styles are *concepts* that are *visually recognized* (see **Style as a visually recognizable concept**), our validation approach involves comparing how the algorithm "perceives" styles and how humans perceive it. We outline below the key ideas for testing if the algorithm *selects, partitions*, and *flags* correctly.[2]

Select: Theoretically, an algorithm that selects correctly would prioritize selection of clusters with *a larger number of styles* in it than clusters with *a smaller number of styles* in it. We test if this is true by comparing *the order of selection* of a cluster and *the number of styles perceived* in

[2] Technically it is impossible to *prove* that the algorithm is correct using statistical means. However, we can statistically *reject* algorithms that fail these conditions.

the cluster by humans. The order of selection and the number of styles of a cluster should be *negatively correlated*.

Partition: Here, we want to see if the algorithm *mimics* the way human partitions designs into styles. Two experiments are needed. First, we need to show that the algorithm does not generate outcomes *by chance*. To do this, we compare the partition outcomes of the algorithm with partition outcomes generated by human beings, and show that they are more similar than a *random benchmark*. Second, we need to show that the algorithm does not generate outcomes that are *systematically different* from human beings. To do this, we ask subjects to judge if they can *detect* if an outcome is generated by a *machine*.

Flag: In this experiment, we test if humans would *consider clusters contained in an organization as styles*. If they do so, then the organization in question configures designs into styles and forms part of our desired style construct.

Before we go into the individual experiments, we discuss some *principles of sampling* that requires attention due to the nature of the clusters obtained, and also *the subject pool* that we used.

Principles of Sampling

The clusters obtained from the algorithm are *unevenly sized.* In fact, the cluster size distribution looks like it follows a *power-law distribution* – a commonly observed phenomenon in real-world networks and not unexpected of in styles. To ensure that we do not bias our results against larger clusters, in all the experiments we adopted a sampling technique that *weighs each design equally*. Also, the uneven size suggests that clusters may be a mix of large and small styles (e.g. a cluster may contain a large style covering 80% of the designs, and other smaller styles covering the remaining 20%). To efficiently uncover the *heterogeneity* of a sampled cluster, we run the algorithm on the sampled clusters (which would identify two parts of the cluster that are most different), and sample designs from each of the two parts.

Each sample (for all the experiments that follow in this paper) has a size of 10 designs. This number is arbitrary – it can vary within a small range (say 6 to 16) – but is chosen to balance between the cognitive load required of subjects while maintaining a minimum complexity.

Steps of Sampling:

1) An organization (from the five organizations identified earlier) is selected
2) A design (from the entire body of designs) is randomly selected. The cluster at which the design belongs to is noted.
3) Starting from the configuration selected in Step 1, we trace back the dendrogram to look for *the first instance* at which the cluster is partitioned. The cluster prior to partitioning is the *sampled cluster*.
4) Five designs are randomly selected from each side of the partition of the sampled cluster to create a sample of 10 designs.

A diagrammatic representation of the sampling process is included in **FIGURE 3**.

Insert **FIGURE 3** Here

Subject Pool

In all the experiments the subject pool is obtained from *Amazon Mechanical Turk*. Based on a demographic study by (Paolacci, Chandler, & Ipeirotis, 2010), the majority of Turk workers are based in the US (47%) followed by India (34%). There are significantly more females (65%) than males (35%). Turk workers also self-report to be younger (median age 36), have a higher educational level, but a lower income level compared to the US population. Despite these differences, their experiments show that Mechanical Turk workers show the “classical heuristics and biases" and pay attention to directions at least as much as subjects from traditional sources (i.e. universities and internet boards).

Is the Algorithm Selecting Correctly?

Idea of Experiment: The hypothesis we want to test is:

H_1: The number of styles contained in a cluster is negatively correlated with the order of which it is selected by the algorithm.

Design of Experiment testing H_1: For this experiment, a total of 25 clusters are sampled representing a total of 250 designs. 20 subjects responded to our survey. Each respondent is presented with 5 samples (randomly selected and presented in random order, at any point in time a respondent sees only a single sample of 10 designs in front of him/her). The respondents were given the following instructions:

"In each of the following questions you will be shown a set of 10 product designs. Your task is to categorize the designs into *groups with distinct styles*. You can form *as many (or as few)* groups as you like, and the groups *need not be of the same size*." (See **Appendix D** for a sample)

Results & Discussion: The specification for OLS is:

$$Groups\ formed = \beta_0 + \beta_1 Order\ of\ Cluster + \sum_{i=1}^{k} \beta_i\, Subject + \epsilon$$

The regression yielded $\beta_1 = -6.5 * 10^{-4}$. This coefficient is negative and statistically significant ($p < 0.01$ one-tailed, based on errors clustered around samples). The interpretation of this coefficient is that clusters *on average* contain one less style after every 1,550 partitions[3]. The negative coefficient suggests that the algorithm does not violate the ordering correctness criterion.

[3] Note that β_1 may be attenuated and the actual drop in the number of perceived styles is faster than that. This is due to *normalization* of perception, i.e. subjects observe finer differences when presented with designs that are stylistically closer. See Tversky & Gati (1978) for evidence in the psychology literature.

Is the Algorithm Partitioning Correctly?

Idea of experiments: Here, we want to see if the algorithm *mimics* the way human partitions designs into styles. We need to show two things: first, that the algorithm is *better than chance* in generating outcomes that is close to human beings; second, that the algorithm does not generate outcomes that is *systematically different* from human beings.

Formally stated, our hypotheses on the algorithm outcomes are:

> *H_2: The outcome generated by the algorithm is closer to outcomes generated by humans compared to a set of randomly generated outcomes.*
>
> *H_3: The outcome generated by the algorithm cannot be distinguished from outcomes generated by humans.*

Variation of Information: H_2 requires a way to measure the distance between two ways of organizing the same set of designs. For this, we leverage on *Variation of Information Index (VI)* (Meila, 2003), chosen amongst the many possible ways of comparing organizations because it is a *metric* and it makes *no prior assumptions* on cluster sizes. The technical details of the VI are left in **APPENDIX E**. Its intuition is best gained through an example: suppose the algorithm configures a set of 10 designs into two clusters $\{D_1, D_2, D_3, D_4, D_5\}$ and $\{D_6, D_7, D_8, D_9, D_{10}\}$. If a human subject configures it *the same way*, then knowing how the algorithm configures tells you exactly how the human subject configures, i.e. there is no *information gap* $(VI = 0)$. Suppose instead the human subject configures a little differently, say by switching D_5 and D_{10}. Knowing how the algorithm configures and knowing that the subject would keep 4 out of 5 designs together in the same clusters is not enough to describe how subject behaves – an information gap exists $(VI = 1.01)$ because we would still need to know *which* designs get switched.

Design of Experiment testing H_2: 25 samples are created (5 samples for each organization, representing a total of 250 designs). Each sample is composed of 10 designs that were partitioned by the algorithm into 2 clusters of 5 designs.

22 subjects responded to our survey. Each respondent is presented with 5 samples (randomly selected and presented in random order, at any point in time a respondent sees only a single sample of 10 designs in front of him/her). The respondents were given the following instructions: "In each of the following questions you will be shown a set of 10 product designs. Your task is to categorize the designs into *2 groups* (with each group containing exactly 5 designs), so that each group contains designs *that have a distinct style.*" (See **Appendix D** for a sample)

Results & Discussion: Statistics for the random benchmark is obtained through bootstrapping the VI score between 2 randomly generated organizations. Given the small sample size, the VI score distribution is highly discrete and asymmetric. Thus, we make no assumptions of normality. Instead, the significance measures reported in **TABLE 1** are done through bootstrapping the *distribution of the means* of the VI score, for a total of 10,000 rounds.

Insert **TABLE 1** Here

Note that the mean VI score for O_1 to O_4 are all significantly below the random benchmark at $p < .01$, while the VI score for O_5 is significantly below the random benchmark at $p < .02$.

A possible weakness of this method is the *assumption of independence* – the experiment is designed such that multiple readings are obtained from the same sample, which may introduce correlation between readings. We conducted a statistical test assuming the *worst case*, i.e. all readings using a sample are *perfectly correlated*, meaning that multiple readings of the same sample are uninformative and do not improve significance. Using this assumption, the significance level did not change for O_1 to O_3. For O_4 it is significant at $p < .05$, but for O_5 it is not significant ($p = .13$).

The high mean VI score for O_5 is indicative that the presented designs are all stylistically very close and differences are hard to perceive. Imagine being given a set of 10 designs of the same style and asked to configure them into two groups of 5 each. Lacking any clear way of organizing we will observe random outcomes. The outcome suggests that O_5 is *too fine*, and that there is no reason to execute the algorithm further to obtain finer styles.

Is the algorithm systematically different from humans? The motivation to ask this question comes from the Turing Test, originally devised by Alan Turing (1950) as a replacement for the question “Can machines think?” – see (Pinar Saygin, Cicekli, & Akman, 2000) for a discussion of the recent works on the Turing Test. In this context, the algorithm obviously *cannot* think *independently* – it requires as input stylistic similarities generated by human beings. What we wish to find out here is the answer to a more limited question: given such inputs, can the algorithm configure designs in a similar way as a human being?

The previous experiment gave us an indication that the algorithm is close to human behavior (at least closer than a random benchmark), but there is the possibility that, while close, the algorithm exhibit *deviant* behavior that is systematically different from human beings. If such systematic differences exist, then human subjects would be able to *detect* such differences. To find out, we asked human subjects *if they can tell* if an organization is generated by the algorithm.

Design of Experiment testing H_3: Samples for this experiment come from the output of the previous experiment. For each unique sample of 10 designs, we randomly sample organizations generated by human beings. Note that we often observe that humans configure *exactly the same way* as the algorithm (in fact, out of the 25 samples, we observe complete agreement between how humans configure and how the algorithm configures in 5 samples). We combine these 25 human-generated organizations with the 25 corresponding algorithm-generated organizations for a total of 50 samples.

20 subjects responded to our survey. Each respondent is presented with 10 samples (randomly selected and presented in random order, at any point in time a respondent sees only a single sample of 10 designs in front of him/her). The respondents were given the following instructions: "In each of the following questions you will be shown a set of 10 product designs. The designs have previously been given to both *a human* and *a machine*, tasked to categorize the designs into two groups (of 5 designs each) so that each group has a *distinct style*. In each question, we randomly select one of the outcomes of this categorization exercise and present it to you. Your task is to decide whether this outcome is generated by a human or a machine." (See **Appendix D**)

Results & Discussion: to ascertain if an algorithm generated outcome is associated with a higher probability of being judged as a "machine-generated outcome", we ran a logit model specified below:

$$P(Response\ is\ Machine) = \frac{\exp(\boldsymbol{\beta X})}{1 + \exp(\boldsymbol{\beta X})},$$

$$\boldsymbol{\beta X} = \beta_0 + \beta_1 Algorithm\ Generated + \sum_{i=1}^{k} \beta_i\ Subject$$

The regression yielded $\beta_1 = -2.6$. This coefficient is negative and not statistically significant (p = 0.44, clustered errors by sample), suggesting that an actual algorithm-generated outcome *does not* have a higher probability of being identified as such. Thus, we conclude that the algorithm is *human-like*, and since both H_2 and H_3 are not rejected, we conclude that the algorithm does not violate partitioning correctness.

Is the Algorithm Flagging Correctly?

The previous sections showed that the algorithm "perceive" styles in a way that is not different from how humans would perceive it. Thus, O_1 to O_5 can potentially form a hierarchy of styles. In this section we discuss the experiment to identify *which* of these organizations contain clusters that have *sufficiently narrow concepts* to be considered styles.

One complication in this experiment is that styles are *existing* concepts in the perceiver's mind. Culturally different people may have different ways of considering styles – for example, Uzbekistan tribesmen who are more attuned to nature than consumer products would associate round shapes to a full moon (Luria, 1976). While unlikely, this would mean that a proportion of the population would fundamentally consider styles differently and reject *any* organizations obtained here as styles.

One way to *gain objectivity* is to capture how *the largest group* of the population considers styles. A priori, without knowing the *size* of this group, we can adopt a *simple majority rule* – that is, we rule that an organization contains styles if *more than 50% of the population* judges its clusters to be styles. Note that achieving 50% may not be possible if the population is composed of groups with very different views of what a style is. However, if achieved, it is a *sufficient* (but not *necessary*) condition to establish that an organization represents how the largest group of the population considers as styles.

The idea of the experiment is to present designs sampled from a cluster to observers, and have them judge if the designs constitute a style. If our assumption that *a majority of the population have the same way of perceiving styles* holds, then we should be able to observe *a greater than 50% agreement* that the clusters are styles.

Design of Experiment: 10 clusters are sampled for each organization, representing a total of 500 designs (5 organizations * 10 clusters * 10 designs per cluster). 89 subjects responded to the survey. Each respondent is presented with 10 questions. For each question, respondents are shown 10 designs from a cluster (randomly selected and presented in random order, at any point in time a respondent sees only a single sample of 10 designs in front of him/her).

The respondents were given the following instructions: "In each of the following questions you will be shown a set of 10 product designs. The designs could be taken from a single style or from multiple styles. Your task is to answer: do you agree that the designs are from a single style?" The answers are rated on a "Yes-No" scale. (See **Appendix D** for a sample)

Results & Discussion: **TABLE 2** shows the summary statistics for the 5 organizations. At O_3 to O_5, more than 50% of subjects agree that the clusters are styles (significant at the 0.05 level, tested using a binomial model for the location of the *median* response).

Insert **TABLE 2** here

The results suggest that we should constitute a hierarchy of styles using O_3, O_4, and O_5. However, note that the partitioning experiment indicated that humans cannot differentiate clusters in O_5, suggesting that clusters may lack divergence – see **Is the Algorithm Partitioning Correctly**. This indicates that O_3 and O_4 should be used to constitute a hierarchy of styles.

CONCLUSIONS

Chic, Zen, clean, classical…. The element of style is part of every design, and an analysis of style can profoundly impact the way we view designs and the design creation process. Yet, an analysis of style has not been possible because *styles have not been identified empirically in designs*. In this paper we proceed in a step-by-step manner to identify styles empirically by first building a theoretical foundation on *what a style is* and *how designs are configured into styles*.

We then execute an *algorithm* to identify styles in designs, and lastly use *experiments* to show that the algorithm obtains styles.

The *notion* of style is complex – while a style is embodied in the *physical characteristics* of a design, the set of physical characteristics that is *salient* to the style depends on how it is *perceived* by an observer. We argue that styles are *visually recognizable concepts*, or *mental representations* that help us gain a coherent understanding of the *meaning* of designs. The concept of a style is composed of complex associations with values, emotions, and the time and place where the style is encountered. This *conceptualization* helps to simplify understanding, and leads to the dichotomy that we observe in styles – complex to define physically, but intuitive and instantaneous to recognize.

We argue that the best way to configure designs into styles is through a *hierarchy*. Within this hierarchy, designs with similar concepts are placed in the same group, while designs with different concepts are placed in different groups. At the top of the hierarchy, where the groups are largest, each group should still have a *sufficiently narrow concept* to be recognized as a style.

Using the complete design patent database from the US Patent & Trademark Office, we proceeded to *find* this hierarchy using a graph clustering algorithm. The design patent is uniquely suitable for identifying styles because it is purely focused on how things *look*. In addition, it has references which *convey the stylistic similarity* between designs. The graph approach to clustering, while complex, is integral here because the *input* to clustering (the stylistic similarity between designs) is a graph.

We pull together the theory of styles and the algorithmic approach in our design of experiments to *validate* the outcome of the algorithm. We obtain from our theoretical construct *necessary and sufficient conditions* that the algorithm has to meet to obtain styles. The outcomes of the experiments validating these conditions suggest that the algorithm is capable of *selecting* the right cluster to partition, and then proceeds to partition it in a manner that is *akin* to how a human would partition. We end our analysis by finding the set of organizations that would represent a hierarchy of styles.

This work uses *design patents* as the body of designs from which styles are formed. A large stream of research in innovation used patent data –there is a wealth of papers that use the idea of the *idea (the patent)* as a unit of analysis, a wealth of papers that use the *inventor*, the *firm*, or the *industry* as a unit of analysis. Each of these streams has become the foundation of empirically validated theory of different managerial and economic realities. But there is no work that analyzes patents at the level of *style*. This unit of analysis has not been available in the past. This work aims to change that. The study of styles would have profound impact on our understanding of designs and the design process, yet an empirical approach to study styles has not been possible because there is no validated style construct that could be applied to the USPTO database. By combining knowledge from the fields of art, engineering, psychology, and artificial intelligence, we built an empirical basis from which styles can now be analyzed.

REFERENCES

Burdek, B. E. (2005). *Design: History, theory and practice of product design.* Birkhauser.

Chan, C. S. (2000). Can style be measured? *Design Studies*, 21(3): 277-291.

Chung, F. (1997). *Spectral Graph Theory.* American Mathematial Society.

Gardner, H., & Kleiner, F. (2009). *Gardner's Art Through the Ages: The Western Perspective, Volume 2.* Wadworths Publishing Company.

Hanna, S. (2006). Representing style by feature space archetypes. *Design Computing and Cognition*, 3-22.

Jaccard, P. (1901). Etude coparative de la distribution florale dans une portion des alpes et des jura. *Bulletin de la Socit Vaudoise des Sciences Naturelles*, 37: 547-579.

Jain, A., Murty, M., & Flynn, P. J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3): 264-323.

Jupp, J., & Gero, J. (2006). Visual style: Qualitative and context-dependent categorization. *AIE EDAM*, 20(03): 247-266.

Kessler, M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 15: 10-25.

Koning, H., & Eizenberg, J. (1981). The language of the prairie: Frank Lloyd Wright's prairie houses. *Environment and Planning B*, 8(3): 295-323.

Luria, A. R. (1976). *Cognitive Development: its Cultural and Social Foundations.* Harvard University Press.

McCormack, J., Cagan, J., & Vogel, C. (2004). Speaking the Buick language: capturing, understanding and exploring brand identity with shape grammars. *Design Studies*, 25(1): 1-29.

McMahon, J. (2003). Perceptual constraints and perceptual schemata: the possibility of perceptual style. *The Journal of Aesthetics and Art Criticism*, 61(3): 259-273.

Meila, M. (2003). Comparing clusterings by the variation of information. *Learning theory and kernel machines*, 173-187.

Milligan, G. (1996). *Clustering and Classification.* New Jersey: World Scientific Publisher.

Munro, T. (1946). Style in the arts: a method of stylistic analysis. *Journal of Aesthetics and Art Criticism*, 128-158.

Ng, A., Jordan, M., & Weiss, Y. (2002). On spectral clustering: analysis and an algorithm. *Advances in neural information processing systems*, 2: 849-856.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: the role of global image features in recognition. *Progress in brain research*, 155, 23.

Pachys, F. (2012). *Patent citation: the inventor, examiner, application, differences. USPTO - case report: Patent no 4,237,224.* Unpublished Manuscript.

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running Experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5: 411-419.

Pinar Saygin, A., Cicekli, I., & Akman, V. (2000). Turing Test: 50 Years Later. *Minds and Machines*, 10(4): 463-518.

Potter, M. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2(5): 509.

Pugliese, M., & Cagan, J. (2002). Capturing a rebel: modeling the Harley-Davidson brand through a motorcycle shape grammar. *Research in Engineering Design*, 13(3): 139-156.

Stacey, M. (2006). Psychological Challenges for the Analysis of Style. *Artificial Intelligence in Engineering Design, Analysis, and Manufacturing*, 20(2): 167-184.

Stiny, G. (1980). Introduction to shape and shape grammars. *Environment and Planning B*, 7(3): 343-351.

Stiny, G., & Mitchell, W. (1978). The Palladian Grammar. *Environment and Planning B*, 5(1): 5-18.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59(236): 433-460.

Tversky, A., & Gati, I. (1978). Studies of Similarity. *Cognition and Categorization*, 1: 79-98.

USPTO. (2006). *Manual of Patent Examining Procedures, 8th Ed.* United States Patent & Trademark Office.

Verma, D., & Meila, M. (2003). Comparison of spectral clustering methods. *Advances in Neural Information Processing Systems*, 15.

von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395-416.

Zahn, C. T. (1971). Graph-theoretical methods for detecting and describing gestalt clusters. *IEEE Transactions Computing*, C-20: 68-86.

TABLES AND FIGURES

TABLE 1
Mean VI Between Algorithm-generated outcomes and human-generated outcomes

Organization (# of Clusters)	$\overline{VI}(Algorithm, Human)$
O_1 (3,129)	0.777**
O_2 (5,749)	0.869**
O_3 (9,690)	0.990**
O_4 (15,463)	1.037**
O_5 (22,065)	1.175*
* p < .05 ** p < .01	$\overline{VI}(Random, Random) = 1.267$

TABLE 2
Percentage of Subjects who responded "yes - designs from a cluster forms a style"

Organization (# of Clusters)	% responding yes
O_1 (3,129)	42%
O_2 (5,749)	36%
O_3 (9,690)	61%*
O_4 (15,463)	64%*
O_5 (22,065)	68%*
* p < .05, ** p < .01	

FIGURE 1
A depiction of how the algorithm progresses (o_4, o_5 not shown)
(Graph edges are hidden for clarity - longer distance implies weaker relationships)

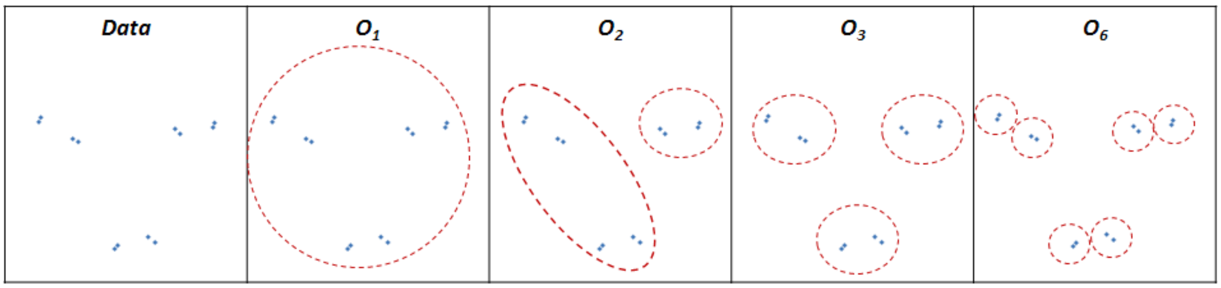


FIGURE 2
A Chart of the Difference in Conductance (Cond) against the Number of Clusters Generated

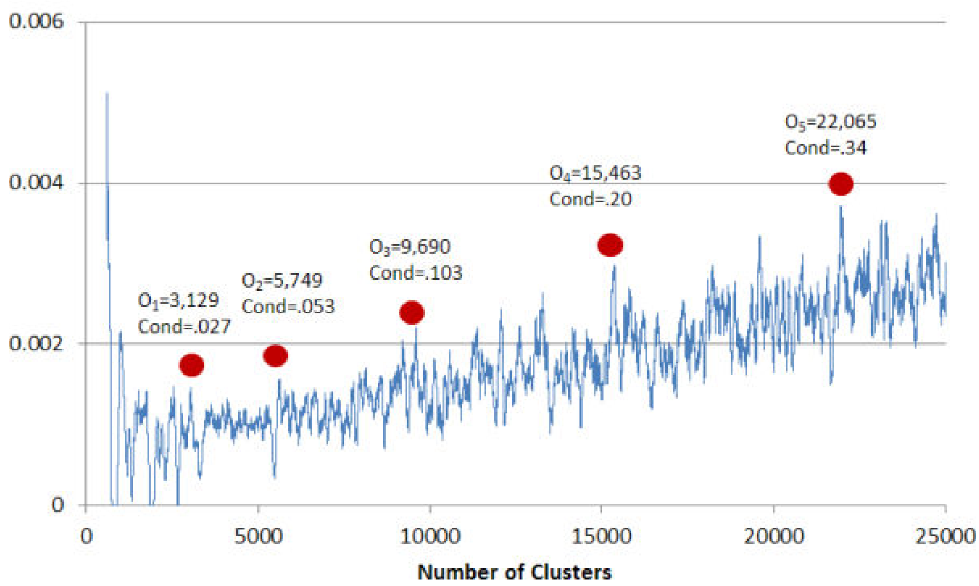
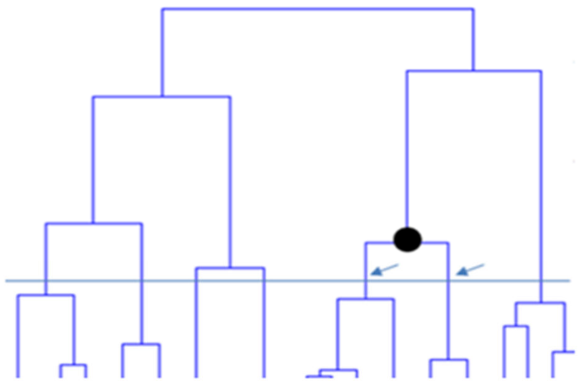


FIGURE 3
A diagrammatic representation of the sampling process: 5 designs are sampled from each of the clusters represented by the arrows, to form a sample of 10 designs for the cluster represented by the black dot.



APPENDIX A

SAMPLE OF A DESIGN PATENT

US00D329019S

United States Patent [19]

Olds et al.

[11] **Patent Number: Des. 329,019**

[45] **Date of Patent: ** Sep. 1, 1992**

[54] **BOTTLE**

[75] Inventors: **Sandra H. Olds**, Lilburn; **Clifford R. Wade**, Tucker, both of Ga.

[73] Assignee: **The Coca-Cola Company**, Atlanta, Ga.

[**] Term: **14 Years**

[21] Appl. No.: **521,472**

[22] Filed: **May 9, 1990**

[52] **U.S. Cl.** **D9/551**

[58] **Field of Search** .... D9/392, 406, 390-399, D9/349, 355; 215/1 R, 1 C

[56] **References Cited**

U.S. PATENT DOCUMENTS

D. 109,928	5/1938	Williams	D9/392
D. 139,564	11/1944	Byrd et al.	D9/392
D. 193,638	9/1962	Goldstein	D9/392
D. 216,096	11/1969	Lawton	D9/392
D. 226,703	4/1973	Broadhead	D9/406

OTHER PUBLICATIONS

Packaging, Sep. 1987, p. 13, Returnable Coca-Cola Pet Bottle, Continental Can Co.

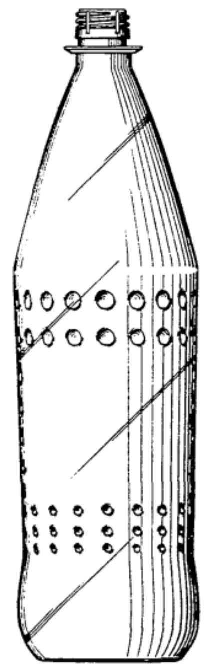
Primary Examiner—Bernard Ansher
Assistant Examiner—Lucy Lieberman
Attorney, Agent, or Firm—Lynne R. O'Brien; W. Dexter Brooks; William C. Lee, III

[57] **CLAIM**

The ornamental design for a bottle, as shown.

DESCRIPTION

FIG. **1** is a top front perspective view of a first embodiment of a bottle showing our new design;
FIG. **2** is a front elevational view thereof;
FIG. **3** is a rear elevational view thereof;
FIG. **4** is a right side elevational view thereof;
FIG. **5** is a left side elevational view thereof;
FIG. **6** is a top plan view thereof;
FIG. **7** is a bottom plan view thereof;
FIG. **8** is a top, front perspective view of a second embodiment of the bottle;
FIG. **9** is a front elevational view thereof;
FIG. **10** is a rear elevational view thereof;
FIG. **11** is a right side elevational view thereof;
FIG. **12** is a left side elevational view thereof;
FIG. **13** is a top plan view thereof; and,
FIG. **14** is a bottom plan view thereof.



APPENDIX B

PROOF OF PROPOSITIONS

Proof to Proposition 2: Uniqueness

We approach the proof by *contradiction*. Suppose there exist two ways of organizing the same set of designs (denote as Ω_1 and Ω_2) given a threshold value ϵ_T. If Ω_1 and Ω_2 are different, then there must exist pairs of designs D_A and D_B such that (1) D_A and D_B are in the same cluster under Ω_1 and (2) D_A and D_B are in different clusters under Ω_2 (or the reverse). (1) implies that $f(D_A, D_B) \leq \epsilon_T$ and (2) implies that $f(D_A, D_B) > \epsilon_T$, a contradiction.

Proof to Proposition 3: Nested-ness

To prove this proposition we just need to prove that given two organizations Ω_1 and Ω_2, corresponding to thresholds $\epsilon_1 > \epsilon_2$, that Ω_2 *partitions* Ω_1. This can again be proven by *contradiction*. Suppose the converse is true, that Ω_2 *does not* partition Ω_1, then there must exist pairs of designs D_A and D_B such that D_A and D_B are in the same cluster under Ω_2 (implying $f(D_A, D_B) \leq \epsilon_2$), but in different clusters under Ω_1 (implying $f(D_A, D_B) > \epsilon_1$). This is a contradiction.

Proof to Proposition 4: $SI(a, b) \in [0, 1]$

First, we note that the graph formed by references is a *directed-acyclic graph (DAG).*
Consider the following cases:

1) $R(a, b) = 0$ (no references to each other). Then $SI(a, b) = J(a, b)$, and $J(a, b) \in [0,1]$.
2) Without loss of generality, consider $R(a, b) = 1$ (a refers to b). Then $N(a, b) = 1/|A|$.

Since the graph is a *DAG*, b does not refer to a, and b does not refer to itself. Thus, $b \in A \backslash B$

We can express $|A \cup B| = |A \cap B| + |B \backslash A| + |A \backslash B|$, so $J(a, b) = \frac{|A \cap B|}{|A \cup B|} \leq \frac{|A \cap B|}{|A \cap B| + |A \backslash B|} = \frac{|A \cap B|}{|A|}$

Since b does not refer to itself, $b \notin B$, implying $|A \cap B| \leq |A| - 1$. Thus, $J(a, b) \leq \frac{|A|-1}{|A|}$.

This means that $SI(a, b) = |N(a, b)| + J(a, b) \leq \frac{1}{|A|} + \frac{|A|-1}{|A|} = 1$

Proof to Proposition 5: Necessary & Sufficient Requirements on Algorithm

We use a *direct* approach to prove this proposition. We will first prove the *sufficient* condition – algorithms that satisfy both partial ordering and partial partitioning correctness will obtain a hierarchy of styles, and then the *necessary* direction – algorithms that fail either criterion *will not* be able to obtain a hierarchy of styles.

Sufficient condition: We first prove that an algorithm with partial ordering and partial partitioning correctness will obtain Ω_1 (the first organization that contains styles).

Note that Ω_1 must satisfy:

1) *Convergence*: $f(D_A, D_B) \leq \epsilon_1, \forall D_A, D_B \in C_i, \forall i$
2) *Divergence*: $f(D_A, D_B) > \epsilon_1, \forall D_A \in C_i, D_B \in C_j, \forall\, i \neq j$

Note that an algorithm with partial ordering correctness only selects clusters with $H(C) = max_{D_A, D_B \in C} f(D_A, D_B) > \epsilon_1$ for partitioning. If the algorithm also satisfied partial partitioning correctness, then the minimum quality of the partition resulting in clusters C_A and C_B must be $Q(C_A, C_B) = min_{D_A \in C_A, D_B \in C_B} f(D_A, D_B) \geq \epsilon_1$.

This means that the divergence criterion of the algorithm is *always met* provided that there are clusters with heterogeneity $H(C) > \epsilon_1$.

As the algorithm continues partitioning, it will eventually reach an organization where no clusters are more heterogeneous than ϵ_1. By Proposition 2, we know that this organization is unique to ϵ_1 and thus must be Ω_1.

Because of Proposition 3, we know that Ω_1, $\Omega_2, \ldots, \Omega_k$ are nested, and because the algorithm reaches Ω_1, the same proof above can be used to show that it will reach Ω_2, and so on.

Necessary condition: Suppose we have an algorithm that does not satisfy partial ordering correctness. Without loss of generality assume that this failure to observe partial ordering correctness occurs before Ω_1. This means that the algorithm selects a cluster $H(C) \leq \epsilon_1$ for partitioning at some stage of the algorithm where there are still clusters with heterogeneity $H(C) > \epsilon_1$. If it does so, no matter how the clusters are partitioned, we will have two clusters C_A and C_B such that $f(D_A, D_B) \leq \epsilon_1, \forall D_A \in C_A, D_B \in C_B$. This violates the divergence property of Ω_1, and thus Ω_1 will not be reached.

On the other hand, suppose we have an algorithm that does not satisfy partial partitioning correctness. Again, without loss of generality assume that this failure to observe partial ordering correctness occurs before Ω_1. This failure would mean that the algorithm generated clusters C_A and C_B, with at least one pair of designs $D_A \in C_A, D_B \in C_B$ such that $f(D_A, D_B) \leq \epsilon_1$. Again, this violates the divergence property of Ω_1.

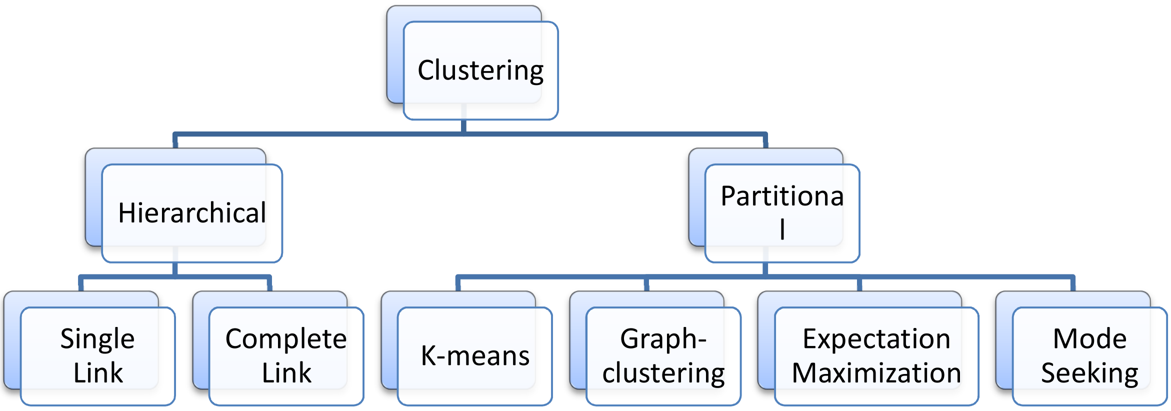
APPENDIX C

THE SPECTRAL CLUSTERING ALGORITHM

Why Use a Spectral Clustering Approach?

The clustering method uses the similarity index to find a set of partitions, such that intra-cluster similarity is maximized (convergence) while inter-cluster similarity is minimized (divergence). **FIGURE C1**, adapted from (Jain, Murty, & Flynn, 1999), show a general classification of clustering methods.

FIGURE C1
A Classification of Clustering Methods



The choice of a method is dependent on our underlying beliefs about cluster structures. Particularly, we desire a method that produces *nested* clusters. We also require a method that can take a graph – the similarity matrix – as an input. K-means, expectation maximization, and mode seeking methods typically require data to be *points in a metric space*. However, because graph data generally cannot be expressed as points in a metric space, we narrow our search to *hierarchical* or *graph clustering methods*. Hierarchical methods fall into a class of agglomerative heuristics whereby all clustering elements are separate clusters in the beginning, and at each step clusters are combined to form bigger clusters, until a predefined stopping point is reached. Graph clustering methods fall into a class of divisive heuristics that take the inverse approach. Treating the similarity matrix as a graph, it aims to find a way to partition the matrix into multiple parts.

We choose graph clustering methods over hierarchical methods because of *global consideration* – graph methods takes into account the structure of the entire data and decide how best to partition, whereas hierarchical methods starts at the local level, looking at nearest neighbors and joining them. The global-ness of graph theoretic methods makes it less sensitive to local errors. Graph clustering methods can further be divided into *heuristic* approaches and *optimizing* approaches. Heuristic approaches, such as the minimal spanning tree (MST) (Zahn, 1971), use simple rules of thumb, such as breaking edges with the weakest similarities. In comparison, optimizing approaches formalizes the clustering problem by first defining an *objective function*, and then find methods to solve or approximately solve for the objective function. Because the solution method typically requires an analysis of the *spectrum* (or the eigenvalues and eigenvectors of the adjacency matrix of the graph), such methods are also called *spectral clustering*. The choice of an optimizing approach over heuristics is driven by the clarity of the objective function, and thus of the nature of solutions obtained.

As for the choice of a specific spectral method, von Luxburg (2007) provided a comparative overview of the different methods of spectral clustering and recommends the use of *normalized spectral clustering*. Verma and Meila (2003) compared 4 popular spectral clustering algorithms and concluded that there is no clear winner in performance, but the *Ng-Jordan-Weiss* (Ng, Jordan, & Weiss, 2002) *algorithm* (NJW) – a normalized spectral clustering method – has more robust performance when the node degrees are highly variable. This is the case in our data: the most connected patent has a node degree of 126.5, and the least connected one has a node degree of 0.005. Thus, we choose the NJW algorithm as the clustering method.

Steps of the Algorithm

Adapted from (Ng, Jordan, & Weiss, 2002)

Input: Similarity matrix $S \in R^{n \times n}$

1) Compute the normalized graph Laplacian matrix $L = I - D^{1/2}SD^{1/2}$, where I is the identity matrix, D is a diagonal matrix with each entry the degree of the vertex.
2) Compute the 2 smallest eigenvectors (u_1, u_2)
3) Let $U \in R^{n \times 2}$ be the matrix containing (u_1, u_2) as columns.
4) Normalize the rows of U to unit lengths of 1.
5) Treat every row as a point in space. Cluster the n data points using k-means.

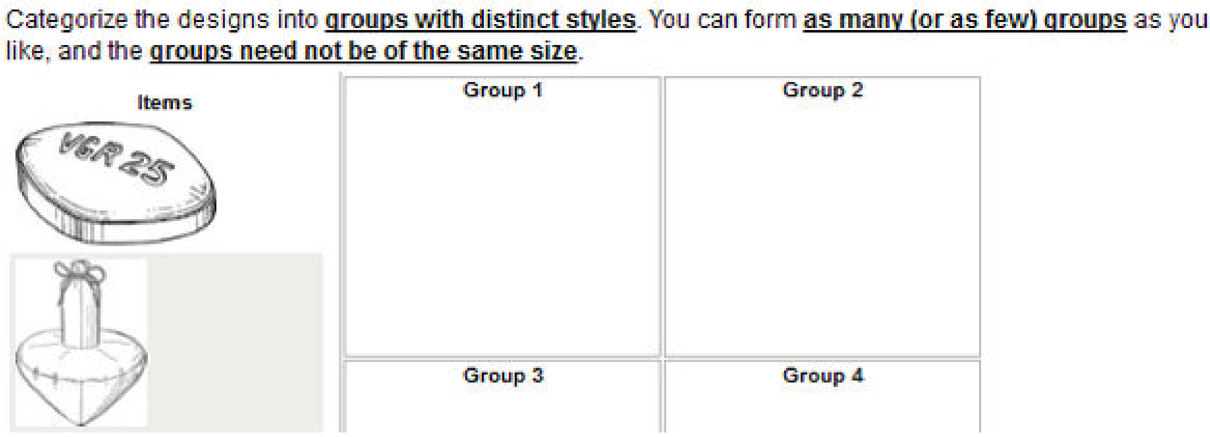
The idea of the method is best illustrated in the ideal case where there are *multiple components* in the graph (and thus the goal of the partitioning is to find these components). The *graph Laplacian matrix* is a key construct in spectral graph partitioning. It is a simple normalization of the similarity matrix. Chung (1997) showed that the smallest eigenvalue of the Laplacian matrix is always 0, and the multiplicity of this eigenvalue depends on the *number of components* the graph has. The interesting result is that a 1-0 vector with 1s corresponding to the locations of vertices in one of the component and 0s everywhere else is an eigenvector (readers can try this themselves). Thus, translated to the space of eigenvectors, the positions of 1s and 0s immediately identify the location of these components.

The algorithm *approximately* solves for the sub-graphs A and $\bar{A}$ such that conductance is minimized – for readers who are interested, see (von Luxburg, 2007) for a proof. Intuitively, the approximation is due to the algorithm relaxing the constraint that a design belongs to either A or $\bar{A}$, i.e. the solution assumes that designs can be, e.g. 80% from A and 20% from $\bar{A}$.

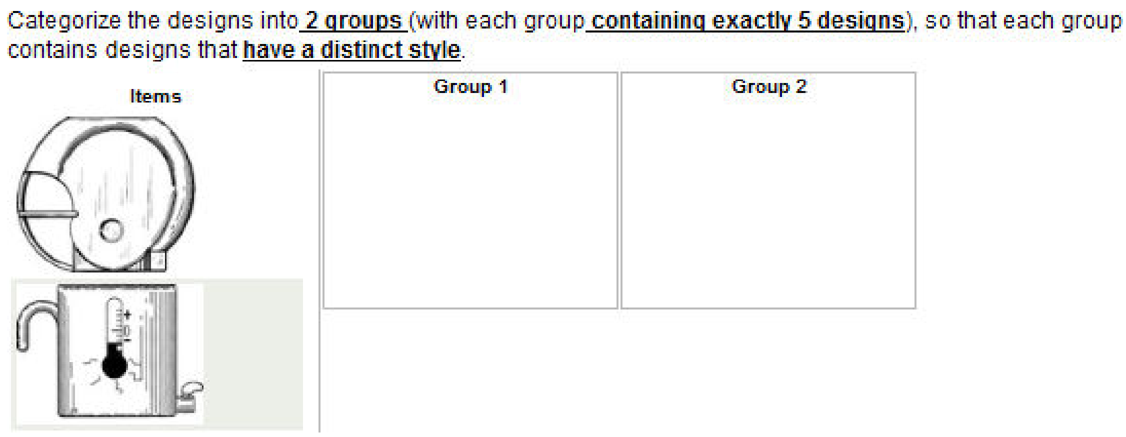
APPENDIX D

SAMPLE OF SURVEYS

1) Sample survey question for *Experiment 1 (Is the Algorithm Selecting Correctly)*.



2) Sample survey question for *Experiment 2 (Partitioning is non-random)*.



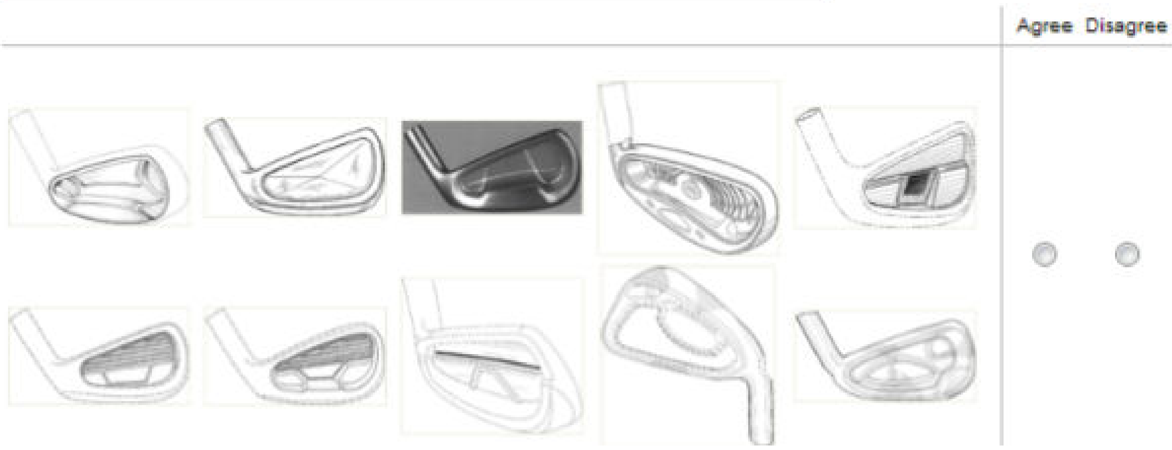
3) Sample survey question for *Experiment 3 (Partitioning is human-like)*.

The following 10 designs have been categorized into **two groups** (the top row of 5 designs forming one group, and the bottom row of 5 designs the other group). The categorization is done so that **each group has a distinct style**.

Do you think this categorization is done by a human being or a machine?

4) Sample survey question for *Experiment 4 (Cluster is a style)*.

Do you agree with the statement: these designs are from a **single style**?



APPENDIX E

VARIATION OF INFORMATION

Suppose we are given a set of designs *D*, and two organizations of *D* - $O_1 = \{C_1, \dots, C_K\}$and $O_2 = \{C'_1, \dots, C'_K\}$. The organizations can be compared using a *contingency table* (see **FIGURE E1**). Each cell in the table is a number n_{ij} representing the number of designs that is in the *intersection* of clusters C_i and C'_j.

FIGURE E1
A Contingency Table

Cluster	C'_1	C'_2	...	$C'_{K'}$
C_1	n_{11}	n_{12}	...	$n_{1K'}$
C_2	n_{21}	n_{22}	...	$n_{2K'}$
.	.	.		.
.	.	.		.
.	.	.		.
C_K	n_{K1}	n_{K2}	...	$n_{KK'}$

Formally, the VI is defined through the entropy associated with an organization. Suppose we randomly pick a design $e \in D$, the membership of *e* under O_1 can be seen as a random variable with *K* different outcomes, where *K* corresponds to the number of clusters in O_1. The probability of realizing a particular outcome *k* (the design falls into cluster C_k) is given by $P(k) = |C_k|/|D|$. Entropy measures the uncertainty of the outcome associated with this random variable. Formally, entropy for an organization is defined as $H(O_1) = -\sum_{k=1}^{K} P(k) log P(k)$.

The *VI* is the sum of the *entropy* about O_2 given that we know O_1 *and* the contingency table, and vice versa. Formally, $VI(O_1, O_2) = H(O_1|O_2, contingency) + H(O_2|O_1, contingency)$

Where $H(O_1|O_2, contingency) = \sum_{k'=1}^{K'} P(e \in C_{k'}) H(O_1|e \in C_{k'})$

$$= \sum_{k=1}^{K} \sum_{k'=1}^{K'} P(k, k') log \frac{P(k')}{P(k,k')}$$

Europe Campus
Boulevard de Constance
77305 Fontainebleau Cedex, France
Tel: +33 (0)1 60 72 40 00
Fax: +33 (0)1 60 74 55 00/01

Asia Campus
1 Ayer Rajah Avenue, Singapore 138676
Tel: +65 67 99 53 88
Fax: +65 67 99 53 99

Abu Dhabi Campus
Muroor Road - Street No 4
P.O. Box 48049
Abu Dhabi, United Arab Emirates
Tel: +971 2 651 5200
Fax: +971 2 443 9461

www.insead.edu

INSEAD

The Business School
for the World®

Printed by INSEAD