Q-Analysis and Human Mental Models: A Conceptual Framework for Complexity Estimate of Simplicial Complex in Psychological Space

Dr. Konstantin DEGTiAREV
Faculty of Business Informatics (BI)
School of Software Engineering (SSE)
The National Research University – Higher School of Economics (HSE)
Moscow / Russian Federation

Antalya / Türkiye | September 01-02, 2011
Understanding structural features of systems,

Revealing and understanding of such features are grounded on categories of the whole and its parts cognitive activity (observations, reasoning, development of models (formal/informal)),

System’s structure complexity,

The notion of complexity is many-sided and rich,

Classification of systems as simple or complex is based on several factors, and variety of elements and connections between them are among the most significant ones,

Analysis of systems structure(s)

Estimation of structure’s complexity

diverse mathematical methods,

- Presence of humans “on the scene” (active implicit use of their unintelligible cognitive processes like thinking, perceiving, making decisions, etc.),

- Cognitive modeling is not exclusively linked to knowledge fields concerned with «process of thought» in large – it utilizes mathematical and computer languages to describe and analyze particularities of human information processing,

- Simplicial complex can be represented as $K = (V, S)$, where $V$ is a finite set of vertices, and $S$ are simplexes of complex $K$:
  - $\forall \{u\} \subseteq S$ (any vertex is a simplex of complex $K$)
  - $\forall \sigma \in S, \forall \tilde{\sigma} \subseteq \sigma | \tilde{\sigma} \neq \emptyset, \tilde{\sigma} \in S$ (simplex $\tilde{\sigma}$ being a face of simplex $\sigma$ is also a simplex of complex $K$)
Simplicial complex K is formed by regularly adjoining faces called simplexes; the intersection of two simplexes is “either empty, or a common face of each”,

Simplex is a convex hull of its (q+1) vertices, q ≥ 0 → q-dimensional simplex (dim(σ) = q | σ_q ) or q-simplex in short; σ_0 is a point (vertex), σ_1 is a line segment, σ_2 - triangle (with its interior), σ_3 - tetrahedron, etc.

Dimension of K (dim(K)) → maximal dimension of K’s simplexes,

Simplicial complex K as an aggregate of simplexes of different dimensions → formal representation (model) of the system under study,

Q-analysis (R. Atkin): analysis of complex K is performed consecutively at each dimensional level q, q = dim(K), ..., 1, 0, through determining the number of clusters of simplexes joined by chains of q-connectivity
The usual notion of *connectivity*—*q-connectivity* in Atkin’s approach (complex is viewed as topological space),

- **Q-analysis** is aimed at discovering multidimensional chains of connectivity (i.e. *q-connectivity components* formed by simplexes of particular dimensions at each level *q*, *q* = *dim(K)*, …, 1,0),

- Two simplexes of complex *K* are said to be *q-near*, if they share a common face having dimension equal to *q*.

---

**Graph**: *The Dynamics of Complex Urban Systems*, Springer, 2008
Two simplexes of complex $K$ are \textit{q-connected} if they are q-near, or there is a chain of pairwise q-near simplexes that links them together.

$$\Lambda = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1
\end{bmatrix}$$

$$\dim(K) = \max\left\{ \dim(\sigma^{(i)} = s_i) \mid i = 1,4 \right\} = 5$$

\textbf{Q-analysis} of complex $K$  $\Rightarrow$ $Q = (Q_{\dim(K)}^{2}, \ldots, Q_{1}, Q_{0})^\top$  (Q-analysis is based on studying the way simplexes are connected to each other by means of chains of q-connectivity  $\Rightarrow$  \textit{multidimensional structure}.)
Each element $Q_q$ of the structural vector $Q$ (global characteristic of complex $K$) is the number of connectivity components at the dimensional level $q$.

\[
\Lambda = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\]

\[Q = (1,3,4,3,1,1)^T\]

Structural complexity (connectivity) estimate $\Psi = \Psi(K)$ (J.L. Casti):

\[
\Psi(K) = \frac{2}{(N+1) \cdot (N+2)} \cdot \sum_{q=0}^{N} (q+1) \cdot Q_q, \quad N = \dim K
\]
Q-analysis Procedure. Complexity Estimate

Based on the following version of axioms:

[1] a system \((complex\ K)\) consisting of a single simplex has \(\Psi(K) = 1\),
[2] the complexity of subsystem \((subcomplex\ K^* \subseteq K)\) \(\leq \Psi(K)\),
[3] the combination of two complexes \((K_1 \text{ and } K_2)\) results in obtaining
new complex \(K\), for which \(\Psi(K) \leq \Psi(K_1) + \Psi(K_2)\) (following \(J.L.\ Casti\)),

Implicit assumption → system is connected at 0-level \((Q_0 = 1)\),

\((Casti\ J.L.,\ On\ System\ Complexity,\ 1985)\):
“… System complexity is a contingent property arising out of the
intersection \(I\) between a system \(S\) and an observer/decision-maker \(O\).
Thus, any perception and measure of complexity is necessarily a
function of \(S, O\) and \(I\).”

\[vector\ Q \rightarrow \text{measure (estimate) of complexity}\]
Structural complexity (connectivity) estimate $\Psi$
(within the scope of used mathematical representation) expression on the
strength of domain expert’s diverse considerations and prerequisites
George Kelly’s personal constructs - none of humans “has neutral access
to reality”, anticipation of ambient events psychologically channelizes
person’s processes; the impact of relativity and subjectivity factors on
both results interpretation and carried out formal calculations becomes
tangible,

Results summarized in vector $Q$ substantially masked declarative knowledge derived from facts and formal
representation of system’s structural connectivity; these
components are mentally apprehended, placed and processed
in certain area of mind (psychological space, or in short, P-space)
Questions to consider…

[1] will experts pay attention to apparently «missing» pieces of essential information while expressing their opinions about $\Psi(K)$ estimate?

[2] will they attempt to extract all available data to bring them into play at the stage of personal constructs formation to replicate these percepts as dimensions in P-space?

[3] will experts proceed along the path of combining findings into some representation form that is convenient for both perception and comparison (a kind of anchoring)? …

… Q-analysis results NOT only in obtaining q-connectivity vector
Results of Q-analysis. Typical features of K

Q-analysis results

- the number of connectivity components at the dimensional level q ($Q_q$)
- the number of simplexes having dimension q or greater ($S_q$)
- total number of non-empty simplexes of all dimensions in complex K ($s(K)$)

typical features of the level q

current dimensional level of complex’s analysis (q)
Turning these accessible portions of information (see prev. slide) into convenient forms qualified for storage and further processing that confine to conventional models used in research fields of psychology ➔ handy **structured base** to simplify comprehension as a *complex human mental ability* to represent, understand and interpret accumulated pieces of information ➔ (as a possible approach) **spatial representation** allowing to take also account of grouped entities + domain expert’s knowledge,

- **P-space** and its “geometry” ≠ concept of space in mathematics,
- The idea to represent objects (**stimuli**) as **points in space** and estimate **similarity of those stimuli through distance** between corresponding points are firmly rooted in psychological studies for many decades
Assumption: description of the concept of connectivity (structural complexity of K) is put into effect on the strength of term dictionary composed of variable q-level feature vectors

\[ A_q = \left( \frac{s_q}{s(K)}, \frac{Q_q}{s_q} \right) = \left( A_q^{(1)}, A_q^{(2)} \right) \]

- Fraction of complex K’s non-empty simplices analyzed at the level q
- Average number of simplices in one q-connectivity component

Basic premises and thought regulations prevailing in human perception:
“...when estimating numbers, most people start with a number (anchor) that comes easily to mind and adjust up/down from that initial state”
"Number sense" human mental ability to grasp the meaning of numbers and relationships between them, quickly perform approximation of quantities that arise from basic to more complex numerical calculations,

Limiting values:

\[
\begin{align*}
\text{if } s_q = s(K) & \Rightarrow \tilde{A}_q^{(1)} = 1 \\
\text{if } Q_q = 1 & \Rightarrow \tilde{A}_q^{(2)} = \frac{1}{s_q}
\end{align*}
\]

the most “severe” case that can be rarely observed in practice at particular q-level

Implication *IC-rules* (Idealized Case rules; refer to anchors mentioned above):

\[
\Psi_q \rightarrow \text{squiggle rule}
\]

\[
\tilde{A}_q = \left( \tilde{A}_q^{(1)}, \tilde{A}_q^{(2)} \right) = \left( 1, \frac{1}{s_q} \right) \Rightarrow \tilde{\Psi}_q = 1
\]
Q-analysis. Feature Vectors [3]

- Implication *AC-rules* (Actual Case rules; computed estimate of q-connectivity): 
  \[ \hat{A}_q = \left( \hat{A}_q^{(1)}, \hat{A}_q^{(2)} \right) = \left( \frac{s_q}{s(K)} \cdot \frac{Q_q}{s_q} \right) \Rightarrow \hat{\Psi}_q = <value> \]

- Theory of memory retrieval if patterns (*IC-rules*) are stored in the memory for a time needed to analyze results at particular q-level, then cognizable stimuli in the form of \( \Psi_q \) – cap rule association rule resonates with them through «invocation» of *IC- and AC-rules* left parts and assessment of their similarity.

Prepared by K. Degtiarev / August 2011
Perceived similarity is not evidently believed to be invariant different models are proposed to measure similarity between two patterns (exemplars),

Geometric models that assume inverse distance measures in a metric space as a basis of proximity (similarity) between exemplars remain the most influential and commonly used ones their results in low-dimensional space appear to be quite convincing and interpretable,

For vectors $\tilde{A}_q$ and $\check{A}_q$ we may notice certain resemblance propositions of prototype theory put forward by Eleanor Rosch (Harvard University)

Prototype mentally represented pattern of knowledge (model of concept)

- **Distance** between two objects \( \tilde{A}_q \) and \( \hat{A}_q \) can be converted to **proximity** (nearness) of these objects \( D(\tilde{A}_q, \hat{A}_q) \rightarrow P(\tilde{A}_q, \hat{A}_q) = f(D(\cdot)) \), where \( f \) is a function determining chosen similarity model on respective space,

- In respect to \( K \) and its subcomplexes, category «maximally attainable complexity (or, connectivity)» at a given q-level can be evinced by IC-rules (see slide 14),

- Rules can be considered as marked anchors in human comparison/categorization of objects based on their perceptual resemblance; each calculated in turn \( \tilde{A}_q \) is appraised for the purpose of its **proximity** to the prototypical exemplar (representation built on \( \tilde{A}_q \)),

prepared by K.Degtiarev / August 2011
Emphasis: psychological space (P-space) endued with metric preference to specific geometric model of cognition; it focuses on analysis of similarity (difference) of data objects complying with Roger Shepard’s Universal Law of Generalization that suggests to take a view of similarity as “a function of the distance between psychological representations”.

Calculated vector (stimulus) \( \hat{A}_q \) is represented on two emerged dimensions in P-space thus enabling to attribute \( \hat{A}_q^{(1)} \) and \( \hat{A}_q^{(2)} \) to corresp. values on psychological dimensions.

Distance measures \( D_{q}^{[r]} \) r-Minkowski metric

\[
P_q(\hat{A}_q, \hat{A}_q) = \exp\left[-a\left(D_{q}^{[r]}(\hat{A}_q, \hat{A}_q)\right)^n\right]
\]

\( a \) is sensitivity parameter (determined by experts), \( n=2 \).
Feature Vectors \([7]\). Perceived Similarity. Prototypes \([4]\)

- Feature vectors allow us to turn *basic information granules* (as shown in slide 11) into geometrical units (vect. \(A_q\)) – the latter relies on similarity notion while dealing with classification,

- Verbal valuation of the q-level by force of *proximity degree* requires utilization of extra *conceptual categories of connectivity* (e.g. «marginal», «close to medium», etc.) that don’t possess *sharp boundaries*,

- The value of \(\Psi_q\) can be obtained under the assumption of existence of latent dependency \(\Psi_q = P_q(\tilde{A}_q, \tilde{A}_q) = f(D^{[r]}_q)\) characterized by conditions:
  
  1. if \(D^{[r]}_q = 0\) \(\Rightarrow\) \(\Psi_q = \tilde{\Psi}_q = 1\)
  2. if \(D^{[r]}_q > 0\) \(\Rightarrow\) \(\Psi_q\) is decreasing \(\rightarrow 0\) with the distance’ growth
Proposed approach makes stress on simple transition forms having clear expression from the viewpoint of representation of human’s information perception and processing.

Steps of action chain «[1] data obtained from Q-analysis procedure – [2] grouping for the purpose of representation in low-dimensional conceptual space – [3] conferring metric upon space – [4] forming proximity (connectivity) estimates followed by their verbal interpretation» also have something in common with individual Gestalt principles of perception (similarity, proximity, etc.) in the sense that Q-analysis data are “organized into a stable and coherent form” named feature vectors.

The computational process

\[
\{\hat{A}_{q}^{(k)}, \bar{A}_{q}^{(k)}\}_{k=1,2} \rightarrow D_{q}^{[r=2]} \rightarrow \hat{\Psi}_{q}
\]

\(r = 2\) – dependency of dimensions
Geometry of situation based on our perception dimensions \( \rightarrow \) distance in vector conceptual space (close connection between psychological and geometrical spaces is based on the order of their primitives – through geometrical structures we are trying to grasp the grounds of information representation and processing at conceptual level)
Q-analysis. Distance and Proximity. Example [1]

- Structural vector $\mathbf{Q} = (1, 3, 2, 1)$

* (following Casti) $\Psi(K) = 1.8$

$s_3 = 1$, $s_2 = 3$, $s_1 = 5$, $s_0 = 5$

$a = 18$, $n = r = 2$

$D^{[r=2]}(\mathbf{D}_3^{[2]}, \ldots, \mathbf{D}_0^{[2]}) = (0.8, 0.78, 0.2, 0) \rightarrow (\hat{\Psi}_3, \ldots, \hat{\Psi}_0) = (0.32, 0.34, 0.93, 1)$
Turning acceptable range of calculated into unit interval affords good opportunity for interpretation of resultant numeric values by means of verbal terms (e.g. weak (low), medium, rather strong (high), strong (high), extremely strong (high)) association with fuzzy intervals,

What about aggregation of individual $\widehat{\Psi}_q$ values into single relative (subjective) estimate $\widehat{\Psi}_{\text{agg}}(K)$ of K’s complexity?

... take into account dimensions of complexes under consideration

weighted sum of individual values (using standardized weights / significance factors)

scale of verbal priorities (\textit{SVP} / T. Saaty)

- Option q-level significance factors  
  \( \varphi(q) = \bar{c} \cdot q^2 + 1, \ q = \text{dim}(K),...,0 \), where \( \bar{c} \) value is directly related to psychological level of human perception of q-levels significance; standardized weight of q-level  
  \( \bar{w}_q = \varphi(q) / \sum_{k=0}^{N} \varphi(k), \ N = \text{dim}(K) \); e.g. under \( \bar{c} = 0.05 \) significance factors for q-levels of complex’s K analysis are 1.45 (\( q = 3 \)), 1.2, 1.05 and 1 (\( q = 0 \)); weights of q-levels are 0.31 (\( q = 3 \)), 0.26, 0.22 and 0.21 (\( q = 0 \)),

- Aggregated measure  
  \( \hat{\Psi}_{\text{agg}}(K) = \sum_{k=0}^{N} \bar{w}_k \cdot \hat{\Psi}_q \Rightarrow \hat{\Psi}_{\text{agg}}(K) \approx 0.6 \)

- Conjugate complex \( K^* \)  
  \( Q^* = (2, 3, 1) \); \( \hat{\Psi}_2, \hat{\Psi}_1, \hat{\Psi}_0 = (0.29, 0.82, 1) \)

- Aggregated value  
  \( \hat{\Psi}_{\text{agg}}(K^*) \approx 0.68 \); [?] \( \hat{\Psi}_{\text{agg}}(K) \)
Scale of verbal priorities proposed by Thomas Saaty can be utilized to express judgments concerning perceived difference between dimensionalities of arbitrary complexes A and B,

- The scale uses five base marks (numbers 1, 3, 5, 7 and 9) that reflect the human ability to perform confident differentiations and four transitional, more «diffused» comparative marks (2, 4, 6 and 8);
- Base scale ticks represent differences as «virtually absent» (1), «insignificant» (3), «non-negligible» (5), «substantial» (7) and «absolute» (9) - such approach exploits rough classification of stimula by three key signs “rejection- indifference-acceptance”, each of which is endowed with specializing shades “low-average-high”

Utility 1: Graph Builder

Utility 2: Graph Builder

Utility 3: Fuzzy Calculator
Integration of information granules (features) obtained at the stage of Q-analysis into compact and interpretable form clears the way to form (calculate) structural complexity estimate making partial feasible use of perception theory and ideas derived from cognitive science,

Development of geometric models proves their adequacy when dealing with similarity that is a base for human cognitive abilities – such models are intuitive and have a good potential in being adopted in formal models of cognition (e.g. connectionist type of models),

The approach sets up a milestone in thorough elaboration of different approaches (fuzzy sets and systems, abstract mechanisms of human information processing described by cognitive models, comparison of complexes followed by making decisions concerning their observed differences (similarities))
Thank You for attention