Characterizing Classes of Structured Objects by Means of Information Inequalities

Matthias Dehmer
Center for Mathematics
University of Coimbra, Apartado 3008
3001-454 Coimbra, Portugal
dehmer@math.uc.pt

Stephan Borgert
Darmstadt University of Technology
Telecooperation Group
Hochschulstr. 10
D-64289 Darmstadt, Germany
borgert@tk.informatik.tu-darmstadt.de

Abstract

The problem to characterize and investigate structured objects by using information theory is currently of considerable interest. In this paper, we describe a method for characterizing structured objects representing graphs by means of information inequalities. For this, we deal with information inequalities which describe relations between information measures for graphs. Additionally, we sketch an approach for comparing such information measures qualitatively.

1 Introduction

In cybernetics, sociology, and computer science, the theory of structured objects provides powerful tools to represent structures and processes by means of graphs (Cook and Holder 2007; Hage and Harary 1995; Jablonski and Lupanow 1980; Lewin 1936; Sobik 1982; Sobik 1986; Sommerfeld and Sobik 1994; Sommerfeld 1994; Wasserman and Faust 1994). By structured objects, we refer to objects which can be represented by relational algebras. If one restricts to the analysis of two-digit relations, such relational objects can be easily described by graphs (Harary 1969; Sobik 1986).

In this paper, we want to address the abstract question to characterize classes of structured objects by using information inequalities. In our case, an information inequality describes a relation between information measures for graphs. In terms of dealing with structured objects, information theory (Aczél and Daróczy 1975; Shannon and Weaver 1997) offers a lot of quantitative methods to investigate information processing and information transmission in graphs (Dehmer 2008a; Solé and Valverde 2004). In particular, if we want to characterize structured objects as well as classes of those by applying information measures, we have to answer the question to which aspect of information we are referring to, e.g., structural information or semantic information. Throughout this paper, if we speak about abstract information measures for graphs, we always use the concept of determining structural information (Dehmer 2008a; Konstantinova 2006; Mowshowitz 1968d; Mowshowitz 1968a; Mowshowitz 1968b; Mowshowitz 1968c; Rashevsky 1955; Solé and Valverde 2004), e.g., by inferring structural characteristics of graphs (Dorogovtsev and Mendes 2003; Skorobogatov and Dobrynin 1988). Then, this task is equivalent to measure the so-called structural information content (Bonchev 1983; Mowshowitz 1968d) of a graph representing the entropy of the underlying graph topology (Bonchev 2003; Dehmer 2008a; Mowshowitz 1968d). This paper aims to outline an approach for characterizing classes of structured objects by means of information inequalities. We will finally say that a class of structured objects is characterized by a certain information measure iff for each pair of graphs a certain (information) inequality holds. One possibility to do so, is to apply a given information measure. As a result, we get a system of information inequalities. Another possibility is to infer such inequalities theoretically. We give an example for doing so in Section (3.1).

This paper is organized as follows: In Section (2), we briefly state some mathematical preliminaries. The approach to characterize a class of structured objects by means of information inequalities is expressed in Section (3). In Section (4), we give an example for performing a qualitative analysis of information measures to characterize structured objects. For this, we outline a method to trace back this problem to the problem of comparing certain graphs.

2 Mathematical Preliminaries

In the following, we first express some graph-theoretical and information-theoretic definitions (Aczél and Daróczy 1975; Harary 1969; Dehmer, Emmert-Streib, and Gesell 2008; Harary 1969; Shannon and Weaver 1997; Sobik 1986).

Definition 2.1 *Let $A_1, A_2, \ldots, A_n$ be arbitrary sets. Then, a n-digit relation R is a subset of the Cartesian product of the corresponding sets $A_1, A_2, \ldots, A_n$,*

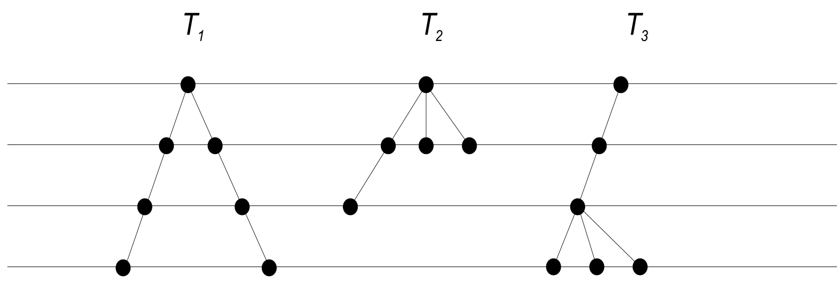

Figure 1: A graph class $C := \{T_1, T_2, T_3\}$. Each H_i represents a rooted tree.

i.e.,

$$R \subseteq A_1 \times A_2 \times \cdots \times A_n, \tag{1}$$

and

$$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, a_2, \ldots, a_n) | \, a \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\}. \tag{2}$$

Definition 2.2 *Let V be the set of elementary objects. Let E be a n-digit relation between the elementary objects of V. We call $O := (V, E)$ a general, finite structured object iff $|V| < \infty, V \neq \emptyset$ and $E \subseteq V \times V \times \cdots \times V$.*

Definition 2.3 *Let V be the set of elementary objects which we call vertices. Let E be a 2-digit relation between the vertices of V. If O is defined as $O := (V, E)$, $|V| < \infty, V \neq \emptyset$ and $E \subseteq V \times V$, we call O a finite directed graph. In contrast, if it holds $O := (V, E)$, $|V| < \infty$, $E \subseteq \binom{V}{2}$, we call O a finite undirected graph.*

Definition 2.4 *Let $(\Delta_n)^k, \mathbb{N} \ni n \geq 2$ be the set*

$$(\Delta_n)^k := \left\{ P = (p_1, p_2, \ldots, p_n) \middle| \, p_i \in [0,1]^k, \; i = 1, 2, \ldots, n, \sum_{i=1}^{n} p_i = \underline{1} \right\}. \tag{3}$$

The addition of the vectors is defined componentwise and $\underline{1}$ denotes the k-dimensional vector $(1, 1, \ldots, 1)^T$. The sequence of functions

$$I_n : (\Delta_n)^k \longrightarrow \mathbb{R} \tag{4}$$

is called a k-dimensional information measure. If $k = 1$,

$$I_n : (\Delta_n)^1 \longrightarrow \mathbb{R} \tag{5}$$

are called entropies.

3 Characterization of Structured Objects: Information Inequalities

In this section, we express an approach to characterize classes of structured objects by using information inequalities. For this, we assume C to be a certain class of directed or undirected graphs. Further, we now set $k = 1$ and define the domain of Equation (5) as C because we interpret the quantities $p_1, p_2, \ldots, p_n$ as probability values inferred from the structured objects under consideration. That means, for a given $G \in C$, we first derive probability values $p_1, p_2, \ldots, p_n$ associated to G. Finally, this implies that $I_n(G) \in \mathbb{R}$ quantifies structural information of G. In the following, we always write $I(G)$ instead of $I_n(G)$ because we apply such information measures to structured objects with different orders. Starting from these preliminaries, we now state the main definition.

Definition 3.1 *Let C be a class of directed or undirected graphs. We say that C is characterized by an information measure $I : C \longrightarrow \mathbb{R}$ iff for each possible pair (G, H) with $G \neq H$ the inequality system*

$$I(G) \underset{(<)}{\geq} I(H), \tag{6}$$

holds.

Example 3.1 *We consider the graph class C depicted in Figure (1). Now, to quantify structural information of these graphs, we exemplarily apply the following information measure (Emmert-Streib and Dehmer 2007): Because a rooted tree T can be naturally divided into horizontal levels representing vertex partitions, we define the probability value for each level i by*

$$p_i^V := \frac{|V_i|}{|V| - 1}, \tag{7}$$

$1 \leq i \leq h$. *Starting from $T \in C$, $|V|$ denotes the total number of vertices, $|V_i|$ denotes the number of vertices on the i-th level and h denotes the height of T, respectively. Then, the vertex entropy of a rooted tree T is defined by using the* Shannon-*entropy (Shannon and Weaver 1997)*

$$I^V(T) := -\sum_{i=1}^{h} p_i^V \log(p_i^V). \tag{8}$$

We yield,

$$I^V(T_1) = \log(3) = 1.5850 \ bits, \tag{9}$$

$$I^V(T_2) = -\frac{3}{4}\log\left(\frac{3}{4}\right) - \frac{1}{4}\log\left(\frac{1}{4}\right) = 0.8113 \ bits, \tag{10}$$

$$I^V(T_3) = -\frac{2}{5}\log\left(\frac{1}{5}\right) - \frac{3}{5}\log\left(\frac{3}{5}\right) = 1.3710 \ bits, \tag{11}$$

where we took the logarithm to the base 2*. Finally, we get the inequality system*

$$I^V(T_1) > I^V(T_2), \tag{12}$$

$$I^V(T_1) > I^V(T_3), \tag{13}$$

$$I^V(T_2) < I^V(T_3). \tag{14}$$

3.1 Example for Inferring Information Inequalities

In Section (3), we claimed that C is characterized by an information measure I iff a certain inequality system holds. In Example (3.1), we applied a special information measure to characterize the given class C. As a direct consequence, we got the inequality system described by the Inequalities (12)-(14). In contrast, we now show that it is also possible to infer information inequalities for graphs theoretically. For giving an example, we state the following assertions which have been recently proven in (Dehmer 2008b).

Theorem 3.2 *Let C be a class of directed or undirected graphs. Further, let $G = (V_G, E_G) \in C$ and $H = (V_H, E_H) \in C$ be non-isomorphic graphs. If $|G| > |H|$ and the following system of vertex probabilities*

$$\begin{aligned}
p(\bar{v}_{j_1}, G) &> p(v_{i_1}, H),\\
p(\bar{v}_{j_2}, G) &> p(v_{i_2}, H),\\
&\vdots\\
p(\bar{v}_{j_{|V_H|}}, G) &> p(v_{i_{|V_H|}}, H),\\
p(\bar{v}_{j_{|V_H|+1}}, G) &> p(v_{\mu_1}, H),\\
&\vdots\\
p(\bar{v}_{j_{|V_G|}}, G) &> p(v_{\mu_k}, H),
\end{aligned} \tag{15}$$

holds, then we obtain the information inequality

$$I(G) < I(H) + d_1, \quad d_1 > 0, \tag{16}$$

where

$$d_1 := -\sum_{i}^{k} p(v_{\mu_i}, H) \log \left(p(v_{\mu_i}, H)\right). \tag{17}$$

If $|G| < |H|$, the relation symbol in Inequality System (15) must be changed to '$<$'. Then, we infer

$$I(G) + d_2 > I(H)\,, \quad d_2 > 0, \tag{18}$$

where

$$d_2 := -\sum_{i}^{k} p(\bar{v}_{\mu_i}, G) \log \left(p(\bar{v}_{\mu_i}, G)\right). \tag{19}$$

Starting from Theorem (3.2), we also get

Theorem 3.3 *It holds*

$$I(G) > I(H) \quad or \quad I(G) < I(H) \quad or \quad I(G) = I(H). \tag{20}$$

We notice that in Theorem (3.2) and Theorem (3.3), the information measures $I(G)$ and $I(H)$ represent the corresponding Shannon-entropies. We interpret these assertions in a way that one is able to predict information inequalities for structured objects by assuming certain constraints. As constraint, we here chose an inequality systems that describes relations between the vertex probabilities (Dehmer 2008b).

4 Qualitative Analysis of Information Measures

This section aims to give an example to define a method for performing a qualitative analysis of information measures to characterize graphs. A similar approach has been defined in (Sobik 1986) for the qualitative comparison of graph similarity measures. In general, while doing a quantitative analysis means to examine numerical differences between information measures, a qualitative analysis examines relationships between the used information measures starting from a class of structured objects. First, that means we identify the pairs of structured objects (G, H) for which, e.g., $I(G) \geq I(H)$ holds. This procedure results in a graph where the set of vertices equals C whereas two vertices will be connected by an edge iff for the corresponding information measures the relation '$\geq$' holds. Second, if we apply this procedure to the same class C by using a different information measure $\bar{I}$, one can finally compare I and $\bar{I}$ qualitatively by comparing the resulting graphs structurally (Bunke and Neuhaus 2007; Sobik 1986).

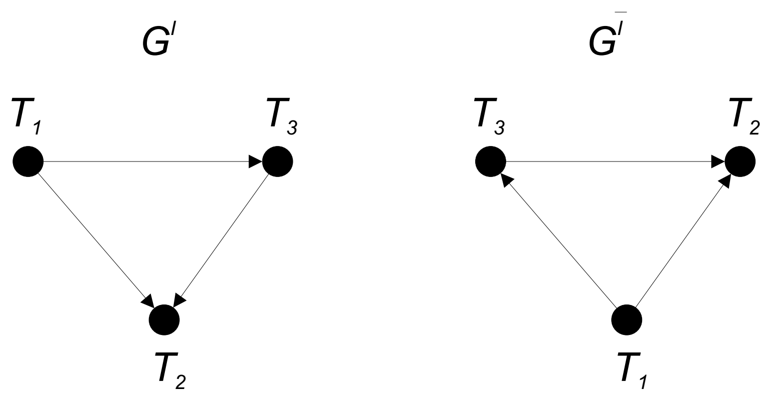

Figure 2: The resulting graphs G^I and $G^{\bar{I}}$ for comparing I and $\bar{I}$ qualitatively.

Definition 4.1 *Let C be a class of directed or undirected graphs and let $I : C \longrightarrow \mathbb{R}$ be an information measure. We define the directed graph*

$$G^I = (V^I, E^I), \tag{21}$$

by setting $V^I = C$ and

$$E^I := \{(G, H) | \, G \neq H, \, I(G) \geq I(H)\}. \tag{22}$$

Example 4.1 *Once again, we look at the trees shown in Figure (1). Similarly to Information Measure (8), we define the edge entropy of a tree as follows (Emmert-Streib and Dehmer 2007). Let T be a rooted tree with height h. $|E|$ denotes the total number of edges and $|E_i|$ denotes the number of edges on the i-th level, respectively. Now, we set*

$$p_i^E := \frac{|E_i|}{2|E| - \delta(r)}. \tag{23}$$

Then, the edge entropy of T is defined by

$$I^E(H) := -\sum_{i=1}^{h} p_i^E \log(p_i^E). \tag{24}$$

By using this information measure, we obtain

$$I^E(T_1) = -\frac{4}{5} \log\left(\frac{2}{5}\right) - \frac{1}{5} \log\left(\frac{1}{5}\right) = 1.5219 \text{ bits}, \tag{25}$$

$$I^E(T_2) = -\frac{4}{5} \log\left(\frac{4}{5}\right) - \frac{1}{5} \log\left(\frac{1}{5}\right) = 0.7219 \text{ bits}, \tag{26}$$

$$I^E(T_3) = -\frac{2}{9} \log\left(\frac{2}{9}\right) - \frac{4}{9} \log\left(\frac{4}{9}\right) - \frac{3}{9} \log\left(\frac{3}{9}\right) = 1.5305 \text{ bits}. \tag{27}$$

Finally, we obtain the same inequality system, i.e.,

$$I^E(T_1) > I^E(T_2), \tag{28}$$

$$I^E(T_1) > I^E(T_3), \tag{29}$$

$$I^E(T_2) < I^E(T_3). \tag{30}$$

If we now apply Definition (4.1) to T_1, T_2 *and* T_3*, we get the graphs* G^I *and* $G^{\bar{I}}$ *which are depicted in Figure (2) for comparing* I^V *and* I^E *qualitatively. In this case we see that the obtained graphs are isomorphic.*

5 Summary and Conclusion

This paper dealt with suggesting a method to characterize classes of structured objects by means of information inequalities. By using an information measure that quantifies structural information of a graph, we called a graph class characterized by an information measure iff for each possible pair of graphs a certain information inequality holds. In the present paper, we outlined two possibilities to do so. First, we started with a special information measure and calculated the corresponding inequality system. The second possibility we sketched to characterize structured objects by means of information inequalities was to infer such relations theoretically. For this, we expressed an example for a method to generate information inequalities starting from certain constraints.

In the future, we want to investigate different kinds of information measures to characterize large sets of graphs.

6 Acknowledgments

We are grateful to Danail Bonchev and Abbe Mowshowitz for fruitful discussions.

References

Aczél, J. and Z. Daróczy (1975). *On Measures of Information and Their Characterizations*. Academic Press.

Bonchev, D. (1983). *Information Theoretic Indices for Characterization of Chemical Structures*. Research Studies Press, Chichester.

Bonchev, D. (2003). *Complexity in Chemistry. Introduction and Fundamentals*. Taylor and Francis.

Bunke, H. and M. Neuhaus (2007). Graph matching. exact and error-tolerant methods and the automatic learning of edit costs. In D. Cook and L. B. Holder (Eds.), *Mining Graph Data*, pp. 17–32. Wiley-Interscience.

Cook, D. and L. B. Holder (2007). *Mining Graph Data*. Wiley-Interscience.

Dehmer, M. (2008a). Information processing in complex networks: Graph entropy and information functionals. *Applied Mathematics and Computation 201*, 82–94.

Dehmer, M. (2008b). Relations between the topological complexities of complex networks. In *Proceedings of the 2008 International Conference on Machine Learning: Models, Technologies & Applications (MLMTA'08), Las Vegas/USA, 2008*. in press.

Dehmer, M., F. Emmert-Streib, and T. Gesell (2008). A comparative analysis of multidimensional features of objects resembling sets of graphs. *Applied Mathematics and Computation 196*, 221–235.

Dorogovtsev, S. N. and J. F. F. Mendes (2003). *Evolution of Networks. From Biological Networks to the Internet and WWW*. Oxford University Press.

Emmert-Streib, F. and M. Dehmer (2007). Information theoretic measures of UHG graphs with low computational complexity. *Applied Mathematics and Computation 190*, 1783–1794.

Hage, P. and F. Harary (1995). Eccentricity and centrality in networks. *Social Networks 17*, 57–63.

Harary, F. (1969). *Graph Theory*. Addison Wesley Publishing Company.

Jablonski, S. W. and O. B. Lupanow (1980). *Diskrete Mathematik und mathematische Fragen der Kybernetik*, Volume 71 of *Lehrbücher und Monographien aus dem Gebiete der exakten Wissenschaften: Mathematische Reihe*. Birkhäuser Publishing.

Konstantinova, E. V. (2006). On some applications of information indices in chemical graph theory. In R. Ahlswede, L. Bäumer, N. Cai, H. Aydinian, V. Blinovsky, C. Deppe, and H. Mashurian (Eds.), *General Theory of Information Transfer and Combinatorics (Lecture Notes in Computer Science*, Lecture Notes of Computer Science, pp. 831–852. Springer.

Lewin, K. (1936). *Principles of Topological Psychology*. McGraw-Hill, New York.

Mowshowitz, A. (1968a). Entropy and the complexity of graphs II: The information content of digraphs and infinite graphs. *Bull. Math. Biophys. 30*, 225–240.

Mowshowitz, A. (1968b). Entropy and the complexity of graphs III: Graphs with prescribed information content. *Bull. Math. Biophys. 30*, 387–414.

Mowshowitz, A. (1968c). Entropy and the complexity of graphs IV: Entropy measures and graphical structure. *Bull. Math. Biophys. 30*, 533–546.

Mowshowitz, A. (1968d). Entropy and the complexity of the graphs I: An index of the relative complexity of a graph. *Bull. Math. Biophys. 30*, 175–204.

Rashevsky, N. (1955). Life, information theory, and topology. *Bull. Math. Biophys. 17*, 229–235.

Shannon, C. E. and W. Weaver (1997). *The Mathematical Theory of Communication*. University of Illinois Press.

Skorobogatov, V. A. and A. A. Dobrynin (1988). Metrical analysis of graphs. *MATCH 23*, 105–155.

Sobik, F. (1982). Graphmetriken und Klassifikation strukturierter Objekte. *ZKI-Informationen, Akad. Wiss. DDR 2*(82), 63–122.

Sobik, F. (1986). Modellierung von Vergleichsprozessen auf der Grundlage von Ähnlichkeitsmaßen für Graphen. *ZKI-Informationen, Akad. Wiss. DDR 4*, 104–144.

Solé, R. V. and S. Valverde (2004). Information theory of complex networks: On evolution and architectural constraints. In *Lecture Notes in Physics*, Volume 650, pp. 189–207.

Sommerfeld, E. (1994). *Kognitive Strukturen. Mathematisch-psychologische Elementaranalysen der Wissensstrukturierung und Informationsverarbeitung*. Waxmann Publishing.

Sommerfeld, E. and F. Sobik (1994). Operations on cognitive structures - their modeling on the basis of graph theory. In D. Albert (Ed.), *Knowledge Structures*, pp. 146–190. Springer.

Wasserman, S. and K. Faust (1994). *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge University Press.