Combinatorics, Probability and Computing (2001) 10, 501–504. © 2001 Cambridge University Press
DOI: 10.1017/S0963548301004783 Printed in the United Kingdom

A Note on Counting Very Different Sequences

EMANUELA FACHINI and JÁNOS KÖRNER

Department of Computer Science, University of Rome I 'La Sapienza',
Via Salaria 113, 00198 Roma, Italy
(e-mail: fachini, korner@dsi.uniroma1.it)

Received 7 August 2000; revised 8 January 2001

We strengthen the Cohen–Körner–Simonyi upper bound on counting 'very different sequences'.

1. Introduction

Let V be an arbitrary finite set and let $\mathcal{G}$ be a family of graphs with common vertex set V. Given a natural number n, two sequences $\mathbf{x} = x_1, \ldots, x_n$ and $\mathbf{y} = y_1, \ldots, y_n$ are called very different or $\mathcal{G}$-different if, for every $G \in \mathcal{G}$, there is a coordinate $i \in [n]$ for which we have $\{x_i, y_i\} \in E(G)$. This concept was introduced by Cohen, Körner and Simonyi [1] to tackle various problems in extremal set theory via information theory, and led eventually to the solution of Rényi's problem on counting qualitatively independent partitions [3]. Let $M(n, \mathcal{G})$ denote the maximum number of pairwise $\mathcal{G}$-different sequences from V^n. This number can be upper-bounded in terms of graph entropy, and the resulting bound is tight when, for example, all the graphs in the family are perfect.

Graph entropy $H(G, P)$ is an information-theoretic functional of a graph G with a probability distribution P on its vertex set, introduced in Körner [4]. It is defined by

$$H(G, P) = \min_{X \in Y \in S(G),\, P_X = P} I(X \wedge Y),$$

where $S(G)$ denotes the family of the stable sets of vertices in G. (A subset of the vertex set is called stable if it does not contain any edge. For the basics of information theory the reader is referred to the book [2]. We recall that the mutual information $I(X \wedge Y)$ of the random variables X and Y equals $H(X) + H(Y) - H(X, Y)$, where $H(X, Y)$ is the entropy of the random variable (X, Y). Notice that the entropy of a random variable is the entropy of its distribution.) The interested reader can find a more complete introduction to graph entropy in the survey of Simonyi [7].

https://doi.org/10.1017/S0963548301004783 Published online by Cambridge University Press

It follows from the results of [1] that

$$\limsup_{n\to\infty} \frac{1}{n} \log M(n, \mathscr{G}) \leqslant \max_{P} \min_{G\in\mathscr{G}} H(G, P),$$

where P runs over all the probability distributions on V. (Note that all the logs are binary.)

A particularly interesting case of the above is when all the graphs in the family have a single edge. Then it is useful to consider the union graph $U(\mathscr{G})$. Given two arbitrary graphs F and G, their union $F \cup G$ is defined by setting

$$V(F \cup G) = V(F) \cup V(G) \qquad \text{and} \qquad E(F \cup G) = E(F) \cup E(G).$$

As an immediate consequence of the results in [3], we have the following result.

Theorem GK.

$$\limsup_{n\to\infty} \frac{1}{n} \log M(n, \mathscr{G}) = \max_{P} \min_{\{a,b\}\in E(U(\mathscr{G}))} [P(a) + P(b)] h\left(\frac{P(a)}{P(a) + P(b)}\right),$$

where $h(t) = -t \log t + (1 - t) \log(1 - t)$ *is the binary entropy function.*

The aim of this note is to prove a non-asymptotic version of the upper bound. This answers a question of A. Monti. We start with a few prerequisites.

A crucial property of $H(G, P)$ needed in our proof is its sub-additivity with respect to graph union [5]. If F and G have the same vertex set, then, for every P, we have

$$H(F \cup G, P) \leqslant H(F, P) + H(G, P). \tag{1.1}$$

2. New inequalities

A hitherto unexploited property of this functional needed here is the following.

Lemma 2.1. *For a fixed graph* G*, its entropy* $H(G, P)$ *as a function of the distribution* P *on its vertices is cap-convex, that is, for any* $\alpha \in [0, 1]$ *and distributions* P_1, P_2 *on* $V(G)$ *we have*

$$H(G, \alpha P_1 + (1 - \alpha)P_2) \geqslant \alpha H(G, P_1) + (1 - \alpha) H(G, P_2).$$

Proof. Let us consider the graphs G_i, $i \in \{1, 2\}$ with disjoint vertex sets $V_i = V(G_i)$, and both isomorphic to G. Consider further the distributions P_i on V_i together with the distribution P defined on $V_1 \cup V_2$ defined by setting $P(x) = \alpha P_1$ if $x \in V_1$ and, similarly, $P(x) = (1 - \alpha)P_2$ if $x \in V_2$. It is immediate from the definition of graph entropy (or else see [5]) that

$$\alpha H(G, P_1) + (1 - \alpha) H(G, P_2) = H(G_1 \cup G_2, P).$$

Our lemma follows at once on noticing that $H(G, \alpha P_1 + (1-\alpha)P_2)$ is obtained by restricting the minimization in the definition of the entropy of G to stable sets in which the vertices of the G_i, paired by the isomorphism of these two graphs, are not separated. (In other words, the only stable sets with positive probabilities are those containing either both

https://doi.org/10.1017/S0963548301004783 Published online by Cambridge University Press

or neither of any pair of vertices G_1, G_2 corresponding via the isomorphism of the two graphs.) □

In [1] the bounds on very different sequences were obtained by partitioning any set of sequences into classes in which every single sequence had the same 'type', in the sense of being obtainable from any other sequence in its class by a suitable permutation of its coordinates. However, this step is not really needed. The idea of using, alternatively, the 'average type' in similar arguments goes back to Massey [6]; *cf.* also the book by Csiszár and Körner [2].

Theorem 2.2. *For every natural number n and every family of graphs $\mathscr{G}$, we have*

$$M(n, \mathscr{G}) \leqslant \exp\left[n \cdot \max_P \min_{G \in \mathscr{G}} H(G, P)\right],$$

where the P runs over all probability distributions on V.

Proof. Let $C \subseteq V^n$ be a family of $\mathscr{G}$-different sequences attaining $M(n, \mathscr{G})$, and let $L = L_n$ be a complete graph with vertex set C. Further, let P_n be the uniform distribution on $V(L_n)$. Given an arbitrary $G \in \mathscr{G}$, let G_i be the graph with vertex set $V(L_n)$ in which two vertices, $\mathbf{x} \in V(L_n)$ and $\mathbf{y} \in V(L_n)$ are adjacent if, for the coordinate $i \in [n]$, we have $\{x_i, y_i\} \in E(G)$. By our hypothesis,

$$L_n = \bigcup_{i=1}^{n} G_i \quad \text{for} \quad \text{every} \quad G \in \mathscr{G}. \tag{2.1}$$

But then, by the sub-additivity of graph entropy,

$$\log M(n, \mathscr{G}) = H(P_n) = H(L_n, P_n) \leqslant \sum_{i=1}^{n} H(G_i, P_n) = \sum_{i=1}^{n} H(G, P^i), \tag{2.2}$$

where the distribution P^i is defined by setting

$$P^i(a) = \sum_{\mathbf{x}; x_i = a} P_n(\mathbf{x}).$$

Thus, applying Lemma 2.1 to the distribution

$$Q = \frac{1}{n} \sum_{i=1}^{n} P^i,$$

the inequality (2.2) yields

$$\log M(n, \mathscr{G}) \leqslant \sum_{i=1}^{n} H(G, P^i) \leqslant nH(G, Q), \tag{2.3}$$

whence

$$\frac{1}{n} \log M(n, \mathscr{G}) \leqslant \min_{G \in \mathscr{G}} H(G, Q) \leqslant \max_P \min_{G \in \mathscr{G}} H(G, P).$$

This concludes the proof of the theorem. □

https://doi.org/10.1017/S0963548301004783 Published online by Cambridge University Press

Corollary 2.3. *If every graph in $\mathcal{G}$ has a single edge, then, for every natural number n, we have*

$$M(n,\mathcal{G}) \leqslant \exp\left[n \cdot \max_{P} \min_{\{a,b\}\in E(U(\mathcal{G}))} [P(a)+P(b)]h\left(\frac{P(a)}{P(a)+P(b)}\right)\right].$$

Proof. It is sufficient to observe that, if G is a graph with a single edge $\{a,b\} = E(U(\mathcal{G}))$, then, for any distribution on $V(G) = V$, we have

$$H(G,P) = [P(a)+P(b)]h\left(\frac{P(a)}{P(a)+P(b)}\right). \qquad \square$$

Acknowledgement

Thanks are due to Angelo Monti for suggesting that we reconsider these problems, and for useful conversations.

References

[1] Cohen, G., Körner, J. and Simonyi, G. (1990) Zero-error capacities and very different sequences. In *Sequences: Combinatorics, Security and Transmission*, *Advanced International Workshop on Sequences, Positano, Italy, June 1988* (R. M. Capocelli, ed.), Springer, New York, pp. 144–155.

[2] Csiszár, I. and Körner, J. (1981, 1982) *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York (1982) and Akadémiai Kiadó, Budapest (1981).

[3] Gargano, L., Körner, J. and Vaccaro, U. (1993) Sperner capacities. *Graphs and Combinatorics* **9** 31–46.

[4] Körner, J. (1973) Coding of an information source having ambiguous alphabet and the entropy of graphs. In *Trans. 6th Prague Conf. Inform. Theory, etc., 1971*, Academia, Prague, pp. 411–425.

[5] Körner, J. (1986) Fredman–Komlós bounds and information theory. *SIAM J. Alg. Disc. Meth.* **4** 560–570.

[6] Massey, J. L. (1974) On the fractional weight on distinct binary n-tuples. *IEEE Trans. Inform. Theory* **20** 131.

[7] Simonyi, G. (1995) Graph entropy: A survey. In *Combinatorial Optimization* (W. Cook, L. Lovász and P. D. Seymour, eds), Vol. 20 of *DIMACS Series on Discrete Math. and Computer Science*, pp. 399–441.

https://doi.org/10.1017/S0963548301004783 Published online by Cambridge University Press