Poset Entropy Versus Number of Linear Extensions: The Width-2 Case

Samuel Fiorini · Selim Rexhep

Abstract  Kahn and Kim (J. Comput. Sci. 51, 3, 390–399, 1995) have shown that for a finite poset $P$, the entropy of the incomparability graph of $P$ (normalized by multiplying by the order of $P$) and the base-2 logarithm of the number of linear extensions of $P$ are within constant factors from each other. The tight constant for the upper bound was recently shown to be 2 by Cardinal et al. (Combinatorica 33, 655–697, 2013). Here, we refine this last result in case $P$ has width 2: we show that the constant can be replaced by $2 - \varepsilon$ if one also takes into account the number of connected components of size 2 in the incomparability graph of $P$. Our result leads to a better upper bound for the number of comparisons in algorithms for the problem of sorting under partial information.

Keywords  Poset · Entropy · Linear extension · Sorting under partial information

1 Introduction

The entropy of a graph is an information theoretic concept introduced by Körner in 1973 [8]. Since then, links with many interesting combinatorial objects have been found, see the survey paper of Simonyi [10] for more information.

In this paper, we consider the case in which the graph is the incomparability graph $G(P)$ of a (finite) poset $P$. We denote by $H(P) := H(G(P))$ the entropy of this graph. Kahn and Kim [7] have proved that $|P| \cdot H(P)$ is within a constant of $\log e(P)$, the base-2 logarithm of the number of linear extensions of $P$. (Throughout this paper, $\log$ denotes the base-2 logarithm).

This work is supported by the ARC Convention AUWB-2012-12/17-ULB2 (COPHYMA project).

S. Fiorini · S. Rexhep
Université Libre de Bruxelles, Bruxelles, Belgium
E-mail: srexhep@ulb.as.be

S. Fiorini
E-mail: sfiorini@ulb.ac.be
Theorem 1 (Kahn and Kim [7]) For every poset $P$:

$$\log e(P) \leq |P| \cdot H(\overline{P}) \leq c_0 \log e(P)$$

for $c_0 = (1 + 7 \log e) \simeq 11.1$.

Cardinal, Fiorini, Joret, Jungers and Munro [2] improved the constant in the upper bound to 2. This is tight since if $P$ is a two-elements antichain we have $|P| \cdot H(\overline{P}) = 2$ and $\log e(P) = 1$.

Theorem 2 (Cardinal et al. [2]) For every poset $P$:

$$|P| \cdot H(\overline{P}) \leq 2 \log e(P).$$

Our starting point is the observation that the upper bound is tight if every element of $P$ is incomparable to at most one other element, that is, $P$ is the ordinal sum of one-element and two-elements antichains: $P = A_1 \oplus A_2 \oplus \cdots \oplus A_k$ where each $|A_i| \leq 2$. Thus it seems likely that for some small enough constant $\varepsilon > 0$, one can prove that the posets with $|P| \cdot H(\overline{P}) \geq (2 - \varepsilon) \log e(P)$ possess a very constrained structure. Our main result is to establish such a phenomenon for width-2 posets and thus refine Theorem 2 in this case. We recall that the width of poset $P$ is the size of a largest antichain of $P$.

Theorem 3 Let $P$ be a width-2 poset and let $\kappa_2(P)$ denote the number of size-2 connected components of $G(P)$. Then

$$|P| \cdot H(\overline{P}) \leq (2 - \varepsilon) \log e(P) + \varepsilon \kappa_2(P)$$

for $\varepsilon = 2 - \frac{3 \log 3 - 2}{\log 3} \simeq 0.26$.

Note that Inequality (1) can be written

$$|P| \cdot H(\overline{P}) \leq \left(2 - \varepsilon \left(1 - \frac{\kappa_2(P)}{\log e(P)}\right)\right) \log e(P)$$

where $1 - \frac{\kappa_2(P)}{\log e(P)}$ is nonnegative since $e(P) \geq 2^{\kappa_2(P)}$ with equality if and only if the components of $G(P)$ are all of size either 1 or 2. From this we deduce:

Corollary 4 Let $P$ be a width-2 poset, then $|P| \cdot H(\overline{P}) = 2 \log e(P)$ if and only if the maximum degree of $G(P)$ is 1.

We remark also that upper bounds such as those in Theorems 1 and 2 translate to upper bounds on the worst case number of comparisons performed by algorithms for a sorting problem known as sorting under partial information, see e.g. [2],[7] for more details. In the context of this problem, Theorem 3 yields an improvement in the width-2 case (merging under partial information) because after comparing each of the $\kappa_2(P)$ pairs of elements that form connected components of $G(P)$, the constant in front of $\log e(P)$ decreases from 2 to $2 - \varepsilon \simeq 1.74$. Furthermore, we point out that the algorithm given by Cardinal et al. [2] reduces the general problem to the width-2 case, hence Theorem 3 also gives an improvement in the general case.

We begin in Section 2 with a brief account of the definitions and main properties of graph entropy. In Section 3, we specialize this to (in)comparability graphs of posets. In order to help the reader understanding the proof, its general structure is explained in Section 4.