A NOTE ON THE ASYMPTOTIC BEHAVIOUR OF BOTTLENECK PROBLEMS

HANSJÖRG ALBRECHER*

Abstract. We generalize and sharpen results of Burkard and Fincke concerning the asymptotic behaviour of a certain class of combinatorial optimization problems with bottleneck objective function. In this way several open questions are answered.

1. Introduction

Let us consider a family (P_n), $n \in \mathbb{N}$ of combinatorial optimization problems defined on finite ground sets E_n. For each $n \in \mathbb{N}$, the feasible solutions of problem P_n are defined by a non-empty class $\mathcal{S}_n$ of subsets of E_n. A feasible solution S is therefore a subset of E_n. Let $|\mathcal{S}_n|$ denote the cardinality of $\mathcal{S}_n$ and $|S|$ the number of elements $e \in E_n$ belonging to S. Let furthermore $f_n : E_n \to \mathbb{R}^+$ be a weight function. The bottleneck problem is now to find

$$F_{n,\min} = \min_{S \in \mathcal{S}_n} \max_{e \in S} f_n(e). \tag{1}$$

We will consider this class of generalized combinatorial optimization problems in a probabilistic framework, i.e. the weights $f_n(e)$ are assumed to be random variables on some arbitrarily large finite interval $[0, M]$ (problems of the form $F_n^* = \max_{S \in \mathcal{S}_n} \min_{e \in S} f_n(e)$ can be treated similarly and will not be mentioned in the sequel). Throughout the paper we will assume that $|S| = s_n$ for each $S \in \mathcal{S}_n$ (i.e. every feasible solution of P_n has the same cardinality s_n) and $\lim_{n\to\infty} s_n = \infty$, $\lim_{n\to\infty} |\mathcal{S}_n| = \infty$. Let (T) denote the class of optimization problems for which the relative difference between the worst and the best objective function value tends to zero with probability tending to one as the problem size n approaches infinity. In such a case every feasible solution is asymptotically optimal and the problem of finding the optimal solution becomes in some sense trivial for large instances. In particular, for optimization problems of class (T), simple heuristics will incline to give good solutions for high-dimensional problems (note that some of the elements of (T) are known to be NP-hard!).

There are many combinatorial problems that belong to class (T) (such as the quadratic bottleneck assignment problem (QBAP) (cf. Section 3), location problems on graphs and network flow problems (see e.g. [4, 8])). The linear bottleneck assignment problem is not an element of (T) (cf. [6]). Burkard and Fincke [4] have shown the remarkable result that under mild probabilistic assumptions on the random variables, bottleneck problems of type (1) belong to (T) if

$$\frac{\log |\mathcal{S}_n|}{s_n} = o(1) \tag{2}$$

Key words and phrases. Bottleneck Problems, Probabilistic Error Bounds.

* This research has been partially supported by the Spezialforschungsbereich F 003 "Optimierung und Kontrolle" / Projektbereich Diskrete Optimierung.

holds (see also [8]). As a by-product of this note it is shown in Remark 1 that criterion (2) is sharp in the sense that no weaker version of (2) can serve as a sufficient condition for a bottleneck problem to be an element of (T), which answers an old corresponding question.

In order to apply the property of asymptotic optimality of any greedy solution in practice, it is of course important to know the rate of convergence of the relative error. In [3], Burkard and Fincke have derived a bound on the convergence rate of the relative difference between the best and the worst feasible solution for the QBAP. In Section 2 we sharpen and generalize their approach and derive rather tight upper bounds on the convergence rate for arbitrary bottleneck problems and arbitrary bounded distributions for the random variables $f_n(e)$. As a by-product, their bound for the QBAP can be improved.

Given some rather weak combinatorial and probabilistic assumptions, condition (2) is also known to be a (sharp) sufficient condition for optimization problems with sum objective function to belong to (T) (cf. [4, 5, 7]). For a recent thermo-dynamic approach to this phenomenon, we refer to [1].

2. Main Result

Theorem 1. *Let $f_n(e)$ for all $e \in E_n$ and $n \in \mathbb{N}$ be identically distributed random variables in [0,M] and assume that $f_n(e)$, $e \in S$ are independent for every fixed feasible solution $S \in \mathcal{S}_n$, $n \in \mathbb{N}$. Let furthermore*

$$\log|\mathcal{S}_n| + s_n \log \mathbb{P}\left(f_n(e) \leq \frac{M}{1+g(n)}\right) \longrightarrow -\infty \quad \text{as} \;\; n \to \infty \tag{3}$$

for some positive function $g(n)$. Then

$$\mathbb{P}\left(\frac{F_{n,\max} - F_{n,\min}}{F_{n,\min}} \leq g(n)\right) = 1 - o(1), \tag{4}$$

where $F_{n,\max} = \max_{S\in\mathcal{S}_n} \max_{e\in S} f_n(e)$ is the worst possible solution of P_n.
If furthermore

$$\sum_{n=1}^{\infty} |\mathcal{S}_n| \Big(\mathbb{P}(f_n(e) \leq \frac{M}{1+g(n)})\Big)^{s_n} < \infty, \tag{5}$$

then $\frac{F_{n,\max}-F_{n,\min}}{F_{n,\min}} \leq g(n)$ holds almost surely, or equivalently

$$\mathbb{P}\left(\frac{F_{n,\max} - F_{n,\min}}{F_{n,\min}} > g(n) \textit{ infinitely often}\right) = 0. \tag{6}$$

Proof. It suffices to show that

$$\mathbb{P}\left(F_{n,\min} \leq \frac{F_{n,\max}}{1+g(n)}\right) = o(1).$$

Since $F_{n,\max} \leq M$, we have

$$
\begin{aligned}
\mathbb{P}\left(F_{n,\min} \leq \frac{F_{n,\max}}{1+g(n)}\right) &\leq \mathbb{P}\left(F_{n,\min} \leq \frac{M}{1+g(n)}\right) \\
&= \mathbb{P}\left(\exists\, S \in \mathcal{S}_n : \max_{e \in S} f_n(e) \leq \frac{M}{1+g(n)}\right) \\
&\leq |\mathcal{S}_n| \left(\mathbb{P}(f_n(e) \leq \frac{M}{1+g(n)})\right)^{s_n} \qquad (7) \\
&= \exp\left(\log|\mathcal{S}_n| + s_n \, \log \mathbb{P}\left(f_n(e) \leq \frac{M}{1+g(n)}\right)\right),
\end{aligned}
$$

and assertion (4) follows from (3).
Since the Borel-Cantelli lemma (see e.g. [2]) ensures $\mathbb{P}(A_n$ infinitely often$) = 0$ for arbitrary events A_n given that $\sum_n \mathbb{P}(A_n)$ converges, (6) finally follows from (5) and (7). □

An obvious necessary requirement on the distribution of $f_n(e)$ for (3) to hold is that $\mathbb{P}(f_n(e) \leq \frac{M}{1+g(n)}) < 1$ for finite n.

Remark 1: If we choose $g(n) = \epsilon$, then Theorem 1 can be used to determine whether a given bottleneck problem belongs to (T), which is equivalent to the fact that (4) holds for every $\epsilon > 0$ (cf. [4]). Since M is the supremum of $f_n(e)$ for all $n \in \mathbb{N}, e \in E_n$, the constant $\mathbb{P}(f_n(e) \leq \frac{M}{1+\epsilon})$ gets arbitrarily close to 1 for small $\epsilon > 0$. Hence (2) is a necessary and sufficient condition for (3) to hold for arbitrary $\epsilon > 0$. However, Theorem 1 relies on the generous estimate (7) and one might wonder whether a weaker condition than (2) could suffice to imply (4). The following counterexample shows that this is not possible:
Let $E_n = \{1, 2, \ldots, n2^n\}$ and partition E_n into 2^n blocks of length n. These disjoint blocks are the feasible solutions $S \in \mathcal{S}_n$. Let the weights $f_n(e)$, $e \in E_n$ be independent random variables with uniform distribution over [0,1]. We have $|\mathcal{S}_n| = 2^n$, $s_n = n$ and thus $\frac{\log|\mathcal{S}_n|}{s_n} = \log 2$, a constant. But $\mathbb{P}(F_{n,\max} \geq x) = 1 - x^{n2^n} = 1 - o(1)$ for each $x \in (0,1)$ and $\mathbb{P}(F_{n,\min} \leq x) = 1 - (1 - x^n)^{2^n} = 1 - o(1)$ for each $x \in (0,1)$, from which it follows that this bottleneck problem does not belong to the class (T). Hence, the combinatorial condition (2) can not be improved.

3. Application of Theorem 1

The main purpose of Theorem 1 is to provide a satisfying bound $g(n) = o(1)$ for the convergence rate of the relative difference between the worst and the optimal solution for any bottleneck optimization problem in (T). By virtue of (3), the best possible bound that can be achieved by relying on the estimate (7) can be determined. Qualitatively, the faster the growth of s_n, the stronger the convergence rate that can be guaranteed. For instance, if we assume that the random variables $f_n(e)$ are uniformly distributed on [0,1], condition (3) reduces to

$$\log|\mathcal{S}_n| - s_n \log(1 + g(n)) \longrightarrow -\infty \quad \text{as } n \to \infty.$$

Example: Let us consider the quadratic bottleneck assignment problem (QBAP) defined by

$$F^Q_{n,\min} = \min_{\varphi \in \mathcal{S}_n} \max_{1 \leq i,j \leq n} a_{ij} b_{\varphi(i)\varphi(j)},$$

where $\mathcal{S}_n$ is the set of all permutations of $\{1,\dots,n\}$ and $A=(a_{ij})$ and $B=(B_{ij})$ are two $n\times n$ matrices. This problem is NP-hard, but it belongs to (T), since $|\mathcal{S}_n|=n!$ and $s_n=n^2$ and thus (2) holds. Under the assumption that for $n\in\mathbb{N}$, a_{ij},b_{kl}, $i,j,k,l=1,\dots,n$ are independent uniform [0,1] random variables, Burkard and Fincke [3] have shown that (4) holds for

$$g_1(n)=\frac{1}{\sqrt{\frac{n}{2\log n}}-1}. \tag{8}$$

We can now use our general approach to check whether this bound can be improved. From the specific structure of the QBAP it follows that the random variables $f_n(e)$ are the product of two independent uniform [0,1]-variates, i.e. $\mathbb{P}(f_n(e)\le x)=x(1-\log x)$ $(0\le x\le 1)$. Thus condition (3) can be rewritten as

$$\log n!-n^2\log(1+g(n))+n^2\log(1+\log(1+g(n)))\longrightarrow -\infty \quad \text{as } n\to\infty.$$

An asymptotic expansion gives

$$n\,\log n-n+\frac{\log n}{2}-n^2\Big(\frac{1}{2}g^2(n)-\frac{5}{6}g^3(n)+\frac{29}{4}g^4(n)+\mathcal{O}(g^5(n))\Big)\longrightarrow -\infty \quad \text{as } n\to\infty, \tag{9}$$

from which it follows that the optimal dominating asymptotic order of $g(n)$ is determined by $\frac{n^2}{2}g^2(n)=n\log n$, i.e. $g(n)=\sqrt{\frac{2\log n}{n}}$. Since an asymptotic expansion of (8) gives

$$g_1(n)=\sqrt{\frac{2\log n}{n}}+\frac{2\log n}{n}+\mathcal{O}\left(\left(\frac{2\log n}{n}\right)^{\frac{3}{2}}\right),$$

the dominating asymptotic term of the bound (8) found by Burkard and Fincke is optimal. However, asymptotic terms of higher order can be considerably improved. A detailed analysis of (9) gives the bound

$$g_2(n)=\sqrt{\frac{2\log n}{n}}\left(1-\frac{1}{2\log n}-\frac{1}{8(\log n)^2}+\mathcal{O}\Big(\frac{1}{(\log n)^3}\Big)\right),$$

where the constants in the higher order terms are optimal. Since the asymptotic behaviour of (9) with $g(n)=g_2(n)$ is given by $-\frac{n}{8(\log n)^2}+\mathcal{O}(\frac{n}{(\log n)^3})$, we can apply the Borel-Cantelli lemma to obtain the stronger result

Corollary 1. *For the QBAP as defined above we have*

$$\frac{F^Q_{n,\max}-F^Q_{n,\min}}{F^Q_{n,\min}}\le\sqrt{\frac{2\log n}{n}}\left(1-\frac{1}{2\log n}-\frac{1}{8(\log n)^2}\right) \quad \textit{almost surely.}$$

Acknowledgements. The author would like to thank Professor R. Burkard for valuable discussions on the topic.

References

[1] H. Albrecher, R. E. Burkard, and E. Cela. An asymptotical study of combinatorial optimization problems by means of statistical mechanics. *SFB Report 267, Graz University of Technology*, 2003.
[2] P. Billingsley. *Probability and Measure.* Third edition. Wiley, New York, 1995.
[3] R. E. Burkard and U. Fincke. On random quadratic bottleneck assignment problems. *Math. Programming*, 23(2):227–232, 1982.
[4] R. E. Burkard and U. Fincke. Probabilistic asymptotic properties of some combinatorial optimization problems. *Discrete Appl. Math.*, 12(1):21–29, 1985.

[5] J. B. G. Frenk, M. van Houweninge, and A. H. G. Rinnooy Kan. Asymptotic properties of the quadratic assignment problem. *Math. Oper. Res.*, 10(1):100–116, 1985.
[6] U. Pferschy. The random linear bottleneck assignment problem. *RAIRO Rech. Opér.*, 30(2):127–142, 1996.
[7] W. T. Rhee. Stochastic analysis of the quadratic assignment problem. *Math. Oper. Res.*, 16(2):223–239, 1991.
[8] W. Szpankowski. Combinatorial optimization problems for which almost every algorithm is asymptotically optimal! *Optimization*, 33(4):359–367, 1995.

(H. Albrecher) Department of Mathematics, Graz University of Technology, Steyrergasse 30, A-8010 Graz, Austria
E-mail address: `albrecher@TUGraz.at`