A statistical analysis of the knapsack problem

1995 J. Phys. A: Math. Gen. 28 4751

(http://iopscience.iop.org/0305-4470/28/17/011)

J. Phys. A: Math. Gen. 28 (1995) 4751–4759. Printed in the UK

A statistical analysis of the knapsack problem

J F Fontanari

Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos SP, Brazil

Received 31 March 1995, in final form 5 June 1995

Abstract. We investigate the dependence of the multi-knapsack objective function on the knapsack capacities and on the number of capacity constraints P, in the case when all N objects are assigned the same profit value and the weights are uniformly distributed over the unit interval. A rigorous upper bound to the optimal profit is obtained employing the annealed approximation and then compared with the exact value obtained through the Lagrangian relaxation method. The analysis is restricted to the regime where N goes to infinity and P remains finite.

1. Introduction

The simplest and probably the most studied integer programming problem is the knapsack problem (Salkin 1975, Martello and Toth 1990). This interest stems from the many important applications involving the knapsack problem such as capital budgeting, cutting stock and cargo loading, to mention only the most classical applications, as well as from its simple structure that allows fairly simple theoretical analysis. The problem consists of selecting the set that maximizes the total profit from among a collection of N objects $i = 1, 2, \ldots, N$, to each one of which is associated a profit $c_i \geqslant 0$ and weights $a_{ki} \geqslant 0\,(k = 1, \ldots, P)$, and subject to P linear capacity constraints. More specifically, introducing N binary variables s_i so that $s_i = 1$ if object i is selected and $s_i = 0$ otherwise, the problem is to find the vector s that maximizes the objective function or total profit

$$E(s) = \sum_{i}^{N} c_i s_i \tag{1.1}$$

subject to

$$\sum_{i}^{N} a_{ki} s_i \leqslant b_k \qquad k = 1, \ldots, P. \tag{1.2}$$

Here a_{ki} is the weight of object i in knapsack k and b_k is the capacity of knapsack k. This problem, the so-called multi-knapsack problem, is a generalization of the standard knapsack problem ($P = 1$). Both problems belong to the NP-complete class, which basically means that the computational cost of any known deterministic algorithm for finding their exact solutions grows exponentially with the number of variables N (Garey and Johnson 1979).

In this paper we consider the case where the a_{ki} are statistically independent random variables uniformly distributed over $[0, 1]$, $c_i = 1$ for all i and $b_k = b$ for all k. The practical motivation for studying this type of random problem is for their widespread use

0305-4470/95/174751+09$19.50

in the empirical evaluation of heuristic algorithms (Martello and Toth 1990, Ohlsson *et al* 1993). In this sense, analytical bounds for the optimal value of the objective function are highly desirable as they furnish valuable information used to evaluate the quality of the solutions found by heuristic algorithms. Since in practical applications the number of objects typically ranges from 10^4 to 10^5, the asymptotic limit $N \to \infty$, to which our analysis is restricted, is not a too unrealistic assumption.

Statistical mechanics techniques developed in the study of the infinite-range Ising spin-glass model—the SK model (Sherrington and Kirkpatrick 1975, Binder and Young 1986, Mèzard *et al* 1987)—have been employed successfully in the analysis of combinatorial optimization problems. In particular, stochastic versions of classical problems such as the graph partitioning problem (Fu and Anderson 1986), the weighted matching problem (Mèzard and Parisi 1985) and the travelling salesman problem (Vannimenus and Mèzard 1984, Mèzard and Parisi 1986) have been investigated within the replica formalism. The main technical difficulty in using the replica approach to study these problems, as well as the knapsack problem, is the appearance of order parameters that measure the correlations between an arbitrary number of replicas. This makes the analysis very involved when compared with that of the SK model, which has equilibrium properties determined by an order parameter that measures the correlations between two replicas only. In a somewhat different vein, Prügel-Bennett and Shapiro (1994) have succeeded in analysing the performance of a class of general purpose heuristic algorithms—genetic algorithms (Goldberg 1989)—within the statistical mechanics framework. More recently, another method from statistical physics, the finite-size scaling, has proved very useful in the study of threshold phenomena in random graphs and in the k-satisfiability problem (Kirkpatrick and Selman 1994).

Our goal in this paper is to determine the dependence of the optimal profit density $\epsilon_m = E_m/N$ on the knapsack capacities $b = \beta N$ and on the number of constraints P. This problem has recently been investigated through statistical mechanics techniques by Korutcheva *et al* (1994) in the special case where P grows linearly with N and $\beta = 1/4$. As we will see below, however, $\epsilon_m \to 2\beta$ for $P \to \infty$, so this is a rather uninteresting limit as far as the dependence of ϵ_m on β is concerned. Actually, those authors focused on the analysis of the $\mathcal{O}(1/\sqrt{N})$ fluctuations to the average value of ϵ_m. The more interesting case, where P is finite, was investigated by Meanti *et al* (1990) using the Lagrangian relaxation technique (Held and Karp 1970, 1971). However, explicit results were presented only for the cases $P = 1$ and $P = 2$, mainly because those authors considered knapsacks of different capacities.

To tackle the multi-knapsack problem in the finite P case we follow two approaches. First, we employ the annealed approximation within the microcanonical ensemble formalism of statistical mechanics to obtain an upper bound to ϵ_m (Fontanari and Meir 1993). This is a very general technique for generating bounds for maximization or minimization problems that involve a search in a space of *discrete* variables. We note that a similar method was developed in the canonical ensemble formalism by Vannimenus and Mèzard (1984) in the study of the travelling salesman problem. Second, building on the results of Meanti *et al* (1990) we use the Lagrangian relaxation technique to obtain the dependence of ϵ_m on β *for all* P. Besides presenting these original results, the purpose of this paper is to draw the attention of the statistical physics community to the powerful, and well known to the operations research community, Lagrangian relaxation technique to generate rigorous bounds to optimization problems. In particular, although the multi-knapsack problem has recently been considered in physics literature (Korutcheva *et al* 1994, Peterson 1993), there has been no reference to the important results obtained with that technique (Meanti *et al* 1990).

The remainder of this paper is organized as follows. In section 2 we calculate an upper bound to the optimal profit density ϵ_m employing the annealed approximation in the microcanonical ensemble. In section 3 we follow Meanti *et al* (1990) to obtain exact estimates for ϵ_m using the Lagrangian relaxation technique. Finally, in section 4 we discuss our results and present some concluding remarks.

2. Annealed approximation

The basic quantity we must evaluate in the microcanonical ensemble formalism of the statistical mechanics is $\mathcal{N}(E)$, the number of feasible vectors s that possess profit E. It is given by

$$\mathcal{N}(E) = \sum_{\{s\}} \delta\left(E, \sum_i s_i\right) \prod_k \Theta\left(b - \sum_i a_{ki} s_i\right) \tag{2.1}$$

where the summation is over the 2^N binary vectors s, $\Theta(x) = 1$ if $x \geqslant 0$ and 0 otherwise, and $\delta(k, l)$ is the Kronecker delta. A vector s is feasible if it satisfies the constraints (1.2).

According to the method of taking averages over extensive, and therefore self-averaging, quantities only (Binder and Young 1986), the physically meaningful average quantity associated to $\mathcal{N}$ is the average entropy $\mathcal{S}(E) = \langle \ln \mathcal{N}(E) \rangle$, where $\langle \ldots \rangle$ stands for the averages over the random variables a_{ki}. The optimal profit E_m is determined by the conditions $\mathcal{S}(E > E_m) \to -\infty$ and $\mathcal{S}(E \leqslant E_m) \geqslant 0$. The evaluation of the quenched averages, however, involves enormous technical difficulties, so we consider instead the much easier to calculate, though physically meaningless, annealed entropy

$$\mathcal{S}_a(E) = \ln \langle \mathcal{N}(E) \rangle. \tag{2.2}$$

In particular, while $\mathcal{S}$ is clearly non-negative, $\mathcal{S}_a$ can take on negative values. The usefulness of $\mathcal{S}_a$ stems from the inequality $\mathcal{S}_a > \mathcal{S}$, which implies that the profit E_m^a so that $\mathcal{S}_a(E_m^a) = 0$ is an upper bound to E_m (Fontanari and Meir 1993).

We now proceed with the explicit calculation of the annealed entropy, equation (2.2). Using the integral representations of the theta and the Kronecker delta functions, it is straightforward to perform the averages over a_{ki} and the summation over s. The final result is

$$\mathcal{S}_a(E) = \ln\left[\frac{N!}{E!\,(N-E)!}\right] + P \ln\left[\int_{-\infty}^{b} \frac{dx}{\pi} \int_{-\infty}^{\infty} dy\, e^{iy(2x-E)} \left(\frac{\sin y}{y}\right)^E\right]. \tag{2.3}$$

Although the integrations can easily be evaluated for any finite $E = \epsilon N$ (Erdélyi 1954), the exact result, expressed as a finite alternating series of terms containing combinatorial factors and powers of N, becomes useless for any practical purpose since its evaluation for N not too small, say $N = 10$, requires the use of enormous numerical precision. Moreover, taking the asymptotic limit $N \to \infty$, which requires the use of Stirling's formula, is hopeless since the series is alternating. Thus, we decide to take the large N limit already in equation (2.3), which allows us to perform the integrations by the method of steepest descents. We present only the final result, relegating the details of the saddle-point integration to the appendix.

For $\epsilon \geqslant 2\beta$, the average density of annealed entropy $s_a = \mathcal{S}_a/N$ is given by

$$s_a(\epsilon) = -\epsilon \ln \epsilon - (1-\epsilon)\ln(1-\epsilon) + P\epsilon\left(-kz_k + \ln \frac{\sinh z_k}{z_k}\right) \tag{2.4}$$

where z_k is the unique solution of the equation

$$k - \coth z_k + \frac{1}{z_k} = 0 \tag{2.5}$$

and $k = 2\beta/\epsilon - 1$. For $\epsilon < 2\beta$, the constraints play no role at all, being automatically satisfied by any vector s in the asymptotic limit, and so s_a is given simply by the first two terms of equation (2.4). Thus, since the upper bound ϵ_m^a is given by the largest solution of $s_a(\epsilon) = 0$ we find $\epsilon_m^a = 1$ in this case. Note the trivial role played by P in the derivation of equation (2.4). In figure 1 we present ϵ_m^a as a function of $\beta < 1/2$ for $P = 1, 2, 10$ and 100.

3. Lagrangian relaxation

Lagrangian relaxation is a general technique developed by Held and Karp (1970, 1971) for generating bounds to the objective function of combinatorial optimization problems that involve constraints (see Beasley (1993) for a recent review on this subject). The basic idea is to attach Lagrangian multipliers to some set of constraints and then relax these constraints into the objective function. In our specific problem, we relax the constraints (1.2) by introducing P Lagrange multipliers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_P)$ into the objective function (1.1), which is then written as

$$L(\lambda, s) = \sum_i s_i + \sum_k \lambda_k \left(b - \sum_i a_{ki} s_i \right) \tag{3.1}$$

where $\lambda_k \geqslant 0$. As we will see below, the relaxation of the integrality constraint so that $0 \leqslant s_i \leqslant 1$ also has no effect in our analysis. To show that L does in fact give an upper bound to E for any λ we consider the problem of maximizing

$$M(\lambda, s) = \sum_i s_i + \sum_k \lambda_k \left(b - \sum_i a_{ki} s_i \right) \tag{3.2}$$

subject to $b - \sum_i a_{ki} s_i \geqslant 0$. Since $\lambda_k \geqslant 0$, the second term of M is positive for any feasible vector s and, therefore, $M \geqslant E$. Finally, by noting that the relaxation of constraints (1.2) can only increase the objective function we conclude that $L \geqslant M \geqslant E$.

The problem then becomes, for a given λ, one of finding the vector s that maximizes L. Since there are no constraints, its solution is simply $s_i = 1$ if $1 - \sum_k a_{ki}\lambda_k \geqslant 0$, and $s_i = 0$ otherwise, which yields

$$L(\lambda) = b \sum_k \lambda_k + \sum_i \left(1 - \sum_k a_{ki} \lambda_k \right) \Theta \left(1 - \sum_k a_{ki} \lambda_k \right). \tag{3.3}$$

To make this upper bound as tight as possible, we must choose the Lagrangian multipliers λ so as to *minimize* L. At this point we can already realize the main benefit of the Lagrangian relaxation method: the search in the discrete N-dimensional space of s is replaced by a search in the continuous P-dimensional space of λ. This, of course, can be a great advantage when N is large and P is small. The disadvantage is that, in general, we obtain only an upper bound for the optimal profit.

Since the components of the optimal solution s are integral, the relaxation of the integrality requirement has no effect in our analysis. The Lagrangian relaxation of the knapsack problem considered above is then said to have the *integrality* property. This fact has an important consequence: if L_m denotes the minimal value the upper bound $L(\lambda)$ can take, then L_m will be equal to the value of the profit obtained with the linear programming (LP) relaxation of the original problem (Beasley 1993). In the LP relaxation, we relax the integrality requirement on the variables s_i, so that $0 \leqslant s_i \leqslant 1$, and maximize the profit (1.1) subject to constraints (1.2) using some standard algorithm (simplex or interior point). As pointed out by Meanti *et al* (1990), the solution of the LP relaxation has at most P variables that are neither 0 nor 1. Thus, rounding down the values of these variables yields a feasible vector s which has a corresponding profit value, given by L_m, decreased by at most P. Therefore, we have the inequalities

$$L_m \geqslant E_m \geqslant L_m - P. \tag{3.4}$$

Dividing by N and taking the limit $N \to \infty$ with P finite, we conclude that $l_m = L_m/N$ converges with probability one to $\epsilon_m = E_m/N$ for any realization of the random variables a_{ki}. Furthermore, in this limit the statistical independence of the random variables a_{ki} guarantees that $l(\lambda) = L(\lambda)/N$ converges with probability one to its mean value $\langle l(\lambda)\rangle/N = \langle l\rangle(\lambda)/N$ for all λ. This result is due to the property of self-averageness of the *extensive* quantity $L(\lambda)$. In particular, if λ^* is the vector of Lagrangian multipliers that minimizes $\langle l\rangle(\lambda)$ then it follows that l_m, and therefore ϵ_m, will converge with probability one to $\langle l\rangle(\lambda^*)$. We refer the reader to Meanti *et al* (1990) for a rigorous proof of these results.

The above mentioned results make the evaluation of E_m rather straightforward: first we must carry out the averages over a_{ki} in equation (3.3) and then minimize the resulting expression, $\langle L\rangle$, with respect to λ. This was the procedure adopted by Meanti *et al* (1990) who, however, considered knapsacks of different capacities, i.e. $b_k = \beta_k N$. This makes the minimization with respect to λ more complicated and adds, besides the evaluation of the averages, another serious technical difficulty to the analysis. As a consequence, explicit results were presented only for $P = 1$ and $P = 2$.

Since in this paper we consider $\beta_k = \beta$ for all k, we can make the simplifying assumption that $\lambda_k^* = \lambda^*$ for all k because the statistically independent random variables a_{ki} are identically distributed. With this assumption we write ϵ_m as

$$\epsilon_m = P\beta\lambda^* + \left\langle \left(1 - \lambda^* \sum_k a_k\right) \Theta\left(1 - \lambda^* \sum_k a_k\right) \right\rangle \tag{3.5}$$

with λ^* given by the solution of $\partial\epsilon_m/\partial\lambda^* = 0$. The direct evaluation of the averages in equation (3.5) is simple only for small P. For instance, for $P = 1$ we find

$$\epsilon_m = \sqrt{2\beta} \qquad \beta \leqslant 1/2 \tag{3.6}$$

while for $P = 2$ we find

$$\epsilon_m = \frac{3}{6^{1/3}}\beta^{2/3} \qquad \beta \leqslant 1/6 \tag{3.7}$$

and

$$\epsilon_m = -\tfrac{1}{2}x^2 + 2x - 1 \qquad 1/6 < \beta \leqslant 1/2 \tag{3.8}$$

where $x \in (1, 2)$ is a root of the cubic equation

$$x^3 - 3x^2 + 1 + 6\beta = 0. \tag{3.9}$$

The calculations for $P > 2$ are too cumbersome so we try to evaluate the averages for general P by first extracting the random variable $\sum_k a_k$ from the argument of the theta function using a delta function, and then performing the averages. Carrying out this procedure yields

$$\epsilon_m = P\beta\lambda^* + \frac{1}{\pi}\int_{-\infty}^{1/\lambda^*} dx\,(1-\lambda^* x)\int_{-\infty}^{\infty} dy\, e^{iy(2x-P)}\left(\frac{\sin y}{y}\right)^P. \tag{3.10}$$

Though we could use a procedure similar to that presented in the appendix to evaluate the double integral for large P, we follow a much simpler approach based on the approximation

$$\left(\frac{\sin y}{y}\right)^M \approx \exp(-My^2/6) \tag{3.11}$$

which is valid for large M. With this approximation, the integrals in equation (3.10) become Gaussian and, therefore, can easily be performed, yielding

$$\epsilon_m = \frac{1}{\xi + \sqrt{3P}}\left[2\beta\sqrt{3P} + \xi H(-\xi) + \frac{e^{-\xi^2/2}}{\sqrt{2\pi}}\right] \tag{3.12}$$

where

$$\xi = \frac{1/\lambda^* - P/2}{\sqrt{P/12}} \tag{3.13}$$

is the solution of the equation

$$\sqrt{3P}[H(-\xi) - 2\beta] - \frac{e^{-\xi^2/2}}{\sqrt{2\pi}} = 0. \tag{3.14}$$

Here, $H(x) = \int_x^\infty dt/\sqrt{2\pi}\exp\left(-t^2/2\right)$. This equation possesses a unique solution, except at $\beta = 0$ where the spurious solution $\xi \to -\infty$ is also present.

Although equation (3.12) is in principle valid only for large P, in figure 2 we compare the results obtained for $P = 1$ and $P = 2$ using this equation with the exact results obtained using equations (3.6)–(3.8). The agreement between them is remarkably good for β not too small and, in the scale of figure 1, they are practically indistinguishable for $P = 2$. Of course, this agreement is expected to become even better for larger P. Hence, we consider that following the laborious saddle-point integration procedure of the appendix (the prefactors cannot be neglected in this case) is not necessary. We must emphasize, however, that the correct procedure to evaluate the double integral in the limit of large P is the one given in the appendix. In particular, applying approximation (3.11) to the annealed calculation of section 2 gives very poor results for small β and P. However, the results become better as P increases and, in the scale of figure 1, they are indistinguishable from the results obtained through the saddle-point integration for $P > 5$.

Together with the upper bound ϵ_m^a, we present in figure 1 the exact estimate for ϵ_m for $P = 1$ and $P = 2$ obtained using equations (3.6)–(3.8), and the approximate estimate for $P = 10$ and $P = 100$ obtained using equation (3.12). Since ϵ_m reaches its maximal value at $\beta = 1/2$ it will obviously remain at that value for $\beta > 1/2$. We note that $\epsilon_m \to 2\beta$ for $P \to \infty$.

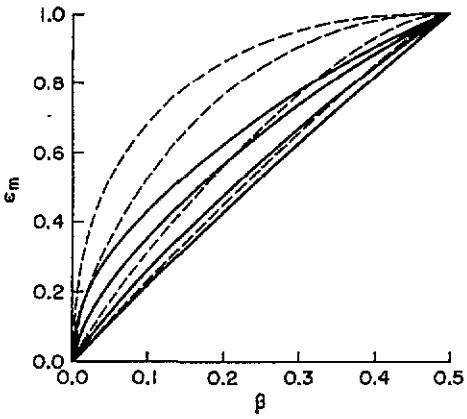

Figure 1. Optimal profit density ϵ_m as a function of the knapsack capacities β for, from top to bottom, $P = 1, 2, 10$ and 100. The broken curves are the annealed upper bounds ϵ_m^a. The full curves are the exact ($P = 1, 2$) and the approximate ($P = 10, 100$) estimates for ϵ_m.

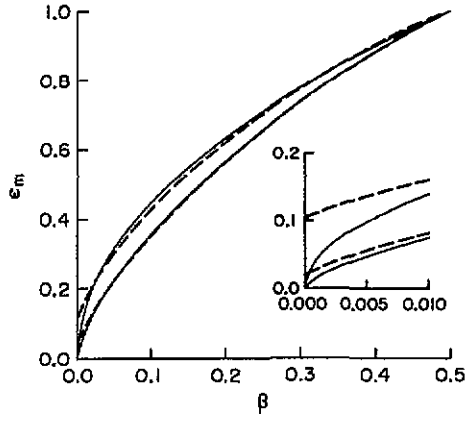

Figure 2. Exact (full curves) and approximate (broken curves) estimates for ϵ_m for, from top to bottom, $P = 1$ and $P = 2$.

4. Conclusion

It is very unfortunate that the annealed approximation, which seems to be the sole general tool of statistical mechanics to generate *rigorous* bounds to integer programming problems, gives such poor results for the multi-knapsack objective function. This conclusion follows from the comparison, presented in figure 1, between the annealed upper bound and the exact value of the optimal profit calculated via the Lagrangian relaxation method. It should be noted, however, that the annealed calculations can easily be extended to the regime where P depends on N (Korutcheva *et al* 1994), whereas the results of Meanti *et al* (1990), on which the Lagrangian relaxation approach is based, are not valid in this regime.

The main result of this paper, which we hope may be of practical use in the evaluation of the performance of heuristic algorithms, is equation (3.12) which, within a reasonable degree of accuracy, gives the optimal profit density as a function of the knapsack capacities β for any finite number of capacity constraints P.

Acknowledgments

It is a pleasure to thank Lidério Ioriatti for useful discussions on the method of steepest descents. This work was supported in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Appendix

In this appendix we evaluate explicitly the double integral

$$\Gamma = \int_{-\infty}^{\beta N} \frac{dx}{\pi} \int_{-\infty}^{\infty} dy\, e^{iy(2x-E)} \left(\frac{\sin y}{y}\right)^E \tag{A1}$$

in the limit $E = \epsilon N \to \infty$. A simple change of variables allows us to rewrite this equation as

$$\Gamma = \frac{E}{2\pi} \int_{-\infty}^{2\beta/\epsilon - 1} \mathrm{d}k \int_{-\infty}^{\infty} \mathrm{d}y \, \mathrm{e}^{EG(k,y)} \tag{A2}$$

where

$$G(k, y) = iky + \ln \frac{\sin y}{y}. \tag{A3}$$

The integration over y can then readily be performed using the method of steepest descents. For $| k | < 1$, the saddle-point is the imaginary $y = y_k = iz_k$, where z_k is the solution of $k - \coth z_k + 1/z_k = 0$. Hence, the function $G(k, y_k) = G(k)$, where

$$G(k) = -kz_k + \ln \frac{\sinh z_k}{z_k} \tag{A4}$$

is real. For $|k| > 1$, there is no saddle-point and the integral vanishes due to the rapid oscillations of the integrand. This last result is in agreement with the result of the exact integration over y (Erdélyi 1954). Thus, equation (A2) reduces to

$$\Gamma = \left(\frac{E}{2\pi}\right)^{1/2} \int_{-1}^{\min[2\beta/\epsilon - 1, 1]} \mathrm{d}k \, \frac{\mathrm{e}^{EG(k)}}{[-G''(k)]^{1/2}} \tag{A5}$$

where $G''(k) = k^2 - 1 + 2k/z_k$. This integral can be evaluated using the method of Laplace. The real function $G(k)$ is negative for all k except for $k = 0$ where it assumes its maximal value, $G(0) = 0$. The value of the integral depends, then, on whether $k = 0$ belongs or not to the interval of integration. In the case when $\epsilon < 2\beta$ it does belong, and the integration yields

$$\lim_{N \to \infty} \frac{1}{N} \ln \Gamma = 0. \tag{A6}$$

For $\epsilon > 2\beta$, the maximum of G occurs at the upper extreme of integration, yielding

$$\lim_{N \to \infty} \frac{1}{N} \ln \Gamma = \epsilon G \, (2\beta/\epsilon - 1) \, . \tag{A7}$$

References

Beasley J E 1993 *Modern Heuristic Techniques for Combinatorial Problems* ed C R Reeves (New York: Wiley)

Binder K and Young AP 1986 *Rev. Mod. Phys.* **58** 801

Erdélyi A 1954 *Tables of Integral Transforms* (New York: McGraw-Hill)

Fontanari J F and Meir R 1993 *J. Phys. A: Math. Gen.* **26** 1077

Fu Y T and Anderson P W 1986 *J. Phys. A: Math. Gen.* **19** 1605

Garey M R and Johnson D S 1979 *Computers and Intractability: A Guide to the Theory of NP-Completeness* (San Francisco: Freeman)

Goldberg D E 1989 *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading, MA: Addison-Wesley)

Held M and Karp R M 1970 *Operations Res.* **18** 1138

Held M and Karp R M 1971 *Mathematical Programming* **1** 6
Kirkpatrick S and Selman B 1994 *Science* **264** 1297
Korutcheva E, Opper M and López B 1994 *J. Phys. A: Math. Gen.* **27** L645
Martello S and Toth P 1990 *Knapsack Problems* (New York: Wiley)
Meanti M, Rinnooy Kan A H G, Stougie L and Vercellis C 1990 *Mathematical Programming* **46** 237
Mèzard M and Parisi G 1985 *J. Phys. Lett.* **46** L771
—— 1986 *J. Physique* **47** 1285
Mèzard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
Ohlsson M, Peterson C and Söderberg B 1993 *Neural Computation* **5** 331
Peterson C 1993 *Physica* A **200** 570
Prügel-Bennett A and Shapiro J L 1994 *Phys. Rev. Lett.* **72** 1305
Salkin H M 1975 *Integer Programming* (Reading, MA: Addison-Wesley)
Sherrington D and Kirkpatrick S 1975 *Phys. Rev. Lett.* **35** 1793
Vannimenus J and Mèzard M 1984 *J. Phys. Lett.* **45** L1145