HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Computer Science
Laboratory for Theoretical Computer Science
T-79.300 Postgraduate Course in Theoretical Computer Science

Timo Latvala

Landscape Families

Contents

1 Introduction

Fitness landscapes is a concept which stems from theoretical biology. It was used as a metaphor for describing population evolution and explaining the link between fitness of genes and multiple population attractors in the space of genotypes.

Currently, fitness landscapes and landscape theory are under active research in physics of disordered systems and combinatorial optimisation besides theoretical biology.

The intuitive idea of landscape is easy. The familiar travelling salesman problem spans a fitness landscape in the following way. Each possible tour of the cities is a point the space of configurations. Other tours can be obtained by e.g. switching the order of two cities. The length or cost of the tour spans a landscape in which one can move by switching the order of two cities.

More formally, the key ingredients of a fitness landscape are

1. a set X of configurations,
2. a notion of neighbourhood, nearness, distance, or accessibility on X, and
3. a fitness function $f : X \to \mathbb{R}$

This work will describe three families of landscapes. The first family comes from the realm of physics, namely spin glasses. Properties of the Ising model with its different variants are covered. The N-K landscapes represent a model from theoretical biology. Both complexity related issues and landscape properties, especially related to adaptive walks, are discussed. Optimisation problems is the last family. Here we focus on describing the properties of the landscape.

2 Spin Glasses

Physics has long sought to study the structure and properties of matter. The field of condensed matter physics is especially interested in the different properties of solids. Our first landscape family comes from this domain.

Spin glasses are magnetic substances which have complicated magnetic interactions. They can exhibit so called ferromagnetic and anti-ferromagnetic interaction, sometimes even both. Spin glasses are also called amorphous magnets. In most cases amorphous substances show structure only for a few tens atom lengths unlike e.g. regular crystalline structures which can be completely regular. The dividing line between amorphous substances and polycrystalline substances is, however, not clear cut.

A model for the magnetic interaction of spin glasses from which properties of the materials can fairly easily be described. The system is modelled as a crystal lattice with N lattice sites. Each site i is assigned a vector quantity $\mathbf{s}_i$, the magnetic spin. The sites interact according to a *coupling constant* J_{ij}. The coupling constant defines the neighbourhood of a site. If $J_{ij} = 0$, the sites i and j are not neighbours. A *configuration* of the system is an assignment to the N magnetic spins. The energy, or the *Hamiltonian*, of the system at particular state $\mathbf{s}$ is given by the equation:

$$H(\mathbf{s}) = -\sum_{i>k} J_{ij}\mathbf{s}_i \cdot \mathbf{s}_k \tag{1}$$

Finding the minimum energy configuration, the *ground state*, is one of the most important problems in statistical physics. We get the following problem.

Definition. The problem GROUND STATE is to find the minimum energy configuration given coupling constants and the problem geometry.

Clearly the spin glass model has all the ingredients of a fitness landscape. The Hamiltonian of the model spans a landscape which has been studied intensively in statistical mechanics.

This general model has several subclasses. In many cases the magnetic spins can only take on two values ± 1. Usually also the possible values for the coupling constant are restricted.

2.1 Ising Models

The model presented above is usually too general for analysis in any manner, which is why simpler forms are usually analysed. In Ising models the spins are restricted to $s_i = \pm 1$ and the geometries are simple. The values of the coupling constant are usually limited to $\{-J, 0, J\}$, where J is a positive integer. The models are named after the German physicist Ising, who studied magnetic properties of materials in the 1920s.

The Ising model can be cast into graph theoretic terms. Consider a graph $G_N = (V, E)$ with N vertexes. Each vertex $i \in V$ is associated with magnetic spin s_i. Each edge $\{i, j\}$ is labelled with a coupling constant J_{ij}. The coupling constant can take on a positive, negative or zero value. The energy of a state $s = (s_1, s_2, \ldots, s_N)$ is given by:

$$H(s) = -\sum_{\{i,j\} \in E} J_{ij} s_i s_j. \tag{2}$$

Here each vertex corresponds to a lattice site and the edges between the vertexes correspond to the bonds.

One can define a cut on the lattice graph in the following way. Let $C^+ = \{i \mid s_i = 1\}$ and $C^- = \{i \mid s_i = -1\}$. These two sets partition the set of vertexes, i.e. $V = C^+ \cup C-$. Similarly the edges can be partitioned into three sets. Let E^+ be the edges with endpoints only in C^+ and E^- in C^- respectively. The edges which go between the two partitions are denoted $E^\pm$. The *weight* of a cut $C = (C^+, C^-)$ is defined as

$$weight(C) = \sum_{\{i,j\} \in E^\pm} J_{ij}$$

Each configuration s corresponds to a specific cut C. Using the notation of cuts the Hamiltonian of the system can be rewritten to another form.

$$\begin{aligned} H(C) = & \quad -\sum_{\{i,j\} \in E^+} J_{ij} - \sum_{\{i,j\} \in E^-} J_{ij} - \sum_{\{i,j\} \in E^\pm} J_{ij} = \\ = & \quad -\sum_{\{i,j\} \in E} J_{ij} + 2weight(C) \end{aligned}$$

Thus minimising the weight of the cut is equivalent to finding the ground state of the spin glass. In arbitrary graphs, minimising the weight of the cut is NP-complete when both positive and negative weights are allowed. We have proved the following theorem (see e.g. [3]).

Theorem. Solving the GROUND STATE problem is NP-complete for arbitrary graphs.

In the special case the coupling constants are positive (the so called ferromagnetic case), the problem is polynomial even for arbitrary graphs, because then problem is equivalent to the well-known polynomial maximum flow problem.

In addition to finding the ground state of a system, physicists are also inter-

ested in the so called *partition function*. It is defined as

$$Z = \sum_{\{s\}} exp(-\beta H(s)). \tag{3}$$

The partition function is of interest because it in one sense completely characterises the system. With knowledge of the partition function the ground state among other things can be efficiently computed. Computing the partition function is at least as hard as finding the ground state or the maximum energy state [3].

A naive approach to solving to compute the partition function would require 2^N summations, one for each configuration. In the general case we cannot hope do better due to the previous NP-completeness result. The situation is even worse than that as Jerrum and Sinclair [4] proved that computing the partition function is $\#P$-complete, even when only positive couplings are allowed.

By examining simpler cases exact analytical solutions have been found to some subclasses of the partition function problem. When the geometry of the problem is one dimensional it is possible to find an exact solution. In the 2D case a solution has been presented for the case when the lattice graph is planar. For decades, generalisations were searched but nothing was found. The 3D version of the problem was proved NP-complete in 1982 by Barhanona [2].

It turned out however that dimensionality was not critical for the complexity of the problem. Remembering that the MINIMUM WEIGHT CUT problem is polynomial when the graph is planar, Istrail [3] showed that the problem is NP-complete for every on planar crystal lattice graph. Then he proved that many general geometries contains a Kuratowksian subgraph, a well-known non-planar graph, effectively proving their NP-completeness.

Theorem. Computing the partition function is NP-hard for every non-planar crystal lattice.

The state of the research is that all problems with exact analytical solutions can be computed in polynomial time.

Several approximation schemes have been developed for finding the ground state and partition function. These include for instance the well-known simulated annealing algorithm (see e.g. [6]).

2.2 Limit Solutions

Although several of the problem in statistical mechanics are intractable, the spin glass problem has also been investigated from a probabilistic view point.

The methods of statistical mechanics provide answers to many of the problems presented earlier. The arguments are probabilistic in nature and therefore only apply when special conditions are met.

The methods work by letting the states of the system adhere to the following probability distribution.

$$P(s) = exp(-\beta H(s))/Z \tag{4}$$

Low energy states have the highest probability in this distribution. The coupling constants must also adhere to some probability distribution. Thus these methods cannot say anything about a specific spin glass system. The solutions are also such that all the results apply only in the limit when $N \rightarrow \infty$. This limit is referred to as the *thermodynamical limit*.

The different methods are explained and analysed e.g. [6]. Results which have been obtained include e.g. that if J_{ij} is 1 or -1 with equal probability the minimum of the Hamiltonian takes on the value $-.7633N^{3/2}$.

These same methods are also applicable to other combinatorial optimisation problems.

3 N-K Fitness Landscapes

N-K fitness landscapes is a family of a landscapes introduced by Stuart Kauffman [5]. The landscapes were presented as a landscape family where the ruggedness – the number of local optima – could easily be tuned. Models of the evolution of genotypes is similar but somewhat more complex than the N-K model. The N-K model has, however, been used to study the qualitative properties of these more complex fitness landscapes of genotypes. In evolutionary computation N-K landscapes are both used as examples and test cases.

3.1 Definitions

The N-K model can be seen as a simple way of generating a tunable fitness function on bit strings. The bit strings $\mathbf{x} = x_1 x_2 \ldots x_N$ consist of N bits. A fitness function f is constructed by generating N component functions f_i. Each f_i depends on $K+1$ bits. In the *adjacent* model f_i depends on i:th bit and K adjacent bits. If $i + K > N$ the addition is computed modulo N. In the *random* model f_i depends on the i:th bit and K other randomly chosen bits. These two variants were originally studied by Kauffman [5]. Wright et al. [12] further introduced the *arbitrary* model where f_i depends only on $K+1$ randomly chosen bits and not necessarily on the i:th bit.

The values for the component functions are constructed by for each function sampling 2^{K+1} values from a random distribution, which usually is Gaussian. The functions can e.g. be implemented as a look-up table which has one entry for each of the 2^{K+1} possible bit strings. The fitness function f is the average of the component functions.

$$f(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} f_i(x_{i_1}, x_{i_2}, \ldots, x_{i_{K+1}}) \tag{5}$$

Moving in the landscape is done by flipping a single bit of the input string $\mathbf{x}$.

3.2 Properties

The properties of the N-K landscape varies heavily with the choice of K. Adjusting the parameter K tunes the ruggedness, i.e. the number of local optima. The computational complexity of the problem is mostly affected by the choice of dependency model for the component functions. Surprisingly the effect on on the statistical properties of the landscape is not as dramatic. Several results are both available on the complexity of the decision problem of finding the optimum and other analytical results on describing the landscape.

For different values of K the landscape has different statistical properties. Altenberger [1] includes a nice summary of known results.

For the case K=0, each component function is independent of the other functions. Consequently, the flipping of one bit either brings you closer to the optimum or farther away. A simple hill climbing will always find the optimum. If the search is started from a random location the expected number

of steps to the global optimum is $N/2$. Naturally the problem is in **P**.

In case when $K = N - 1$ each component function is dependent of all of the bits of the string. Thus, this means that the fitness of each bit string is statistically independent of its neighbours and the fitness function is equivalent to the random assignments of fitnesses over the bit string space. The following properties hold.

- The probability that a random bit string is a local optimum is $\frac{1}{N+1}$.
- The expected number of local optima is $\frac{2^N}{N+1}$.
- A hill climbing algorithm is expected to hit a local optimum in $ln(N-1)$ steps.

When $1 \leq K < N - 1$ the component functions depend on some but not all bits in the string.

- When K is small, many of the bits are the same for the highest local optima. This correlation decreases when K is increased.
- The average Hamming distance between local optima is approximately

$$\frac{Nlog_2(K+1)}{2(K+1)}.$$

3.3 Complexity

For different component function models and different values of K the computational complexity varies.

For case when the component functions f_i depend on i and its adjacent neighbours two polynomial time solutions have been presented [11, 12]. Of these two solutions, Weinberger's has better asymptotic complexity. The algorithm is a simple algorithm based on dynamic programming.

Theorem. The N-K optimisation problem with adjacent neighbourhoods is solvable in $O(2^K N)$ steps and is thus in P.

For arbitrary neighbourhoods, i.e. when f_i can depend on any bits, the problem is difficult for $K \geq 1$. The problem can be reduced to the NP-complete problem MAX2SAT, where a boolean formula in conjunctive normal form

(CNF) is given and the clauses can contain at most two literals. The objective is to maximise the number of true clauses.

Wright et al. [12] give the following reduction. N is chosen to be the maximum of the number variables and number of clauses. Each clause is allocated a term f_i and each variable in the CNF formula a position in the string. The component functions f_i encode the boolean function their corresponding clause express. Let 1 correspond to true and 0 correspond to false. The value of f will be the number of true clauses in the CNF formula. Thus, we have established the following.

Theorem. The N-K optimisation problem with arbitrary neighbourhoods is NP-complete for $K \geq 1$.

Clearly, in the case of random neighbourhoods when $K \geq 2$, MAXSAT can again be encoded easily because each component function has at least two free variable dependencies which can be used to encode the clauses [12].

Theorem. The N-K optimisation problem with random neighbourhoods is NP-complete for $K \geq 2$.

Weinberger [11] independently proved the same for $K \geq 3$ with a reduction to 3SAT. When $K = 1$ the problem again becomes polynomial [12].

Theorem. The N-K optimisation problem with random neighbourhoods is polynomial $K = 1$.

Wright et al. [12] have investigated how well the optimisation problem can be approximated. The arbitrary N-K optimisation problem is similar to the MAXGSAT problem. In the MAXGSAT problem one is given a conjunction Boolean expressions which can be in any form. The objective is to maximise the number expressions.

In the paper an algorithm is developed which for each input bit maximises the average value of the functions. The result is an algorithm which can be seen in Figure 1. A BDD implementation of algorithm could possibly fair well as the algorithm focuses on manipulating sets of binary strings

Theorem. The approximation threshold for the algorithm with $K \geq 2$ is at most $1 - \frac{1}{2^{K+1}}$.

The upper bound follows from simple reasoning concerning the relation of the average value of f to the optimum. However, as the problem is so similar to the MAXGSAT problem, it is unlikely that a significantly better approximation threshold will be found.

```
function N-K-OPTIM () : (bit string) =
    // s is the approximately optimal string
    // S is initialised to the set of all N-bit strings
    // and evolves as the bits of S are assigned;
    // eventually, S = {s}.
        for i from 0 to N − 1 do
            S_0 ← subset of S where the i:th bit is 0
            M_0 ← average of f over S_0
            S_1 ← subset of S where the i:th bit is 1
            M_1 ← average of f over S_1
            if M_0 > M_1 then
                s[i] ← 0
                S ← S_0
            else
                s[i] ← 1
                S ← S_1
        od
        return s //Return the approximate string
```

Figure 1: Approximating the N-K optimisation problem.

4 Optimisation Problems

Optimisation problems are very common in computer science. A university level basic algorithms course will consider many optimisation problems and their solutions. It will soon be clear for the eager student that while some problems easy to solve, for others an efficient solution has eluded researchers for decades.

In the following we will discuss two well-known optimisation problems: the travelling salesman problem (TSP) and ASSIGNMENT. The problems are in many ways similar but they have crucial difference. TSP is NP-complete while ASSIGNMENT is in P.

4.1 Travelling Salesman

Formally an instance of the travelling salesman problem is an $n \times n$ matrix (a_{ij}), where each element is greater than zero. In most cases the matrix is considered to be symmetric but not in all. The problem is to find the permutation π of $\{1, \ldots, N\}$ such that $\sum_{i=1}^{n} a_{\pi(i),\pi(i+1)}$ (where by $\pi(n+1)$

we mean $\pi(1)$) is minimised.

TSP is one of the most studied NP-complete problems. It belongs to the category problems for which no polynomial approximation scheme is possible unless $P = NP$ [7].

TSP was one the first combinatorial optimisation problems to which the probabilistic methods developed for spin glasses were applied. Some rather interesting results have been achieved.

When the links a_{ij} are considered as uniformly distributed random variables on $[0, 1]$ the length of the tour, with probability one, in the large limit is $l = 2.08$. Consult e.g. [6] for more details on the calculations.

The landscape of the TSP has also undergone intensive study. Stadler and Schnabl [10] have studied statistical properties of the TSP landscape for both the symmetric and the asymmetric case. Two different neighbourhoods are studied: transpositions and inversions. Transpositions just exchanges two cities in the tour. The inversion $[s, t]$ not only exchanges the two cities but also reverts the paths from s to t.

The study used random walks and other statistical analysis to examine the landscapes. Surprisingly, the landscape of TSP is fairly regular. It is a model of a so called AR(1) process for both neighbourhoods. This means that most measure in the landscape can be very well approximated with a Gaussian distribution. For asymmetric TSP the result is almost the same. When transposition is used the result is the same. With inversions the result is more complicated. The landscape is an AR(1) landscape with a superimposed random behaviour.

4.2 Assignment

The assignment problem is quite similar to TSP with the important difference that it is in P. Formally an instance of ASSIGNMENT is a $n \times n$ matrix a_{ij} where $a_{ij} \geq 0$. The problem is to find a permutation π on $\{1, 2, \ldots, n\}$ such that

$$E_n^* = \sum_{i=1}^{n} ai, \pi(i) \tag{6}$$

is minimised.

As can be seen from above the formulation is very similar to TSP. The problem has also been analysed with probabilistic methods borrowed from

statistical mechanics. If the a_{ij} are drawn from a common probability density $\rho(a)$ it has been proved that

$$lim_{n\to\infty}\langle E_n^*\rangle = \frac{\pi^2}{6}. \tag{7}$$

When the a_{ij} are drawn from an exponential distribution it has been shown that in the finite case the average optimum is

$$\langle E_n^*\rangle = \sum_k = 1^n \frac{1}{k^2}. \tag{8}$$

Surprisingly, the correlation of the landscape is the same as for TSP [9]. Thus, simple statistical properties cannot completely characterise the hardness of a problem.

References

[1] Lee Altenberger. NK fitness landscapes. In T. Bäck, D. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*. Oxford University Press, Oxford, 1997.

[2] Francisco Barahona. On the computational complexity of Ising spin glass models. *Journal Physic A*, 15:3241–3253, 1982.

[3] Sorin Istrail. Statistical mechanics, three-dimensionality and np-completeness - I. universality of intractability for the partition function of the Ising model across non-planar lattices (extended abstract). In *Proceeding of the 31st ACM Annual Symposium on the Theory of Computing (STOC 2000)*, pages 87 –96. ACM Press, 2000.

[4] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 22:1087–1116, 1993.

[5] Stuart Kauffman. *The Origins of Order*. Oxford University Press, New York, 1993.

[6] Marc Mezard, Giorgio Parisi, and Miguel Angel Virasoro. *Spin Glass Theory and Beyond*. World Scientific, Singapore, 1987.

[7] Christos H. Papadimitriou. *Computational Complexity*. Addison-Wesley, Reading, Massachusetts, 1994.

[8] Christian M. Reidys and Peter F. Stadler. Combinatorial landscapes. Unpublished research report.

[9] Peter Stadler. Correlation in landscapes of combinatorial optimization problems. *Europhysics Letters*, 20:479–482, 1992.

[10] Peter Stadler and Wolfgang Schnabl. The landscape of the travelling salesman problem. *Physics Letters A*, 161:337–344, 1992.

[11] Edward D. Weinberger. NP completeness of Kauffman's N-k model. Technical Report 96-02-003, Santa Fe Institute, Santa Fe, USA, 1996.

[12] A. H. Wright, R. K. Thompson, and J. Zhang. The computational complexity of N-K fitness functions. *IEEE-EC*, 4(4):373, November 2000.