Landscapes and Other Art Forms.

Darrell Whitley
Computer Science, Colorado State University

Introduction to Elementary Landscapes

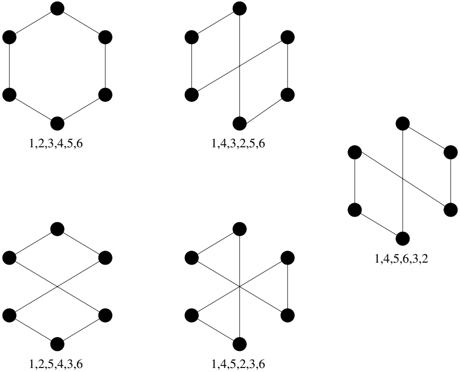
What is a landscape?

- Many different intuitive definitions
- A mathematical formalism of the *search space* of a combinatorial optimization problem

Definition: a landscape is a tuple (X, N, f)

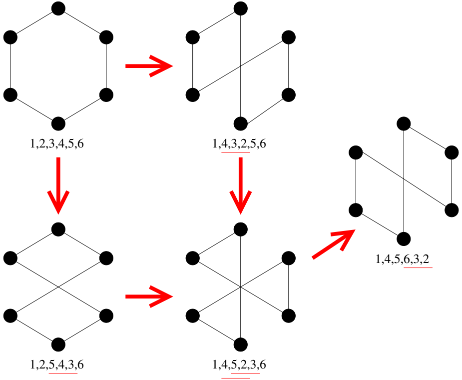
A set of *states*	X
A *neighborhood* operator	$N : X \mapsto \mathcal{P}(X)$
A *fitness* function	$f : X \mapsto \mathbb{R}$

Introduction: a landscape



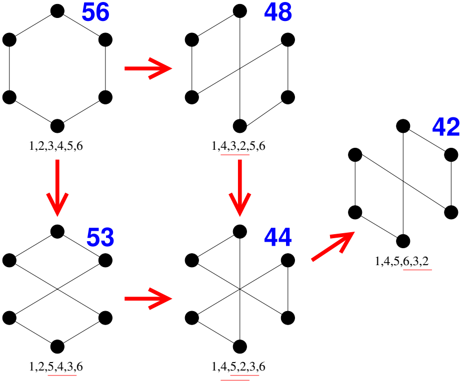
X set of *states*

Introduction: a landscape



X	set of *states*
$N : X \mapsto \mathcal{P}(X)$	neighborhood operator

Introduction: a landscape



X	set of *states*
$N : X \mapsto \mathcal{P}(X)$	neighborhood operator
$f : X \mapsto \mathbb{R}$	objective function

Preliminaries

$$G(X, E)$$

is the underlying graph induced by N.
We assume G is regular with vertices of degree d.

$$\mathbf{A} \in \mathbb{R}^{|X| \times |X|}$$

is the *adjacency matrix* of G.
If x_1 and x_2 are neighbors, $A(x_1, x_2) = 1$.

$$\mathbf{\Delta} = \mathbf{A} - d\mathbf{I}$$

is the Laplacian of G.

The Wave Equation: definition 1

On an arbitrary landscape

- f and N are *unrelated*

On an elementary landscape

The wave equation

$$\mathbf{\Delta} f = \lambda f$$

- where λ is a scalar
- In other words, f is an eigenvector of the Laplacian

The Wave Equation: definition 1

Average change

$$\mathbf{\Delta} f = (\mathbf{A} - d\mathbf{I})f = k(\bar{f} - f)$$

$$\mathbf{\Delta} f(x) = \sum_{y \in N(x)} (f(y) - f(x)) = k(\bar{f} - f(x))$$

Average value

$$\begin{aligned} \operatorname*{avg}_{y \in N(x)} \{f(y)\} &= \frac{1}{d} \sum_{y \in N(x)} f(y) \\ &= f(x) + \frac{1}{d} \left(\sum_{y \in N(x)} f(y) - f(x) \right) \\ &= f(x) + \frac{1}{d} \Delta f(x) \\ &= f(x) + \frac{k}{d} \left(\bar{f} - f(x) \right) \end{aligned}$$

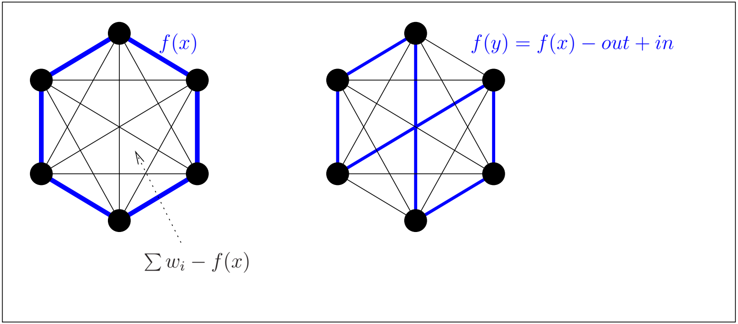
The Wave Equation: definition 2

$$f(x) = \sum \text{ a subset of ``components''}$$

Starting from average...

$$\operatorname*{avg}_{y \in N(x)} \{f(y)\} = f(x) + \operatorname*{avg}_{y \in N(x)} \{\text{components in} - \text{components out}\}$$

Example: TSP under 2-opt



- Components: set of edge weights $w_{i,j}$
- $f(x) =$ sum of edge weights induced by tour x
- There are $n(n-1)/2 - n$ weights not in tour x
- Average value of components out: $\frac{2}{n} f(x)$
- Average value of components in: $\frac{2}{n(n-3)/2} (\sum w - f(x))$

The Components and $\bar{f}$

Let C denote the set of components

$0 < p_3 < 1$ is the proportion of the components in C that contribute to the cost function for any randomly chosen solution

$$\bar{f} = p_3 \sum_{c \in C} c$$

For the TSP:

$$\bar{f} = \frac{n}{n(n-1)/2} \sum_{w_{i,j} \in C} w_{i,j}$$

$$\bar{f} = \frac{2}{n-1} \sum_{w_{i,j} \in C} w_{i,j}$$

The Components and $\bar{f}$

C is the set of components (e.g. from a cost matrix)

x is a solution (e.g. a subset of the cost matrix)

Let $(C - x)$ denote the set of components, excluding those in x

For a move, we then define:
$0 < p_1 < 1$ is the proportion of components in x that change
$0 < p_2 < 1$ is the proportion of components in $(C - x)$ that change

The Components and $\bar{f}$

Theorem

If p_1, p_2 and p_3 can be defined for any regular landscape such that the evaluation function can be decomposed into components where $p_1 = \alpha/d$ and $p_2 = \beta/d$ and

$$\bar{f} = p_3 \sum_{c \in C} c = \frac{\beta}{\alpha + \beta} \sum_{c \in C} c$$

then the landscape is elementary.

The Wave Equation: definition 2

$$
\begin{aligned}
\operatorname*{avg}_{y \in N(x)} \{f(y)\} &= f(x) + \frac{2}{n(n-3)/2}\left(\sum w - f(x)\right) - \frac{2}{n} f(x) \\
&= f(x) + \frac{2}{n(n-3)/2}\left((n-1)/2\bar{f} - f(x)\right) - \frac{2}{n} f(x) \\
&= f(x) + \frac{(n-1)}{n(n-3)/2}(\bar{f} - f(x)) \\
&= f(x) + \frac{k}{d}(\bar{f} - f(x))
\end{aligned}
$$

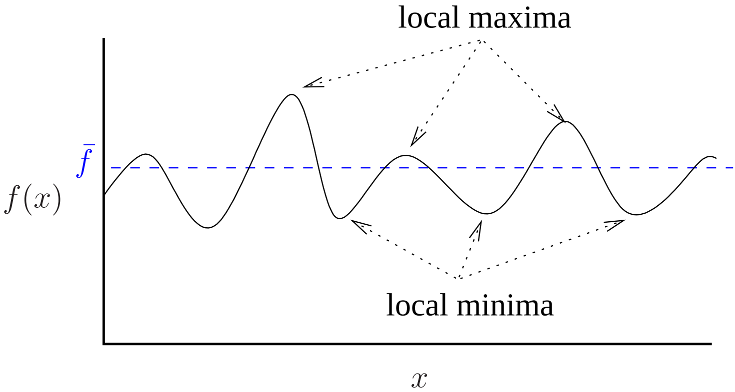
For a 5 city TSP

	ab	bc	cd	de	ae	ac	ad	bd	be	ce
ABCDE	1	1	1	1	1	0	0	0	0	0
ABEDC	1	0	1	1	0	1	0	0	1	0
ABCED	1	1	0	1	0	0	1	0	0	1
ABDCE	1	0	1	0	1	0	0	1	0	1
ACBDE	0	1	0	1	1	1	0	1	0	0
ADCBE	0	1	1	0	1	0	1	0	1	0

Looking at the neighbors in aggregate.

ab	bc	cd	de	ae	ac	ad	bd	be	ce
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	0	0	0	0	0
0	0	0	0	0	1	1	1	1	1
0	0	0	0	0	1	1	1	1	1

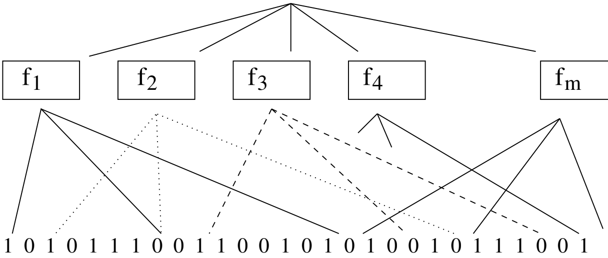
Properties



A General Model for all bounded Pseudo-Boolean Problems

This obviously applies to MAXSAT and NK Landscapes

$$f(x) = \sum_{i=1}^{m} f_i(x, mask)$$



Superpositions of Elementary Landscapes

$$f(x) = f1(x) + f2(x) + f3(x) + f4(x)$$

$$f1(x) = f1_a(x) + f1_b(x) + f1_c(x)$$
$$f2(x) = f2_a(x) + f2_b(x) + f2_c(x)$$
$$f3(x) = f3_a(x) + f3_b(x) + f3_c(x)$$
$$f4(x) = f4_a(x) + f4_b(x) + f4_c(x)$$

$$\varphi^{(1)}(x) = f1_a(x) + f2_a(x) + f3_a(x) + f4_a(x)$$
$$\varphi^{(2)}(x) = f1_b(x) + f2_b(x) + f3_b(x) + f4_b(x)$$
$$\varphi^{(3)}(x) = f1_c(x) + f2_c(x) + f3_c(x) + f4_c(x)$$

$$f(x) = \varphi^{(1)}(x) + \varphi^{(2)}(x) + \varphi^{(3)}(x)$$

MAX-3SAT decomposition

MAX-3SAT is a superposition of 3 elementary landscapes

Walsh span of order p

$$\varphi^{(p)} = \sum_{\{i \,:\, bc(i)=p\}} w_i \psi_i$$

The p^{th} Walsh span is an elementary landscape

$$\Delta\varphi^{(p)} = -2p\varphi^{(p)}$$

With Thanks to Andrew Sutton!

MAX-3SAT decomposition

Recall that we can express f as:

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{2^k} w_{m(i,j)} \psi_{m(i,j)}(x)$$

Grouping the Walsh decomposition results in

$$f(x) = \sum_{p=0}^{3} \varphi^{(p)}(x)$$

Thus MAX-3SAT is a superposition of 3-elementary landscapes

This allows us to compute Statistical Summaries

Properties

Average neighborhood

$$\begin{aligned}\sum_{y \in N(x)} f(y) &= f(x) + \frac{1}{d} \sum_{y \in N(x)} (f(y) - f(x)) \\ &= f + \frac{1}{d} \boldsymbol{\Delta} f \\ &= f + \frac{1}{d} \boldsymbol{\Delta} \left(\sum_{p=0}^{3} \varphi^{(p)} \right) \\ &= f - \frac{1}{d} \left(\sum_{p=0}^{3} 2p\varphi^{(p)} \right)\end{aligned}$$

For 3-SAT

$$f(x) - \frac{1}{d} \left(2\varphi^{(1)}(x) + 4\varphi^{(2)}(x) + 6\varphi^{(3)}(x) \right)$$

This allows us to compute Statistical Summaries

We can compute statistics efficiently over Hyperplanes

We can compute statistics efficiently over Hamming Balls

Random Walks and Autocorrelation

Remark. We have the following identities:

$\langle f, f \rangle = \sum_i w_i^2 \qquad \langle f, \mathbf{T}^s f \rangle = \sum_i \lambda_i^s w_i^2 \qquad \langle \mathbf{1}, f \rangle = w_0$

Random walk process estimates the following equation

$$r(s) = \frac{\langle f, \mathbf{T}^s f \rangle - \langle \mathbf{1}, f \rangle^2}{\langle f, f \rangle - \langle \mathbf{1}, f \rangle^2}$$

Substitutions...

$$r(s) = \frac{\sum\limits_i \lambda_i^s w_i^2 - w_0^2}{\sum\limits_j w_j^2 - w_0^2} = \frac{\sum\limits_{i \neq 0} \lambda_i^s w_i^2}{\sum\limits_{j \neq 0} w_j^2}$$

Random Walks and Autocorrelation

This gives *exact autocorrelation* function

$$r(s) = \frac{\sum_{i \neq 0} \lambda_i^s w_i^2}{\sum_{j \neq 0} w_j^2}$$

where $\lambda_i = \left(1 - \frac{2\langle i,i \rangle}{n}\right)$.

Recall for k-SAT all nonzero w_i can be computed in $O(m)$ time.

THANK YOU

Take Home Message:

PROBLEM STRUCTURE MATTERS.

Black Box Optimizers can never match the performance of an algorithm that efficiently exploits problem structure.

But we need only a small amount of information: **Gray Box Optimization**.