A Polynomial Time Computation of the Exact Correlation Structure of k-Satisfiability Landscapes

Andrew Sutton
Darrell Whitley
Adele Howe

10 July 2009

k-SAT (MAX-k-SAT)

Set of n Boolean variables

$$\{v_1, v_2, \dots, v_n\}$$

Set of m clauses

$$(l_{1,1} \vee l_{1,2}, \vee \dots \vee l_{1,k}), (l_{2,1} \vee l_{2,2}, \vee \dots \vee l_{2,k}), \dots, (l_{m,1} \vee l_{m,2}, \vee \dots \vee l_{m,k})$$

Each consists of exactly k unique *literals* in disjunction

Objective: find an assignment to the set of variables that maximizes the number of clauses that evaluate to true

Local methods

Search the space of complete assignments to all n variables

Perform *local perturbations* in Hamming space

- "Mutations"
- Local search algorithms among state of the art (WalkSat)

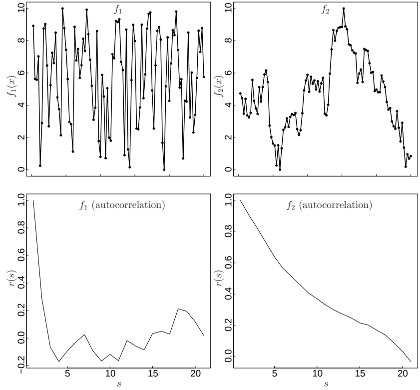
Formalizing the search space

X	set of all 2^n assignments to variables
$f : X \to \mathbb{R}$	fitness function $f(x)$ counts clauses satisfied under x
$N : X \to 2^X$	$N(x) \subset X$ set of Hamming neighbors of x

Empirical correlation

How difficult is a landscape to search locally?

- Correlation structure: "smoothness/ruggedness"
- $\{x_t, x_{t+1}, \ldots, x_{t+s}\}$, $x_{i+1} \in N(x_i)$
- $\{f(x_t), f(x_{t+1}), \ldots, f(x_{t+s})\}$
- $r(s)$: time series *autocorrelation*
- Measured *empirically*



Analytical correlation

Vector/function duality

- Any function $g : X \to \mathbb{R}$ can be characterized as a vector $g \in \mathbb{R}^{|X|}$
- Any vector in $\mathbb{R}^{|X|}$ can be characterized as a function on X.
- Any $|X| \times |X|$ matrix can be used as an operator g.

For example, suppose

$$\mathbf{A} \in \mathbb{R}^{|X|\times|X|}, \quad g, h \in \mathbb{R}^{|X|}$$

Then we have

$$\mathbf{A}g : X \to \mathbb{R}, \quad \mathbf{A}g(x)$$

Furthermore, we can define the *inner product*

$$\langle g, h \rangle = \sum_{x \in X} g(x)h(x)$$

Analytical correlation

$\mathbf{T}$ is *random-walk transition matrix:*

$$\mathbf{T}_{xy} = \begin{cases} \frac{1}{|N(x)|} & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

Analytical correlation

$\mathbf{T}$ is *random-walk transition matrix:*

$$\mathbf{T}_{xy} = \begin{cases} \frac{1}{|N(x)|} & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

The matrix-vector product $\mathbf{T}f$ characterized as function

$$\mathbf{T}f(x) = \frac{1}{|N(x)|} \sum_{y \in N(x)} f(y)$$

where $N(x)$ is the neighborhood of x.

Analytical correlation

$\mathbf{T}$ is *random-walk transition matrix:*

$$\mathbf{T}_{xy} = \begin{cases} \frac{1}{|N(x)|} & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

The matrix-vector product $\mathbf{T}f$ characterized as function

$$\mathbf{T}f(x) = \frac{1}{|N(x)|} \sum_{y \in N(x)} f(y)$$

where $N(x)$ is the neighborhood of x. In general,

$$\mathbf{T}^s f(x) = \frac{1}{|N^s(x)|} \sum_{y \in N^s(x)} f(y)$$

where $N^s(x)$ is the set of states reachable in s steps from x.

Analytical correlation

Random walk process estimates the following equation

$$r(s) = \frac{E[f(x_t)f(x_{t+s})] - E[f(x_t)]^2}{E[f(x_t)^2] - E[f(x_t)]^2}$$

Analytical correlation

Random walk process estimates the following equation

$$r(s) = \frac{E[f(x_t)f(x_{t+s})] - E[f(x_t)]^2}{E[f(x_t)^2] - E[f(x_t)]^2}$$

Expectation	Over all $x \in X$	Inner product

Analytical correlation

Random walk process estimates the following equation

$$r(s) = \frac{E[f(x_t)f(x_{t+s})] - E[f(x_t)]^2}{E[f(x_t)^2] - E[f(x_t)]^2}$$

Expectation	Over all $x \in X$	Inner product
$E[f(x_t)^2]$	$\frac{1}{\lvert X\rvert}\sum_{x\in X} f(x)^2$	$\frac{1}{\lvert X\rvert}\langle f, f\rangle$

Analytical correlation

Random walk process estimates the following equation

$$r(s) = \frac{E[f(x_t)f(x_{t+s})] - E[f(x_t)]^2}{E[f(x_t)^2] - E[f(x_t)]^2}$$

Expectation	**Over all $x \in X$**	**Inner product**
$E[f(x_t)^2]$	$\frac{1}{\|X\|} \sum_{x \in X} f(x)^2$	$\frac{1}{\|X\|}\langle f, f \rangle$
$E[f(x_t)]^2$	$\left(\frac{1}{\|X\|} \sum_{x \in X} f(x) \right)^2$	$\frac{1}{\|X\|}\langle \mathbf{1}, f \rangle^2$

Analytical correlation

Random walk process estimates the following equation

$$r(s) = \frac{E[f(x_t)f(x_{t+s})] - E[f(x_t)]^2}{E[f(x_t)^2] - E[f(x_t)]^2}$$

Expectation	**Over all $x \in X$**	**Inner product**
$E[f(x_t)^2]$	$\frac{1}{\lvert X\rvert}\sum_{x\in X} f(x)^2$	$\frac{1}{\lvert X\rvert}\langle f, f\rangle$
$E[f(x_t)]^2$	$\left(\frac{1}{\lvert X\rvert}\sum_{x\in X} f(x)\right)^2$	$\frac{1}{\lvert X\rvert}\langle \mathbf{1}, f\rangle^2$
$E[f(x_t)f(x_{t+s})]$	$\frac{1}{\lvert X\rvert}\sum_{x\in X} f(x)\left[\frac{1}{\lvert N^s(x)\rvert}\sum_{y\in N^s(x)} f(y)\right]$	$\frac{1}{\lvert X\rvert}\langle f, \mathbf{T}^s f\rangle$

Analytical correlation

Random walk process estimates the following equation

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

Landscape Analysis

Search space is formalized by the landscape structure (X, N, f)

Correlation is a measure of the strength of the relationship between f and N

We will write f "in terms of" N
(actually in terms of $\mathbf{T}$)

Walsh decomposition

Let $x \in \{0,1\}^n$. The i^{th} Walsh function ($0 \le i < 2^n$) is defined as

$$\psi_i(x) = \frac{1}{\sqrt{2^n}}(-1)^{\langle i,x \rangle}$$

Remark. Any pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ can be written as

$$f(x) = \sum_{i=0}^{2^n-1} w_i \psi_i(x)$$

where w_i is a scalar called the Walsh coefficient

Walsh decomposition

In general there are 2^n unique Walsh coefficients w_i

Rana, Heckendorn, and Whitley (1998):

- For k-SAT if $\langle i, i \rangle > k$, then $w_i = 0$
- Walsh coefficients computed by summing up contributions from each clause
- Each clause contributes to $O(1)$ Walsh coefficients (at most 2^k).

Thus, for k-SAT all nonzero Walsh coefficients w_i can be computed in $O(m)$ time.

Main result

Theorem 1. ψ_i is an eigenvector of the random walk transition matrix $\mathbf{T}$.

$$\mathbf{T}\psi_i = \lambda_i \psi_i$$

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

Theorem 2. $\{\psi_i\}$ form an orthonormal basis

So in the Walsh decomposition

$$f = \sum_i w_i \psi_i$$

we are actually writing the fitness function in terms of the eigenbasis of $\mathbf{T}$

Remark. We have the following identities:

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

Remark. We have the following identities:

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

$$\begin{aligned}\langle f, f\rangle &= \langle \sum_i w_i \psi_i, \sum_j w_j \psi_j\rangle \\ &= \sum_i \sum_j w_i w_j \langle \psi_i, \psi_j\rangle \\ &= \sum_i w_i^2 && \text{by Theorem 2}\end{aligned}$$

Remark. We have the following identities:

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

$$\begin{aligned}
\langle f, \mathbf{T}^s f \rangle &= \langle \sum_i w_i \psi_i, \mathbf{T}^s \sum_j w_j \psi_j \rangle \\
&= \sum_i \sum_j w_i \lambda_j^s w_j \langle \psi_i, \psi_j \rangle && \text{by Theorem 1} \\
&= \sum_i \lambda_i^s w_i^2 && \text{by Theorem 2}
\end{aligned}$$

Remark. We have the following identities:

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

$$\begin{aligned}\langle \mathbf{1}, f \rangle &= \langle \mathbf{1}, \sum_i w_i \psi_i \rangle \\ &= w_0\end{aligned}$$

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...

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

This gives *exact autocorrelation* function

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

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.

Conclusion

- Worst case exact correlation for arbitrary function $O(2^n)$
- Exact correlation for k-SAT (MAX-k-SAT) can be computed in polynomial time
- In the paper we also give an expression for computing the *expected* correlation over randomly generated problems
- Found statistical deviation from expectation in filtered problems (future work)