Fitness Landscapes and Graphs: Multimodularity, Ruggedness And Neutrality

Sébastien Verel
Gabriela Ochoa

DOLPHIN team - INRIA Lille-Nord Europe INRIA
I3S laboratory - University of Nice-Sophia Antipolis / CNRS, France
Automated Scheduling, Optimisation and Planning (ASAP) research group - University of Nottingham, UK
http://www.i3s.unice.fr/∼verel

July, 18 2010

Fitness landscapes : Motivations

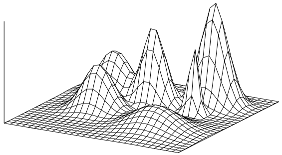
Why using fitness landscapes ?

- To analyse the structure of the search space
- To study problem (search) difficulty in combinatorial optimisation :
  information on runtime for a given problem and a class of LS
- To design effective search algorithms

L. Barnett, U. Sussex, DPhil Diss. 2003

"the more we know of the statistical properties of a class of fitness landscapes, the better equipped we will be for the design of effective search algorithms for such landscapes"

Fitness landscapes in biology

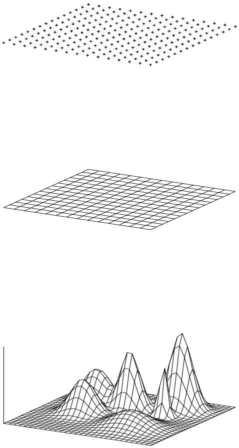


Biological science :
Wright 1930 [35]

Biological evolution :

- a metaphorical uphill struggle across a "fitness landscape"
- mountain peaks represent high "fitness", or ability to survive,
- valleys represent low fitness.
- evolution proceeds : population of organisms performs an "adaptive walk"

Fitness landscapes in biology



In biology :

- Modelisation of species evolution

Used to model dynamical systems :

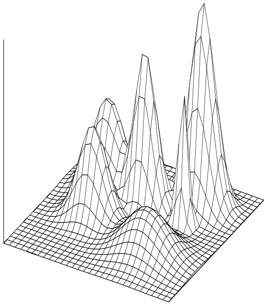
- statistical physic,
- molecular evolution,
- ecology, etc

Fitness landscapes in biology

2 sides for Fitness Landscapes :

- Powerful metaphor : most profound concept in evolutionary dynamics
  - give pictures of evolutionary process
  - be careful of misleading pictures : "smooth landscape without noise"
- Quantitative concept : predict the evolutionary paths
  - Quasispecies equation : mean field analysis with differential equations
  - Stochastic process : markov chain
  - Network analysis

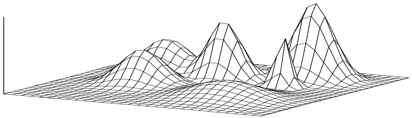
In combinatorial optimization



Fitness landscape $(\mathcal{S}, \mathcal{N}, f)$:

- $\mathcal{S}$: set of admissible solutions,
- $\mathcal{N} : \mathcal{S} \to 2^{\mathcal{S}}$: neighborhood function,
- $f : \mathcal{S} \to \mathbb{R}$: fitness function.

Fitness landscapes for black-box optimisation



Tools for black-box optimisation

Blackbox scenario :
we have only $\{(x_0, f(x_0)), (x_1, f(x_1)), ...\}$ given by an "oracle"

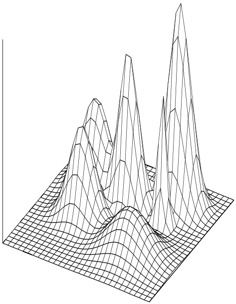
Search space analysis where "no" information is either not available or needed on the definition of fitness function.

Fitness landscapes in evolutionary computation

2 sides for Fitness Landscapes :

- Powerful metaphor : most profound concept
  - give pictures of the search dynamic :
    "if the fitness landscapes have big valleys, I can use this algorithm"
  - be careful of misleading pictures : set of smooth mountains
- Quantitative concept : predict the evolutionary dynamic
  - Quasispecies equation : mean field analysis with differential equations
  - Stochastic process : markov chain
  - Network analysis

What is a neighborhood ?



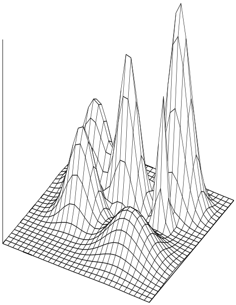
Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

Set of "neighbor" solutions associated to each solution

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op(x)) > 0\}$$

What is a neighborhood ?



Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

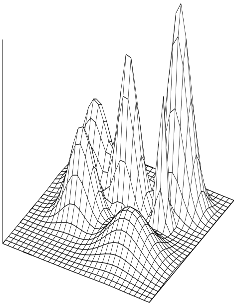
Set of "neighbor" solutions associated to each solution

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op(x)) > 0\}$$

or

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op(x)) > \epsilon\}$$

What is a neighborhood ?



Neighborhood function :

$$\mathcal{N} : \mathcal{S} \rightarrow 2^{\mathcal{S}}$$

Set of "neighbor" solutions associated to each solution

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op(x)) > 0\}$$

or

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op(x)) > \epsilon\}$$

or

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid d(y, x) \leq 1\}$$

Example of neighborhood : bit strings

Search space : $\mathcal{S} = \{0,1\}^N$
Algorithm : simple GA, hill-climbing, or simulated annealing, etc.

$\mathcal{N}(01101) = \{$
01101,
01100,
01111,
01001,
00101,
11101,
$\}$

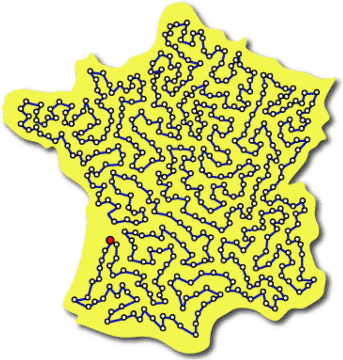
Important !

Definition of neighborhoood must be based on the local search operator used in the algorithm

Neighborhood $\Leftrightarrow$ Operator

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid d_{Hamming}(y,x) \leq 1\}$$

Example of neighborhood : permutations

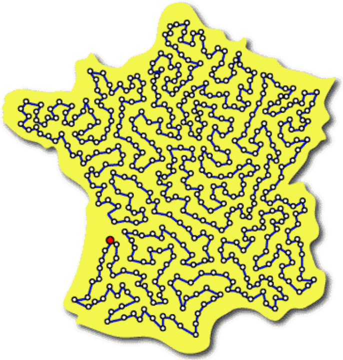


Traveling Salesman Problem : find the shortest tour which cross one time every town

- *Search space* :
  $\mathcal{S} = \{\sigma \mid \sigma \text{ permutations } \}$
- *Algorithm* : simple EA
  operator : 2-opt

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op_{2opt}(x)) > 0\}$$

Example of neighborhood



Traveling Salesman Problem : find the shortest tour which cross one time every town

- *Search space* : $\mathcal{S} = \{\sigma \mid \sigma \text{ permutations } \}$
- *Algorithm* : simple EA
  operators : 2-opt and 3-opt

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid \mathbb{P}(y = op_{2opt}(x)) > 0 \textbf{ or } \mathbb{P}(y = op_{3opt}(x)) > 0\}$$

Example of neighborhood : memetic algorithms

- *Algorithm* : memetic algorithm, EA + operator hill-climbing

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid y = op_{HC}(x)\}$$

Example of neighborhood : memetic algorithms

- *Algorithm* : memetic algorithm, EA + operator hill-climbing

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid y = op_{HC}(x)\}$$

- *Algorithm* : memetic algorithm, EA + operator hill-climbing and bit-flip mutation

2 possibilities :

- Study 2 landscapes :
  one for *HC* operator, one for bit-flip mutation
- Study 1 landscape :
  $\mathcal{N}(x) = \{y \in \mathcal{S} \mid y = op_{HC}(x) \textbf{ or } \mathbb{P}(y = op_{bit-flip}(x)) > \epsilon\}$

Example of neighborhood : memetic algorithms

- *Algorithm* : memetic algorithm, EA + operator hill-climbing

$$\mathcal{N}(x) = \{y \in \mathcal{S} \mid y = op_{HC}(x)\}$$

- *Algorithm* : memetic algorithm, EA + operator hill-climbing and bit-flip mutation

2 possibilities :

- Study 2 landscapes :
  one for *HC* operator, one for bit-flip mutation
- Study 1 landscape :
  $\mathcal{N}(x) = \{y \in \mathcal{S} \mid y = op_{HC}(x)$ **or** $\mathbb{P}(y = op_{bit-flip}(x)) > \epsilon\}$

It depends on what you want to know

Goal of the fitness landscapes study

- "Geometry" (features) of fitness landscape
  $\Rightarrow$ dynamics of a local search algorithm
- Geometry is linked to the problem difficulty :
  - If there are a lot of local optima, the probability to find the global optimum is lower.
  - If the fitness landscape is flat, discovering better solutions is rare.
  - What is the best search direction in the landscape ?

Study of the fitness landscape features
allows to study
the performance of search algorithms

Goal of the fitness landscapes study

1. To compare the difficulty of two search spaces :
   - One problem with 2 (or more) possible codings : $(S_1, \mathcal{N}_1, f_1)$ and $(S_2, \mathcal{N}_2, f_2)$
     different coding, mutation operator, fitness function, etc.

     Which one is easier to solve ?

Goal of the fitness landscapes study

1. To compare the difficulty of two search spaces :
   - One problem with 2 (or more) possible codings : $(S_1, \mathcal{N}_1, f_1)$ and $(S_2, \mathcal{N}_2, f_2)$
     different coding, mutation operator, fitness function, etc.

     Which one is easier to solve ?
2. To choose the algorithm :
   - analysis of global geometry of the landscape

     Which algorithm can I use ?

Goal of the fitness landscapes study

1. To compare the difficulty of two search spaces :
    - One problem with 2 (or more) possible codings : $(S_1, \mathcal{N}_1, f_1)$ and $(S_2, \mathcal{N}_2, f_2)$
      different coding, mutation operator, fitness function, etc.

      Which one is easier to solve ?
2. To choose the algorithm :
    - analysis of global geometry of the landscape

      Which algorithm can I use ?
3. To tune the parameters :
    - *off-line* analysis of structure of fitness landscape

      Which is the best mutation operator ? the size of the population ? etc.

Goal of the fitness landscapes study

1. To compare the difficulty of two search spaces :
   - One problem with 2 (or more) possible codings : $(S_1, \mathcal{N}_1, f_1)$ and $(S_2, \mathcal{N}_2, f_2)$
     different coding, mutation operator, fitness function, etc.

     Which one is easier to solve ?
2. To choose the algorithm :
   - analysis of global geometry of the landscape

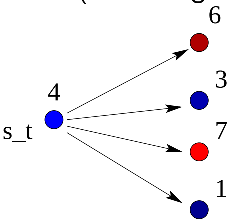
     Which algorithm can I use ?
3. To tune the parameters :
   - *off-line* analysis of structure of fitness landscape

     Which is the best mutation operator ? the size of the population ? etc.
4. To control the parameters during the run :
   - *on-line* analysis of structure of fitness landscape

     Which is the optimal mutation rate according to the estimation of structure ?

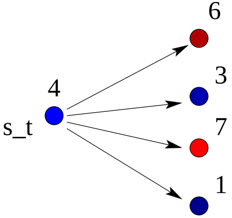
Point of view : Before putting a particular heuristic

FL = (Sol., Neighbors, Fitness)

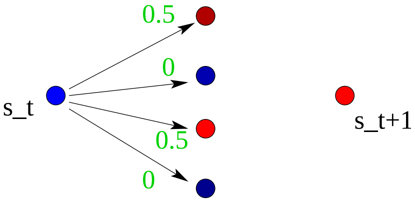


Point of view : Before putting a particular heuristic

FL = (Sol., Neighbors, Fitness)



Put prob. from your heuristic :



- Sample the neighborhood to have information on local features of the search space
- From local information : deduce some global features like general shape of search space, "difficulty", etc.

Goal of the fitness landscapes study

Study of the geometry of the landscape
allows to
study the difficulty, and design a good optimisation algorithm

Fitness landscape is a graph $(\mathcal{S}, \mathcal{N}, f)$ where the nodes have a value (fitness) : can be "pictured" as a "real" landscape

Two main geometries have been studied :

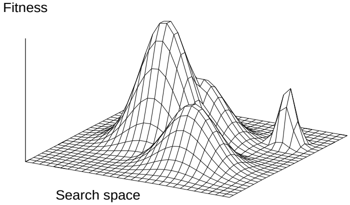
- multimodal and ruggedness
- neutral

Multimodal Fitness landscapes

Local optima s^* :

no neighbor solution with higher fitness value

$$\forall s \in \mathcal{N}(s^*), f(s) < f(s^*)$$



Multimodal Fitness landscapes

Adaptive walk : $(s_0, s_1, \ldots)$ where $s_{i+1} \in \mathcal{N}(s_i)$ and $f(s_i) < f(s_{i+1})$

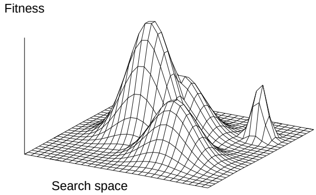
Hill-Climbing (HC) algorithm

Choose initial solution $s \in S$
repeat
  choose $s^{'} \in \mathcal{N}(s)$ such that $f(s^{'}) = max_{x \in \mathcal{N}(s)} f(x)$
  if $f(s) < f(s^{'})$ **then**
    $s \leftarrow s^{'}$
  end if
until s is a Local optimum

Basin of attraction of s^* :

$$\{s \in \mathcal{S} \mid HillClimbing(s) = s^*\}.$$

Multimodal Fitness landscapes



Optimisation difficulty :
number and size of attractive basins (Garnier *et al* [10])

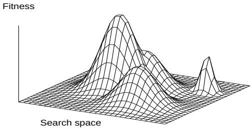
The idea :

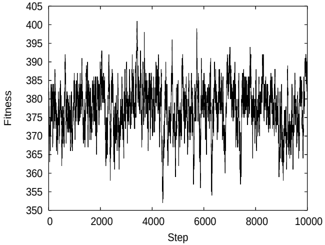
- if the size of attractive basin of global optima is relatively "small"
- the problem is difficult to optimize

The measure :

- Length of adaptive walks (distribution, avg, etc.)

Walking on fitness landscapes



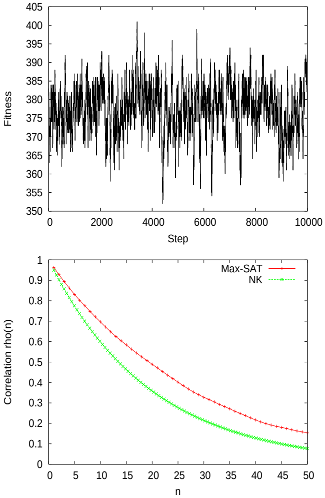


fitness vs. step of a random walk (example of max-SAT problem)

Random walk : $(s_1, s_2, \dots)$ such that $s_{i+1} \in \mathcal{N}(s_i)$ and equiprobability on $\mathcal{N}(s_i)$

- Fitness seems to be very "chaotic"
- Analysis the fitness during the random walk as a signal

Rugged/smooth fitness landscapes



Autocorrelation of time series of fitnesses $(f(s_1), f(s_2), \ldots)$ along a random walk $(s_1, s_2, \ldots)$ [34] :

$$\rho(n) = \frac{E[(f(s_i) - \bar{f})(f(s_{i+n}) - \bar{f})]}{var(f(s_i))}$$

autocorrelation length $\tau = \frac{1}{\rho(1)}$

- small τ : rugged landscape
- long τ : smooth landscape

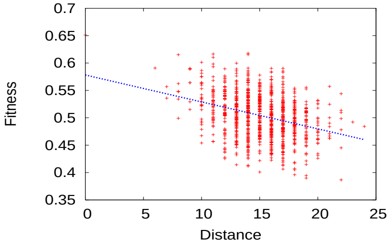
Results on rugged fitness landscapes (Stadler 96 [26])

Problem	parameter	$\rho(1)$
symmetric TSP	n number of towns	$1 - \frac{4}{n}$
anti-symmetric TSP	n number of towns	$1 - \frac{4}{n-1}$
Graph Coloring Problem	n number of nodes α number of colors	$1 - \frac{2\alpha}{(\alpha-1)n}$
NK landscapes	N number of proteins K number of epistasis links	$1 - \frac{K+1}{N}$

Ruggedness decreases with the size of thoses problems :
small variation has less effect on the fitness values

Fitness Distance Correlation (FDC) (Jones 95 [15])

Correlation between distance to global optimum and fitness



Classification based on experimental studies :

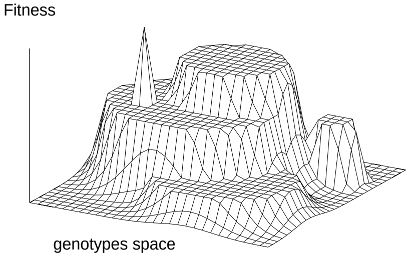
- $\rho < -0.15$, easy optimization
- $\rho > 0.15$, hard optimization
- $-0.15 < \rho < 0.15$, undecided zone

Neutral Fitness Landscapes

Neutral theory (Kimura ≈ 1960 [17])

Theory of mutation and random drift

A considerable number of mutations have no effects on fitness values

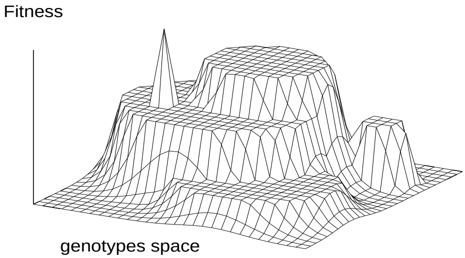


- plateaus
- neutral degree
- neutral networks [Schuster 1994 [25], RNA folding]

Neutral Fitness Landscapes

Combinatorial optimization

- Redundant problem (symetries, ...) (Goldberg 87 [12])
- Problem "not well" defined or dynamic environment (Torres 04 [14])



Applicative problems :

- Robot controler
- Circuit design
- genetic programming
- Protein Folding
- learning problems

Neutrality and difficulty

- In our knowledge, there is no definitive answer about neutrality / problem hardness
- Certainly, it is dependent on the nature of neutrality of the fitness landscape

$\Rightarrow$ Sharp description of the geometry of neutral fitness landscapes is needed

Neutrality and difficulty

We know for certain that :

- No information is better than Bad information :
  Hard trap functions are more difficult than needle-in-a-haystack functions
- Good information is better than No information

Neutrality and difficulty

We know for certain that :

- No information is better than Bad information :
  Hard trap functions are more difficult than needle-in-a-haystack functions
- Good information is better than No information

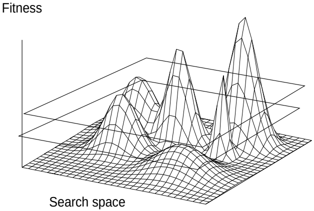
- When there is No information :
  you should have a good method to find it !

In the following

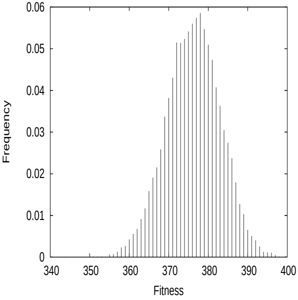
Description of neutral fitness landscapes :

- Neutral sets :
  set of solutions with the same fitness
- Neutral networks :
  add neighborhood information

Neutral sets : Density Of States



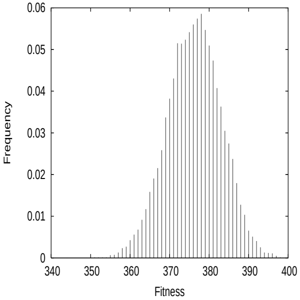
Set of solutions with fitness value



Density of states (D.O.S.)

- Introduce in physics (Rosé 1996 [24])
- Optimization (Belaidouni, Hao 00 [4])

Neutral sets : Density Of States

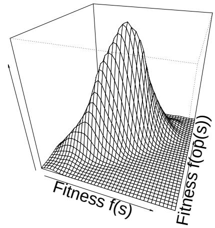


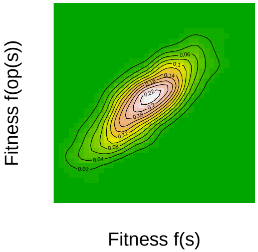
Density of states (D.O.S.)

Informations given :

- Performance of random search
- Tail of the distribution is an indicator of difficulty :
  - the faster the decay, the harder the problem
- But do not care about the neighborhood relation

Neutral sets : Fitness Cloud





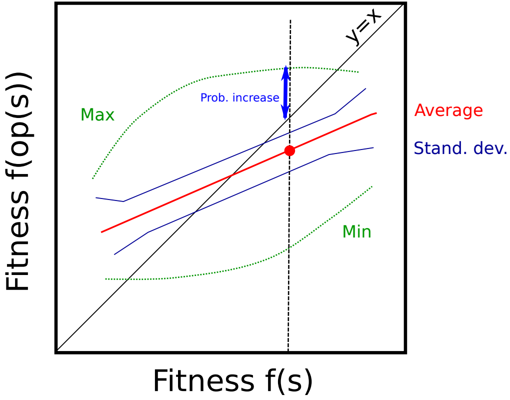
- $(\mathcal{S}, \mathcal{F}, \mathbb{P})$: probability space
- $op : \mathcal{S} \rightarrow \mathcal{S}$ stochastic operator of the local search

- $X(s) = f(s)$
- $Y(s) = f(op(s))$

Fitness Cloud of *op*

Conditional probability density function of Y given X

Fitness cloud : Measure of evolvability

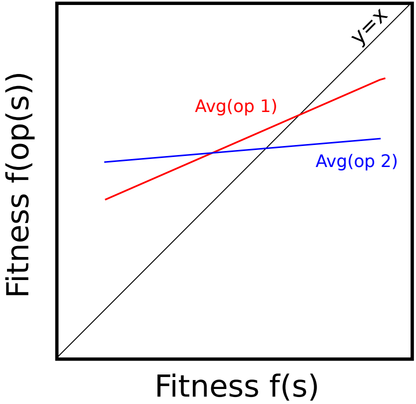


Evolvability

Ability to evolve : fitness in the neighborhood compared to the fitness of the solution

- Probability of finding better solutions
- Average fitness of better neighbor solutions
- Average and standard deviation of fitnesses

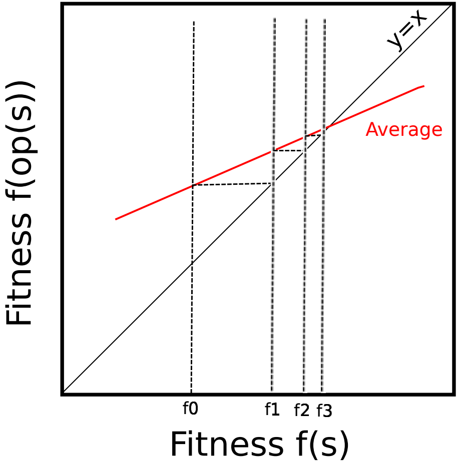
Fitness cloud : Comparaison of difficulty



- Operator 1 > Operator 2
- Because Average 1 more correlated to fitness
- Linked to autocorrelation
- Average is often a line :
  - See works on Elementary Landscapes (D. Wihtley and others)
  - See Negative Slope Coefficient (NSC)

Fitness cloud

Prediction of fitness (CEC 2003)

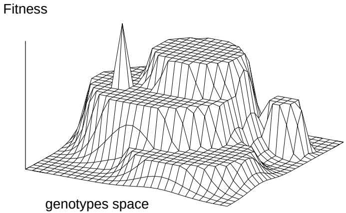


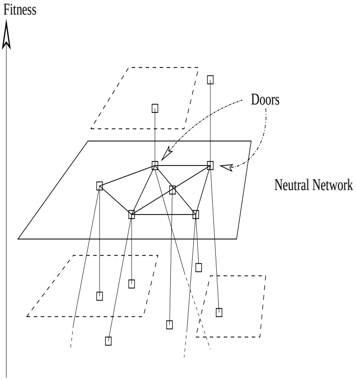
- Approximation (only approximation) of the fitness value after few steps of local operator
- Indication on the quality of the operator

Neutral fitness landscapes

- Neutral sets (done) :
  set of solutions with the same fitness
  $\Rightarrow$ No structure
- Fitness cloud (done) :
  Bivariate density $(f(s), f(op(s)))$
  $\Rightarrow$ Neighborhood relation between neutral sets
- Neutral networks (to be done) :
  $\Rightarrow$ Neighborhood structure into the neutral sets : Graph

Neutral networks (Schuster 1994 [25])





Definitions

Test of neutrality

$$isNeutral : S \times S \rightarrow \{true, false\}$$

For example, $isNeutral(s_1, s_2)$ is *true* if :

- $f(s_1) = f(s_2)$.
- $|f(s_1) - f(s_2)| \leq 1/M$ with M is the search population size.
- $|f(s_1) - f(s_2)|$ is under the evaluation error.

Neutral neighborhood

of s is the set of neighbors which have the same fitness $f(s)$

$$\mathcal{N}_{neut}(s) = \{s' \in \mathcal{N}(s) \mid isNeutral(s, s')\}$$

Neutral degree of s

Number of neutral neighbors : $nDeg(s) = \sharp(\mathcal{N}_{neut}(s) - \{s\})$.

Definitions

Neutral walk

$W_{neut} = (s_0, s_1, \ldots, s_m)$

- for all $i \in [0, m-1]$, $s_{i+1} \in \mathcal{N}(s_i)$
- for all $(i, j) \in [0, m]^2$, $isNeutral(s_i, s_j)$ is *true*.

Neutral Network

graph $G = (N, E)$

- $N \subset \mathcal{S}$: for all s and $s^{'}$ from V, there is a neutral walk belonging to V from s to $s^{'}$,
- $(s_1, s_2) \in E$ if they are neutral neighbors : $s_2 \in \mathcal{N}_{neut}(s_1)$

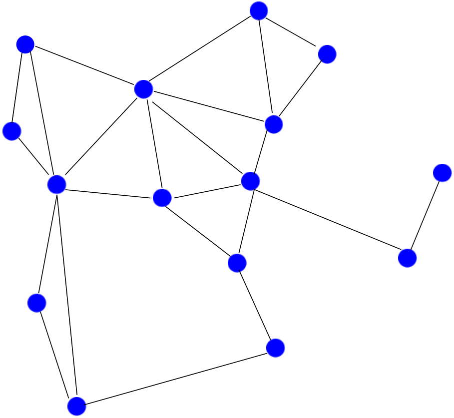
A fitness landscape is neutral
if there are many solutions with high neutral degree.

Neutral Networks (NN) : Inside Metrics

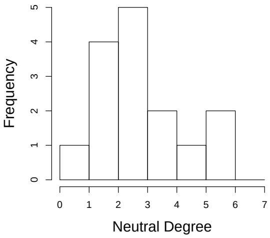
Classical graph metrics :

- Size of NN :
  number of nodes of NN,
- Neutral degree distribution :
  - measure of the quantity of "neutrality"
- Autocorrelation of neutral degree (Bastolla 03 [3]) :
  during neutral random walk
  - comparaison with random graph,
  - measure of the correlation structure of *NN*

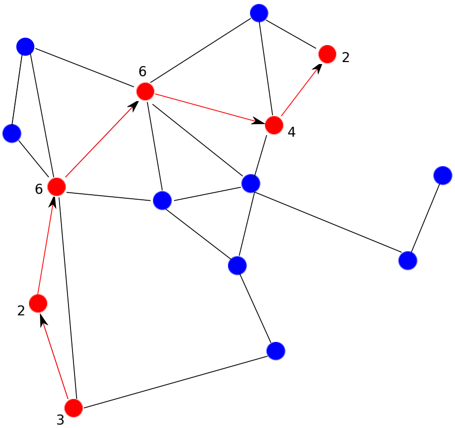
Neutral Networks : Inside Metrics



- Size : 15 solutions
  Distribution of size overall landscapes
- Neutral degree distribution

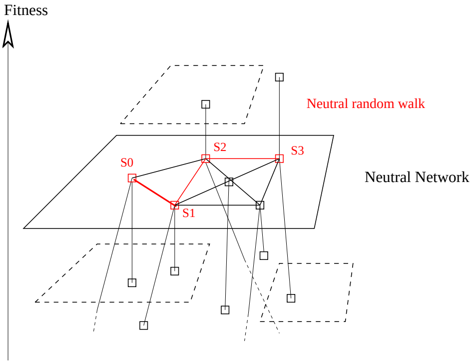


Neutral Networks : Inside Metrics



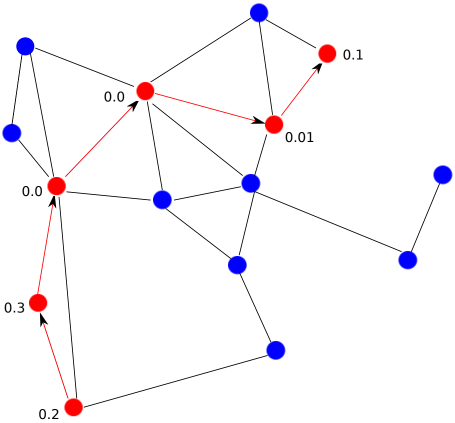
- Size : 15 solutions
  Distribution of size overall landscapes
- Neutral degree distribution
- Autocorrelation of neutral degree :
  - random walk on NN
  - autocorrelation of degrees

Neutral Networks : Outside Metrics



1. Rate of innovation (Huynen 96 [13]) : The number of new accessible structures (fitness) per mutation
2. Autocorrelation of evolvability [32] : autocorrelation of the sequence $(evol(s_0), evol(s_1), \ldots)$.

Neutral Networks : Outside Metrics



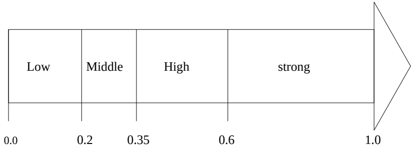
- Autocorrelation of evolvability :
  - Evolvability $evol$ = avg fitness in the neighborhood
  - Autocorrelation of $(evol(s_0), evol(s_1), \ldots)$.
- Informations :
  - if high correlation $\Rightarrow$ "easy" (you can use this information)
  - if low correlation $\Rightarrow$ "difficult"

Summary of metrics

- Neutral degrees distribution :

  "How neutral is the fitness landscape ?"

- Autocorrelation of neutral degrees : network “structure”



- Rate of innovation :

  low information for combinatorial optimization

- Autocorrelation of evolvability :

  information on the links between NN

Basic Methodology of fitness landscapes analysis

- Density of States : pure random search, initialization ?
- Length of adaptive walks : multimodality ?
- Autocorrelation of fitness : ruggedness ?
- Neutral Degree Distribution : neutrality ?
- Fitness Cloud : Quality of the operator, evolvability ?
- Fitness Distance Correlation from best known
- Neutral walks and evolvability : neutral information ?

Basic Methodology of fitness landscapes analysis

- Density of States : pure random search, initialization ?
- Length of adaptive walks : multimodality ?
- Autocorrelation of fitness : ruggedness ?
- Neutral Degree Distribution : neutrality ?
- Fitness Cloud : Quality of the operator, evolvability ?
- Fitness Distance Correlation from best known
- Neutral walks and evolvability : neutral information ?
- ... be creative from your algorithm and problem point of view
- ... be careful on the computed measures : one measure is not enough, and must be very well understand

Sofware to perform fitness landscape analysis

Framework **ParadisEO 1.3**

http://paradiseo.gforge.inria.fr/newWebsite/index.php?n=Doc.Tuto

and tutorials :

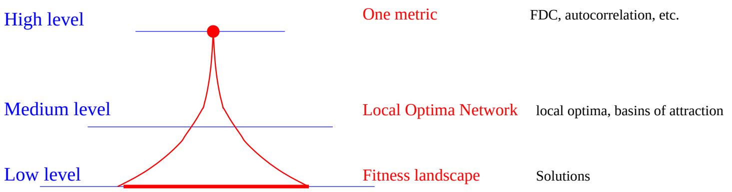
http://paradiseo.gforge.inria.fr/newWebsite/index.php?n=Doc.Tuto

```
moAutocorrelationSampling<Neighbor> sampling(randomInit,
                                             neighborhood,
                                             fullEval,
                                             incrementalEval,
                                             nbStep);

sampling();

sampling.fileExport(str_out);
```

Motivation and general idea : Levels of description



- Fitness landscapes : based on an huge number of solutions
- One metric : based on one real number, or curve to catch all the complexity
- Local optima Network : based on local optima

Overview and Motivation

- Bring the tools of *complex networks* analysis to the study the structure of combinatorial fitness landscapes
- Goals : Understand problem difficulty, design effective heuristic search algorithms
- Methodology : Extract a network that represents the landscape (Inspiration from energy landscapes (Doye, 2002)[1])
  - Vertices : local optima
  - Edges : a notion of adjacency between basins
- Conduct a network analysis
- Relate (exploit ?) network features to search algorithm design

1. J. P. K. Doye, The network topology of a potential energy landscape : a static scale-free network., *Phys. Rev. Lett.*, 88 :238701, 2002.

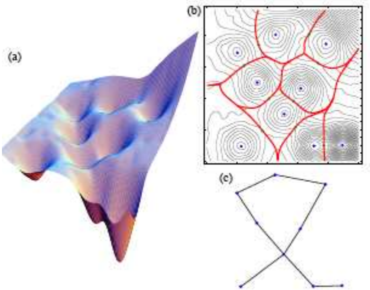
Small – world networks (Watts and Strogatz, 1998)

- Neither ordered nor completely random
- Nodes highly clustered yet path length is small
- Network topological measures :
  - C : clustering coefficient, measure of local density
  - l : shortest path length global measure of separation

Scale – free networks (Barabasi and Albert, 1999)

- The distribution of the number of neighbours (the degree distribution) is *right – skewed* with a heavy tail
- Most of the nodes have less-than-average degree, whilst a small fraction of hubs have a large number of connections
- Described mathematically by a power-law

Energy surface and inherent networks (Doye, 2002)



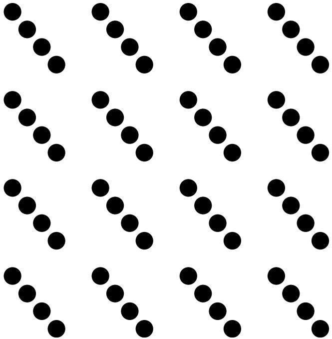
- a Model of 2D energy surface
- b Contour plot, partition of the configuration space into basins of attraction surrounding minima
- c landscape as a network

Inherent network :

- Nodes : energy minima
- Edges : two nodes are connected if the energy barrier separating them is sufficiently low (transition state)

Basins of attraction in combinatorial optimisation

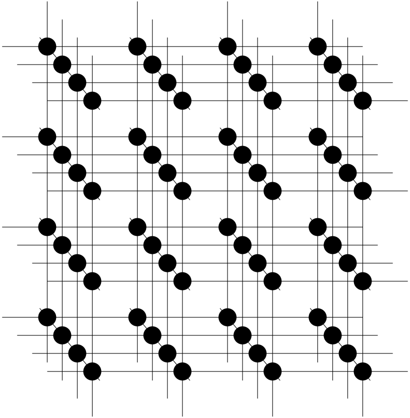
Example of small NK landscape with $N = 6$ and $K = 2$



- Bit strings of length $N = 6$
- $2^6 = 64$ solutions
- one point $=$ one solution

Basins of attraction in combinatorial optimisation

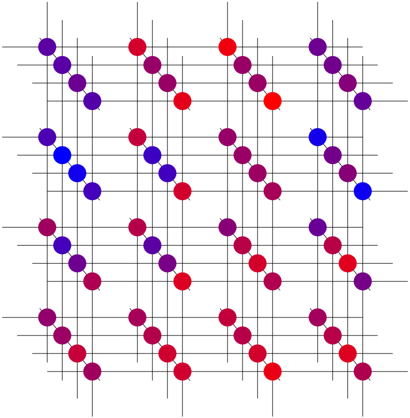
Example of small NK landscape with $N = 6$ and $K = 2$



- Bit strings of length $N = 6$
- Neighborhood size $= 6$
- Line between points $=$ solutions are neighbors
- Hamming distances between solutions are preserved (except for at the border of the cube)

Basins of attraction in combinatorial optimisation

Example of small *NK* landscape with $N = 6$ and $K = 2$

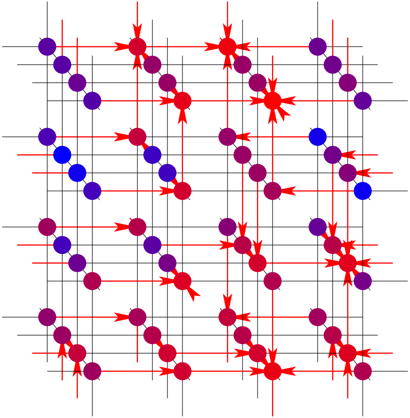


Color represent fitness value

- high fitness
- low fitness

Basins of attraction in combinatorial optimisation

Example of small *NK* landscape with $N = 6$ and $K = 2$



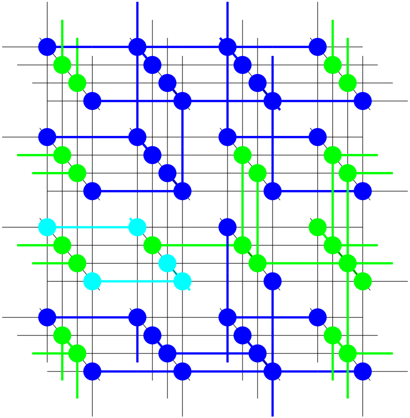
- Color represent fitness value
  - high fitness
  - low fitness
- ⟶ point towards the solution with highest fitness in the neighborhood

Exercise :

Why not make a Hill-Climbing walk on it ?

Basins of attraction in combinatorial optimisation

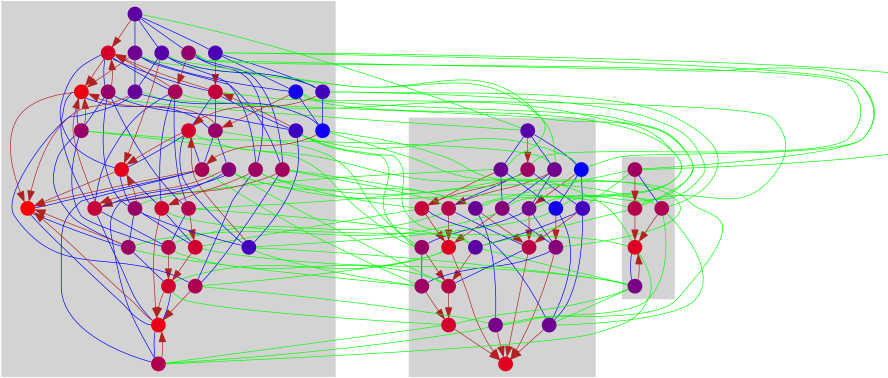
Example of small NK landscape with $N = 6$ and $K = 2$



- Each color corresponds to one basin of attraction
- Basins of attraction are interlinked and overlapped
- Basins have no "interior"

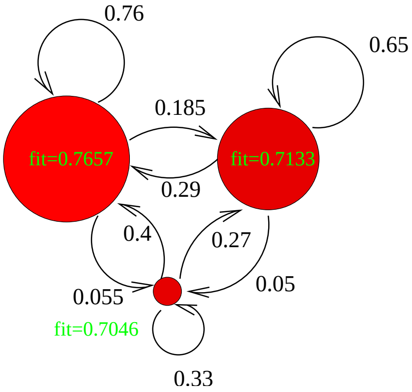
Basins of attraction in combinatorial optimisation

Example of small NK landscape with $N = 6$ and $K = 2$



- Basins of attraction are interlinked and overlapped !
- Most neighbours of a given solution are outside its basin

Local optima network



- Nodes : local optima
- Edges : transition probabilities

Basin of attraction

Hill-Climbing (HC) algorithm

Choose initial solution $s \in S$
repeat
  choose $s^{'} \in \mathcal{N}(s)$ such that $f(s^{'}) = max_{x \in \mathcal{N}(s)} \; f(x)$
  if $f(s) < f(s^{'})$ **then**
    $s \leftarrow s^{'}$
  end if
until s is a Local optimum

Basin of attraction of s^* :

$$\{s \in \mathcal{S} \mid HillClimbing(s) = s^*\}.$$

local optima network

Local optima network

- Nodes : set of local optima $\mathcal{S}^*$
- Edges : notion of connectivity between basins of attraction
  - e_{ij} between i and j if there is at least a pair of neighbours s_i and $s_j \in \mathcal{N}(s_i)$ such that $s_i \in b_i$ and $s_j \in b_j$ (GECCO 2008 [21])
  - weights w_{ij} is attached to the edges, account for transition probabilities between basins (ALIFE 2008 [33], Phys. Rev. E 2008 [30], CEC 2010)

Weights of edges

- From each s and $s^{'}$, $p(s \rightarrow s^{'}) = \mathbb{P}(s^{'} = op(s))$

  For example, $\mathcal{S} = \{0,1\}^N$ and bit-flip operator

  - if $s^{'} \in \mathcal{N}(s)$, $p(s \rightarrow s^{'}) = \frac{1}{N}$
  - if $s^{'} \notin \mathcal{N}(s)$, $p(s \rightarrow s^{'}) = 0$

Weights of edges

- From each s and $s^{'}$, $p(s \rightarrow s^{'}) = \mathbb{P}(s^{'} = op(s))$

  For example, $\mathcal{S} = \{0,1\}^N$ and bit-flip operator

  - if $s^{'} \in \mathcal{N}(s)$, $p(s \rightarrow s^{'}) = \frac{1}{N}$
  - if $s^{'} \notin \mathcal{N}(s)$, $p(s \rightarrow s^{'}) = 0$

- Probability that a configuration $s \in S$ has a neighbor in a basin b_j

$$p(s \rightarrow b_j) = \sum_{s^{'} \in b_j} p(s \rightarrow s^{'})$$

Weights of edges

- From each s and s', $p(s \to s') = \mathbb{P}(s' = op(s))$

  For example, $\mathcal{S} = \{0,1\}^N$ and bit-flip operator

  - if $s' \in \mathcal{N}(s)$, $p(s \to s') = \frac{1}{N}$
  - if $s' \notin \mathcal{N}(s)$, $p(s \to s') = 0$
- Probability that a configuration $s \in S$ has a neighbor in a basin b_j

$$p(s \to b_j) = \sum_{s' \in b_j} p(s \to s')$$

- w_{ij} : Total probability of going from basin b_i to basin b_j is the average over all $s \in b_i$ of the transition prob. to $s' \in b_j$:

$$p(b_i \to b_j) = \frac{1}{\sharp b_i} \sum_{s \in b_i} p(s \to b_j)$$

$\Rightarrow$ local optima network : weighted oriented graph

NK fitness landscapes : ruggedness and epistasis

NK-landscapes : Model of problems

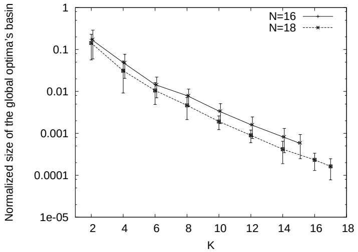
N size of the bit-strings
K from 0 to $N-1$, NK landscapes can be tuned from smooth to rugged (easy to difficult respectively) :

- $K = 0$ no correlations, f is an additive function, and there is a single maximum
- $K = N - 1$ landscape completely random, the expected number of local optima is $\frac{2^N}{N+1}$
- Intermediate values of K interpolate between these two extreme cases and have a variable degree of epistasis (i.e. gene interaction)

Methods

- Extracted and analysed networks
  - $N \in \{14, 16, 18\}$,
  - $K \in \{2, 4, \ldots, N-2, N-1\}$
  - 30 random instances for each case
- Measures :
  - Statistics on basins sizes and fitness of optima
  - Network features : clustering coefficient, shortest path to the global optimum, weight distribution, disparity, boundary of basins

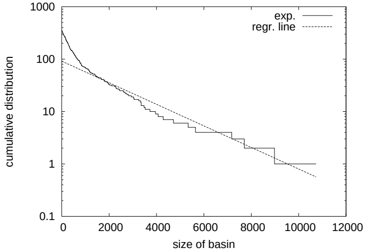
Global optimum basin size versus K



Size of the basin corresponding to the global maximum for each K

- Trend : the basin shrinks very quickly with increasing K.
- for higher K, more difficult for a search algorithm to locate the basin of attraction of the global optimum

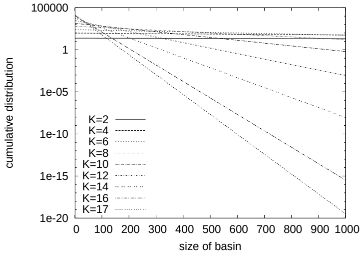
Analysis of basins : basin size



Cumulative distribution of basins sizes for $N = 18$ and $K = 4$

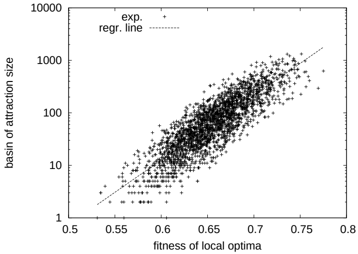
- Trend : small number of large basin, large number of small basin
- Log-normal cumulative distribution : not uniform !
- Slope of correlation increases with K
- When K large : basin sizes are nearly equals the distribution becomes more uniform

Analysis of basins : basin size



- Trend : small number of large basin, large number of small basin
- log-normal cumulative distribution
- slope of correlation increases with K
- when K large : basin sizes are nearly equals

Analysis of basins : fitness *vs.* basin size



Correlation fitness of local optima vs. their corresponding basins sizes

- Trend : clear positive correlation between the fitness values of maxima and their basins' sizes

The highest, the largest

- On average, the global optimum easier to find than one other local optimum
- But more difficult to find, as the number of local optima increases exponentially with increasing K

General network statistics

Weighted clustering coefficient

local density of the network

$$c^w(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{w_{ij} + w_{ih}}{2} a_{ij} a_{jh} a_{hi}$$

where $s_i = \sum_{j \neq i} w_{ij}$, $a_{nm} = 1$ if $w_{nm} > 0$, $a_{nm} = 0$ if $w_{nm} = 0$ and $k_i = \sum_{j \neq i} a_{ij}$.

Disparity

dishomogeneity of nodes with a given degree

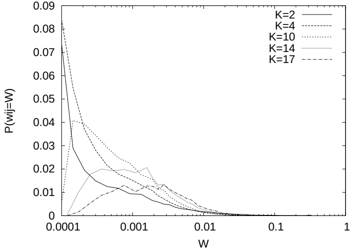
$$Y_2(i) = \sum_{j \neq i} \left(\frac{w_{ij}}{s_i} \right)^2$$

General network statistics $N = 16$

K	$\sharp$ nodes	$\sharp$ edges	$\bar{C}^w$	$\bar{Y}$	$\bar{d}$
2	33_{15}	516_{358}	$0.96_{0.0245}$	$0.326_{0.0579}$	56_{14}
4	178_{33}	9129_{2930}	$0.92_{0.0171}$	$0.137_{0.0111}$	126_{8}
6	460_{29}	41791_{4690}	$0.79_{0.0154}$	$0.084_{0.0028}$	170_{3}
8	890_{33}	93384_{4394}	$0.65_{0.0102}$	$0.062_{0.0011}$	194_{2}
10	$1,470_{34}$	162139_{4592}	$0.53_{0.0070}$	$0.050_{0.0006}$	206_{1}
12	$2,254_{32}$	227912_{2670}	$0.44_{0.0031}$	$0.043_{0.0003}$	207_{1}
14	$3,264_{29}$	290732_{2056}	$0.38_{0.0022}$	$0.040_{0.0003}$	203_{1}
15	$3,868_{33}$	321203_{2061}	$0.35_{0.0022}$	$0.039_{0.0004}$	200_{1}

- **Clustering Coefficient** : For high K, transition between a given pair of neighboring basins is less likely to occur
- **Disparity** : For high K the transitions to other basins tend to become equally likely, an indication of the randomness of the landscape

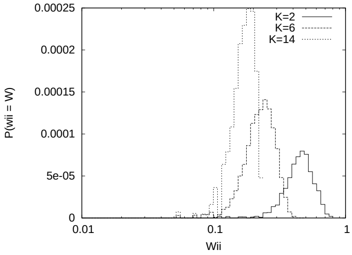
Weights distribution : transition probability between basins



distribution of the network weights w_{ij} for outgoing edges with $j \neq i$ in log-x scale, $N = 18$

- Weights are small
- For high K the decay is faster
- Low K has longer tails
- On average, the transition probabilities are higher for low K (less local optima)

Weight distribution remain in the same basin



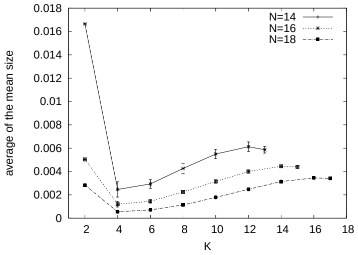
Average weight w_{ii} according to the parameter N and K

Question :

Is it easy to escape a basin ?

- Weights to remains in the same are large compare to w_{ij} with $i \neq j$
- w_{ii} are higher for low K
- Easier to leave the basin for high K : high "natural" exploration
- But : number of local optima increases fast with K

Interior and border size



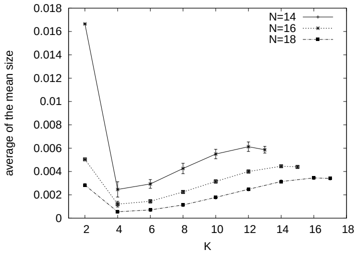
Average of the mean size of basins interiors

Question :

Do basins look like a "montain" with interior and border ?

solution is in the interior if all neighbors are in the same basin

Interior and border size



Average of the mean size of basins interiors

Question :

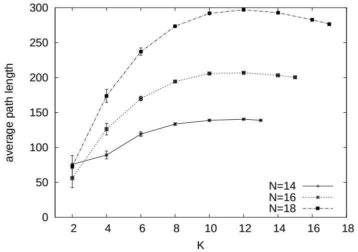
Do basins look like a "montain" with interior and border ?

solution is in the interior if all neighbors are in the same basin

Answer

- Interior is very small
- Nearly all solution are in the border

Shortest path length between local optima



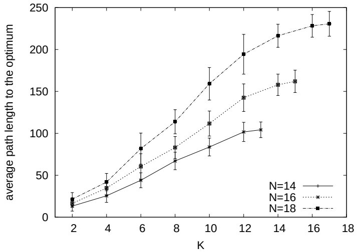
Average distance (shortest path) between nodes

Question :

Are the basins "far" from each other ?

- Increase with N (♯ of nodes increases exponentially)
- For a given N, increase with K up to $K = 10$, then stagnates

Shortest path length to global optima



Average path length to the global optimum from all the other basins

Question :

Is the global optimum basin is far ?

- More relevant for optimisation
- Increase steadily with increasing K

Local Optima Network of the Quadratic Assignment Problem

Please, you can come to the talk on Wednesday, July 21, 11 :50AM, Room 118.

Local Optima Network with other hill-climbing like first-improvement heuristic

Please, you can come to the talk at PPSN 2010, Krakow, Poland, September 2010.

Summary on local optima network

- Medium level of description : proposed characterization of combinatorial landscapes as networks
- a new model for landscape analysis
- New findings about basin's structure : sizes, fitness *vs.* size, etc.
- Related some *network features* to *search difficulty*

Future on local optima network

- Design a method for sampling large search space (under construction)
- Compare the properties of Loc. Opt. Network and the optimal tradeoff between exploration and exploitation
- Study the LON like a fitness landscape
- Deduce some approximation of the runtime from the properties of LON

Summary on fitness landscapes

Fitness landscape is a representation of

- search space
- notion of neighborhood
- fitness of solutions

Summary on fitness landscapes

Fitness landscape is a representation of

- search space
- notion of neighborhood
- fitness of solutions

Goal :

- local description : fitness between neighbor solutions
  Ruggedness, local optima, fitness cloud, neutral networks, local optima networks...
- and to deduce global features :
  - Difficulty !
  - To decide (and control) a good choice of the representation, operator and fitness function

Open questions

- How to control the parameters and/or operators of the algorithm with the local description of fitness landscape ?
- Can fitness landscape describe the dynamics of a population of solutions ?
- Links between neutrality and fitness difficulty ?
- Which intermediate description shows relevant properties of the optimization problem according to the local search heuristic ?
- What is the fitness landscapes for a *multiobjective problem* ?

Integration of the FL tools into the open framework *paradisEO*
http://paradiseo.gforge.inria.fr

L. Barnett.
Ruggedness and neutrality - the NKp family of fitness landscapes.
In C. Adami, R. K. Belew, H. Kitano, and C. Taylor, editors, *ALIFE VI, Proceedings of the Sixth International Conference on Artificial Life*, pages 18–27. ALIFE, The MIT Press, 1998.

Lionel Barnett.
Netcrawling - optimal evolutionary search with neutral networks.
In *Proceedings of the 2001 Congress on Evolutionary Computation CEC2001*, pages 30–37, COEX, World Trade Center, 159 Samseong-dong, Gangnam-gu, Seoul, Korea, 27-30 2001. IEEE Press.

U. Bastolla, M. Porto, H. E. Roman, and M. Vendruscolo.
Statiscal properties of neutral evolution.
Journal Molecular Evolution, 57(S) :103–119, August 2003.

Meriema Belaidouni and Jin-Kao Hao.
An analysis of the configuration space of the maximal constraint satisfaction problem.
In *PPSN VI : Proceedings of the 6th International Conference on Parallel Problem Solving from Nature*, pages 49–58, London, UK, 2000. Springer-Verlag.

P. Collard, M. Clergue, and M. Defoin Platel.
Synthetic neutrality for artificial evolution.
In *Artificial Evolution : Fourth European Conference AE'99*, pages 254–265. Springer-Verlag, 2000.
Selected papers in Lecture Notes in Computer Sciences 1829.

J. C. Culberson.
Mutation-crossover isomorphisms and the construction of discrimination function.
Evolutionary Computation, 2 :279–311, 1994.

J. P. K. Doye.
The network topology of a potential energy landscape : a static scale-free network.
Phys. Rev. Lett., 88 :238701, 2002.

J. P. K. Doye and C. P. Massen.
Characterizing the network topology of the energy landscapes of atomic clusters.
J. Chem. Phys., 122 :084105, 2005.

Ricardo Garcia-Pelayo and Peter F. Stadler.
Correlation length, isotropy, and meta-stable states.
Physica D, 107 :240–254, 1997.
Santa Fe Institute Preprint 96-05-034.

Josselin Garnier and Leila Kallel.
Efficiency of local search with multiple local optima.
SIAM Journal on Discrete Mathematics, 15(1) :122–141, 2002.

P. Gitchoff and G. Wagner.
Recombination induced hypergraphs : A new approach to mutation-recombination isomorphism, 1996.

David E. Goldberg and Philip Segrest.
Finite markov chain analysis of genetic algorithms.
In *ICGA*, pages 1–8, 1987.

M. Huynen.
Exploring phenotype space through neutral evolution.
Journal Molecular Evolution, 43 :165–169, 1996.

E. Izquierdo-Torres.
The role of nearly neutral mutations in the evolution of dynamical neural networks.
In J. Pollack and al, editors, *Ninth International Conference of the Simulation and Synthesis of Living Systems (Alife 9)*, pages 322–327. MIT Press, 2004.

T. Jones.
Evolutionary Algorithms, Fitness Landscapes and Search.
PhD thesis, University of New Mexico, Albuquerque, 1995.

S. A. Kauffman.
The Origins of Order.
Oxford University Press, New York, 1993.

M. Kimura.
The Neutral Theory of Molecular Evolution.
Cambridge University Press, Cambridge, UK, 1983.

J. Lobo, J. H. Miller, and W. Fontana.
Neutrality in technology landscape, 2004.

M. Newman and R. Engelhardt.
Effect of neutral selection on the evolution of molecular species.

In *Proc. R. Soc. London B.*, volume 256, pages 1333–1338, 1998.

Erik Van Nimwegen, James P. Crutchfield, and Martijn Huynen.
Metastable evolutionary dynamics : Crossing fitness barriers or escaping via neutral paths ?
Technical Report 99-07-041, SanteFe institute, 1999.

Gabriela Ochoa, Marco Tomassini, Sébastien Verel, and Christian Darabos.
A Study of NK Landscapes' Basins and Local Optima Networks.
In *Proceedings of the 10th annual conference on Genetic and evolutionary computation Genetic And Evolutionary Computation Conference*, pages 555–562, Atlanta États-Unis d'Amérique, 07 2008. ACM New York, NY, USA.
best paper nomination.

M. Defoin Platel.
Homologie en Programmation Génétique - Application à la résolution d'un problème inverse.
PhD thesis, Université de Nice Sophia Antipolis, France, 2004.

Eduardo Rodriguez-Tello, Jin-Kao Hao, and Jose Torres-Jimenez.
A new evaluation function for the minla problem.
In *Proceedings of the MIC 2005*, pages 796–801, Vienna Austria, 2005.

Helge Rosé, Werner Ebeling, and Torsten Asselmeyer.
The density of states - a measure of the difficulty of optimisation problems.
In *Parallel Problem Solving from Nature*, pages 208–217, 1996.

P. Schuster, W. Fontana, P. F. Stadler, and I. L. Hofacker.

From sequences to shapes and back : a case study in RNA secondary structures.
In *Proc. R. Soc. London B.*, volume 255, pages 279–284, 1994.

Peter F. Stadler.
Landscapes and their correlation functions.
J. Math. Chem., 20 :1–45, 1996.

Peter F. Stadler and W. Schnabl.
The landscape of the traveling salesmen problem.
Phys. Letters, A(161) :337–344, 1992.

Peter F. Stadler and Gunter P. Wagner.
Algebraic theory of recombination spaces.
Evolutionary Computation, 5(3) :241–275, 1997.

Terry Stewart.
Extrema selection : Accelerated evolution on neutral networks.

In *Proceedings of the 2001 Congress on Evolutionary Computation CEC2001*, pages 25–29, COEX, World Trade Center, 159 Samseong-dong, Gangnam-gu, Seoul, Korea, 27-30 May 2001. IEEE Press.

Marco Tomassini, Sébastien Verel, and Gabriela Ochoa.
Complex-network analysis of combinatorial spaces : The NK landscape case.
Physical Review E : Statistical, Nonlinear, and Soft Matter Physics, 78(6) :066114, 12 2008.
89.75.Hc ; 89.75.Fb ; 75.10.Nr.

Vesselin K. Vassilev and Julian F. Miller.
The advantages of landscape neutrality in digital circuit evolution.
In *ICES*, pages 252–263, 2000.

Sebastien Verel, Philippe Collard, and Manuel Clergue.
Measuring the evolvability landscape to study neutrality.

In M. Keijzer and et al., editors, *Poster at Genetic and Evolutionary Computation – GECCO-2006*, pages 613–614, Seatle, 8-12 July 2006. ACM Press.

Sébastien Verel, Gabriela Ochoa, and Marco Tomassini.
The Connectivity of NK Landscapes' Basins : A Network Analysis.
In *Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems Artificial Life XI*, pages 648–655, Winchester France, 08 2008. MIT Press, Cambridge, MA.
tea team.

E. D. Weinberger.
Correlated and uncorrelatated fitness landscapes and how to tell the difference.
In *Biological Cybernetics*, pages 63 :325–336, 1990.

S. Wright.

The roles of mutation, inbreeding, crossbreeding, and selection in evolution.
In *Proceedings of the Sixth International Congress of Genetics 1*, pages 356–366, 1932.