Tutorial: Recent advances in fitness landscapes

Hendrik Richter

HTWK Leipzig University of Applied Sciences

Faculty of Electrical Engineering & Information Technology

Leipzig, Germany

Content: What I will talk about

1. Principles and perspectives of fitness landscapes
2. Landscapes, topology & evolutionary paths
3. Measures & problem hardness
4. Dynamic problems & coevolution
5. Summary, challenges & open questions

Content: What I will NOT talk about

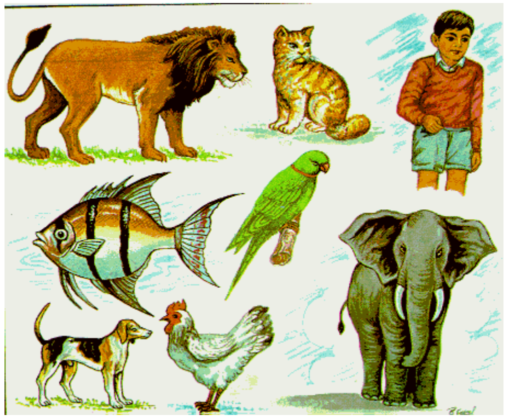
1. Details and comparison of landscape measures
2. Implications of landscape topology for algorithmic design
3. Empirical landscapes
4. Landscapes of specific optimization problems
5. Visualization of fitness landscapes

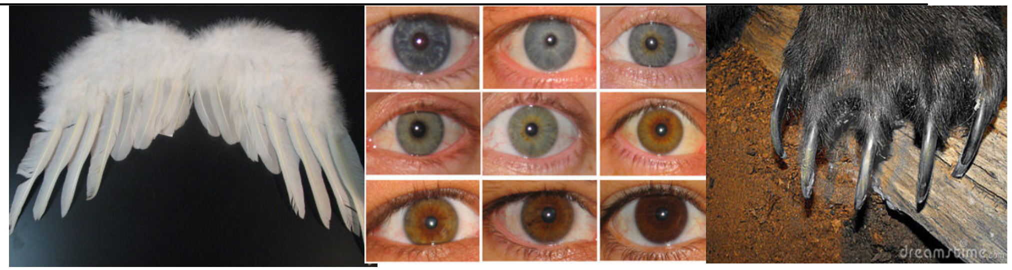
1. Principles and perspectives of fitness landscapes

Background, definitions, terminology, meaning

What is fitness landscape? Some general idea behind evolution

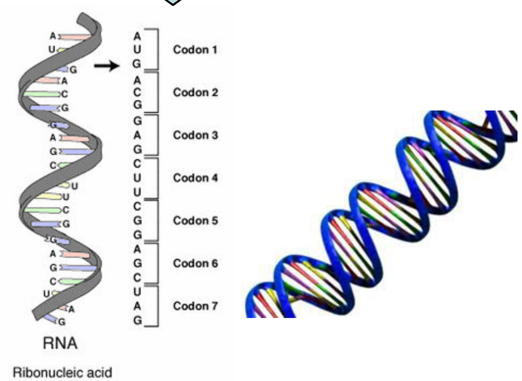
Living entities





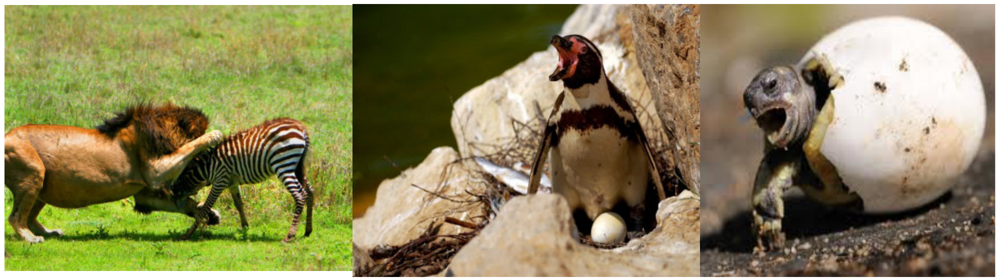
Inheritable features, traits and abilities

Information flow



Specified by (genetic) code

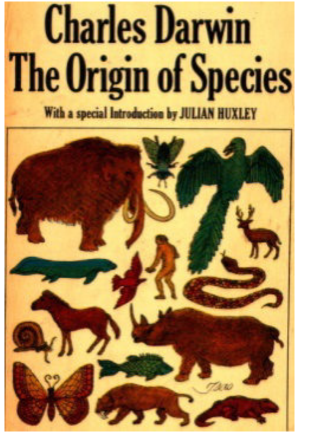
Survival and reproduction success



Source of figures: www.wikipedia.com

Evolution

- **Evolution: Dynamics of inheritable traits over consecutive generations**
  - Traits vary among individuals of a population
  - Differences in traits mean different rates in survival/reproduction labeled as fitness
  - Superior traits are inheritable by (genetic) code



- **Evolutionary algorithms: search methods analogous to natural evolution**
  - Population-based approach with inheritable search point features
  - Fitness evaluation of all individuals of a population
  - Parallelized generational and random-driven search process

Source of figure: www.amazon.com

What is fitness landscape? Some general idea behind evolution

Genetic coding spans a space consisting of all genetically possible individuals
→ **genotype**

Actually existing individuals are an instance of the genotype
→ **phenotype**

Reproduction and survival success (of actually existing individuals)
→ **fitness**

Questions:
How are the possibilities within the genetic code realized by the individuals ?

Which phenotypic realization of the genotype have (or would have) highest fitness?

Do evolutionary paths from a given phenotype to the high fitness region in genotypic space exist? (and where are they?)

What is fitness landscape? Some general idea behind evolution

Additional questions:
What relationships between the optimization problem, its fitness landscape and the evolutionary search method exist?

Is there a relation between certain aspects in the fitness landscape and the expectable behavior of the evolutionary search?

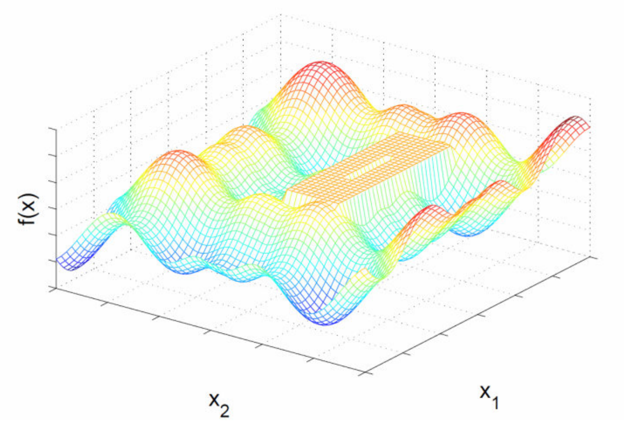
Can these relations be exploited for designing evolutionary search methods?

Evolutionary biology Evolutionary computation

different fields, different issue, different applications
but
sometimes
similar questions and similar mathematical methods

Motivation and background

What is a fitness landscapes?



Relationship between genotype, phenotype and fitness

Geometric interpretation:
2D fitness landscape metaphor
(static topological features: valleys, peaks, ridges, plateaus but also dynamic consequences of topology: lakes and flows)

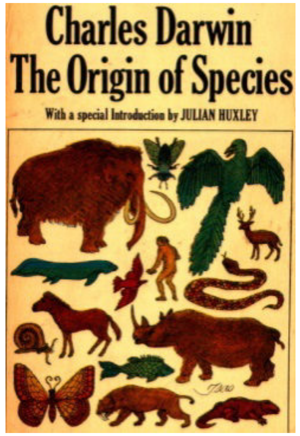
Potentials for (evolutionary) dynamics

i.e. driving forces behind evolutionary processes

→ move operators for phenotypic realizations
→ phenotypic realizations "walk" on the "fitness surface"

Evolution and evolutionary dynamics

- **Evolution** → dynamic process (the walking up the hills or staying close to the hill)



- **Requires:** → differences in fitness over genotypic space
  → Darwinian imperative: move into the direction of increasing fitness
  → codified by generating offspring and (natural and sexual) selection

- evolutionary dynamics studies 'the principles according to which life has evolved and continues to evolve'

- different manifestations of evolutionary dynamics (natural or artificial) can be understood using the same framework → fitness landscapes

- shape of fitness landscape (together with population size and mutation rate) defines evolutionary dynamics

Hendrik Richter
HTWK Leipzig, EIT, MSR

Evolution and evolutionary dynamics

- **another aim of this tutorial:** studying evolutionary dynamics is a topic in its own right

- extends the rather narrow focus of optimally designing evolutionary search algorithms

- considerable amount of recent works on defining, experimenting with, measuring and visualizing natural and theoretical fitness landscapes

- embedded in general trends to redefine foundations of biology in term of mathematization and algorithmization

- forms the base for addressing further questions such as

  - how does a population deal with environmental changes (dynamic landscapes)

  - how do several populations interact and coevolve (coevolutionary and codynamic landscapes)

Motivation and background

How and why is fitness landscape useful ?

- **as metaphor:** gives a “picture” of evolutionary path (= sequence of increasing fitness)
  if landscape has certain feature (e.g. several hills, or big valley or flat regions) than expect certain types of evolutionary dynamics

  however: is bounded to low dimensions (generally 2)
  might be misleading

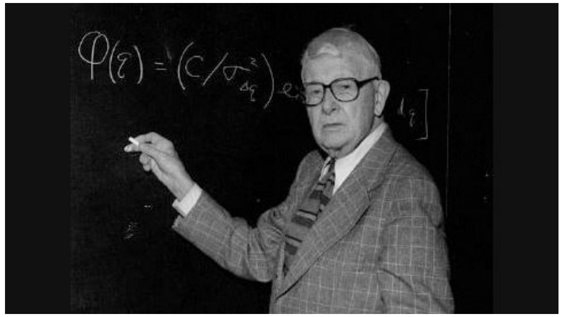
- **as tool for quantification:** turn the metaphorical pictures into mathematical concepts
  → feature extraction
  → landscape measures
  → visualization concepts

What is fitness landscape? Historical remarks

• Fitness landscape concepts were introduced by Sewall Wright in 1932

• Concepts influential for thinkers in evolutionary biology (Fisher, Haldane, Mayr, …)

• Provided an abstract way to understand evolutionary dynamics



Source of figure: http://institucional.us.es/darwin09/Wright.htm

• But: no computational approach

• First instance of computable models: NK landscapes by Kauffman & Levin (1987)

• Paved the way for explaining genetic algorithms by fitness landscape, Manderick et al. (1991), Jones (1995)

• Since then: much work on fitness landscapes to understand evolutionary computation processes

• Most recently: direct experimental approach in evolutionary biology, → empirical fitness landscapes, Poelwijk et al. (2007), Lobokovsky et al. (2011)

• adaptation of bacteria in an antibiotic → triggering the mutations controlling resistance → direct experimental relationship between genotype and fitness

What is fitness landscape? Definition

Relationships between genotype, phenotype and fitness (but also between all possible solutions, candidate solutions, and solution quality of optimization problems)

Mathematical formulation: (static) fitness landscape

$$\Lambda_S = (S, n, f)$$

Configuration space

- genotypic space
- space of all possible solutions

Neighborhood structure

What is next to what in configuration space?

Produces locations

Fitness function

Quality information for configuration space

Gives elevation (or height) to locations

Configuration space

Different application contexts employ (almost) synonymous terms

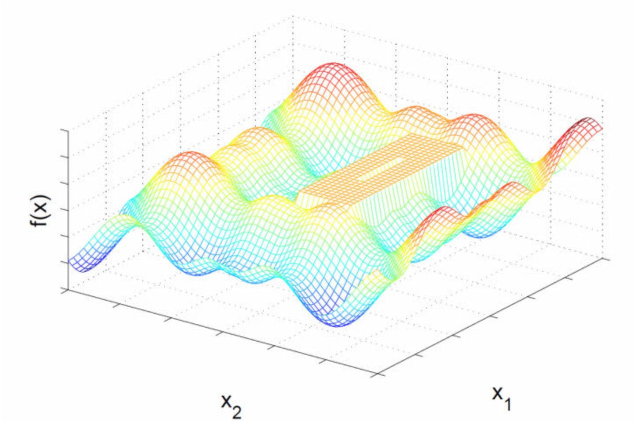
•Configuration space

•Genotypic space

•Sequence space

•Search space

•Representation space



All specify the space over which the landscape is defined (= sets out locations that have fitness allocated as constituting property), but have different origins and (slightly) different meanings.

Configuration space

Configuration space	most general term: made up by the finite (or infinite) number of configurations the genetic description of a natural or artificial biological system can have
Genotypic space	mathematical description of natural biological systems → genetic description defined as genotype → genotypic space
Sequence space	as genotypic space → genetic description as sequences (= string) over a finite alphabet, e.g. the DNA alphabet {A,T,C,G} the RNA alphabet (U = T) or binary genomic equivalents {0,1}
Search space	solutions space of evolutionary search algorithms
Representation space	as search space → underlining the fact that different representations of the search algorithm result in different search spaces

Neighborhood structure: consequence of genetic move operator

from a given point in configuration space
what other points are reached in one genetic move
(= neighborhood is a static description of a dynamic aspect)

Genetic move operators: Mutation and recombination

mutational neighborhoods → mutational landscapes → mutational trajectories
recombinational neighborhoods → recombinational landscapes → recomb. trajectories

Biological background mutational landscapes:
Strong selection/weak mutation model (Gillespie, 1984)

→ Population monomorphic → selection lets beneficial mutations go to fixation before the next mutation occurs → mutations are random flips in the letters of the finite alphabet at single points on the string

Neighborhood structure

Mutational landscapes: single bit flip neighbors of $\{00\}$ → $\{01\}\,\{10\}$

Binary representation and Hamming distance =1 → simplest model (and limited biological relevance)

$$\{A, T, G, C\} = \{00, 01, 10, 11\} \qquad \{00\} \rightarrow \{01\}\,\{10\}\,\{11\}$$

Recombinational landscapes:

Even more complicated and generally an open problem:
Needs polymorphism → allows jumps through genotypic space

some results that show for non-homologous recombination (no similar snippets of DNA are exchanged) the resulting genotypic space is not metric (Shipak & Wagner, 2000, Stadler et al., 2001)

Also search space neighborhood structure depends on move operator → 'one operator, one landscape'

Fitness

Fitness in biology → connected to longevity, fertility and ultimately reproduction success

Assigning fitness to actual (micro-)biological entities is highly debatable

Fitness is phenotype's viability expressed as the fact of surviving to the age of reproduction and actually reproducing → only after lifetime of phenotype

Moreover: fitness is not only connected to reproducing but reproducing offspring that itself survives and reproduces disproportionally

→Assigning fitness requires a time window of overlapping generations

In short: assigning fitness is complicated and might be controversial

Possible remedy:
- considering fitness as axiomatic property of the landscape
- fitness proxies (empirical landscapes), e.g. growth rates of bacteria, resistance to antibiotic etc.

Fitness

Fitness in evolutionary computation: usually more easy to define

Search space consists of a finite (or infinite) number of candidate solutions

For each of them → calculating an objective function

Possible complications: dynamic or coevolutionary fitness → discussed later

What is fitness landscape? Summary

$$\Lambda_S = (S, n, f)$$

Configuration Space	Neighborhood Structure	Fitness Function
Biology		
- Genotypic space	- Mutation - Recombination	- definable only via phenotypic space - axiomatic or empirical property
Optimization (search space)		
- real (integer) numbers	- Inherent metrics	
		- objective function
- binary trees	- Hamming distance	
- trees	- branch of subtrees	

2. Landscapes, topology & evolutionary paths

Geometric conception and computational experiments

Motivation and background

Structure and topology of the fitness landscape → predictor of evolutionary dynamics

Fuel for discussion about fitness landscapes

A) Geometric intuition and conception

B) Computational experiments with designed models of landscapes

C) (Computational) experiments with observed models of landscapes

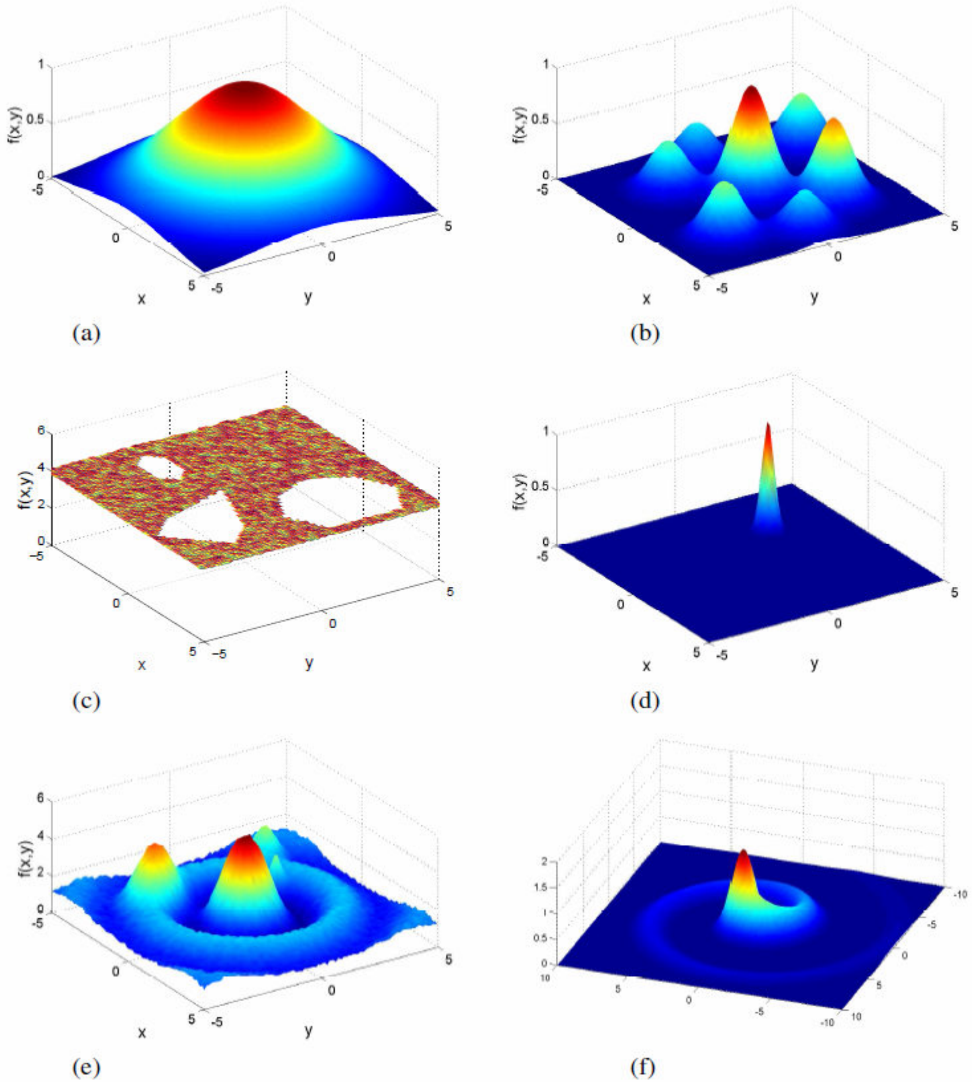
A) and B) are shared in evolutionary biology and evolutionary computation

C) takes different forms → empirical landscapes
→ landscapes of optimization problems

Fitness landscapes: Geometric intuition and conception

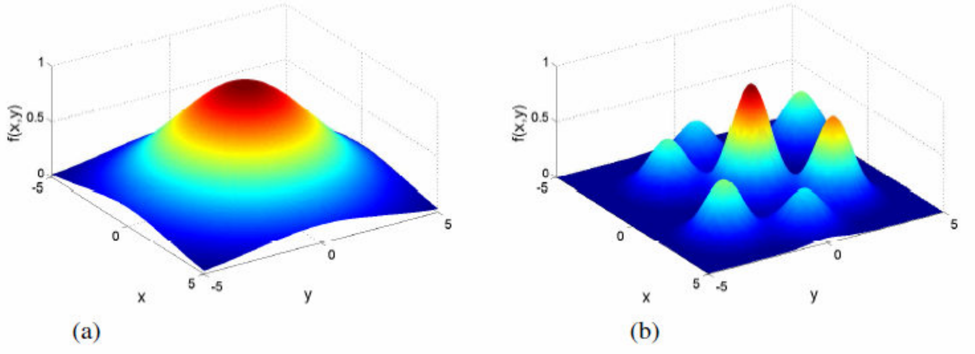
Topological features and consequences for evolutionary pathways: How likely are the paths?

a) Single smooth peak (Mt. Fuji) → Evolutionary hill climbing
b) Rugged landscapes with multiple peaks → Basins of attraction and valley crossing
c) Holey landscapes → Finding ridged between optima
d) Neutral landscape with needle-in-the-haystack → Neutral drift, jumps in fitness and unguided search
e) Barrier landscape → Conditional valley crossing
f) Detour landscape (long-path problem) → Small paths



Fitness landscapes: Geometric intuition and conception

Minimal and maximal ruggedness:



Mt. Fuji

vs.

House-of-cards-model

Source of figures: www.wikipedia.com

→Fitness of neighboring points in landscape is uncorrelated random
→Performing a genetic move is like adding a card to a house-of-cards: it may either extend or collapse it

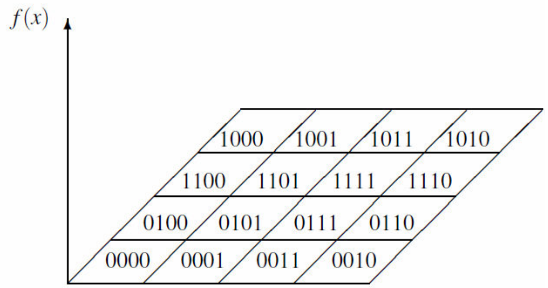
Example: NK fitness landscape by Kauffman & Levin, 1987

Genotypes coded with a string of length *N* over a given alphabet
N = number of components in the system

For instance: binary alphabet, *N=4* → genotype of in total 16 elements
Neighborhood structure = Hamming distance 1

0000	**0001**	**0011**	**0010**
0100	**0101**	**0111**	**0110**
1100	**1101**	**1111**	**1110**
1000	**1001**	**1011**	**1010**

Configuration space
+
Neighborhood structure
=
'Location'



Defining fitness
K number of (epistatic) interactions = degree of interaction between the building elements of each genotype
$0 \leq K \leq N - 1$

Example: NK fitness landscape by Kauffman & Levin

$$f(x) = \frac{1}{N}\sum_{i=0}^{N-1} f_i(x_i; n(x_i, K))$$

$$x = x_0 x_1 x_2 \ldots x_{N-1}$$

0000	0001	0011	0010
0100	0101	0111	0110
1100	1101	1111	1110
1000	1001	1011	1010

$f_i(x_i; n(x_i, K))$ Fitness contribution of bit x_i

$n(x_i, K)$ Neighborhood function of bit x_i specifying *K* neighboring bits

→*K* specifies how many interactions contribute to fitness

Two types of neighborhood functions

Nearest neighbor interaction	Random interaction
Takes $\frac{K}{2}$ neighbors left and right	Takes random neighbors (no repetition)

Example: NK fitness landscape by Kauffman & Levin

$$f(x) = \frac{1}{N}\sum_{i=0}^{N-1} f_i(x_i; n(x_i, K))$$

$\frac{K}{2}$ odd → right hand bias

0000	0001	0011	0010
0100	0101	0111	0110
1100	1101	1111	1110
1000	1001	1011	1010

N=4 $x = x_0 x_1 x_2 x_3 = 0110$

K=0

0110
0110
0110
0110

$$f(0110) = \frac{1}{4}\big(f_0(0) + f_1(1) + f_2(1) + f_3(0)\big)$$

K=1

interaction

0110
0110
0110
0110

$$f(0110) = \frac{1}{4}\big(f_0(01) + f_1(11) + f_2(10) + f_3(00)\big)$$

K=2

0110
0110
0110
0110

$$f(0110) = \frac{1}{4}\big(f_0(001) + f_1(011) + f_2(110) + f_3(100)\big)$$

Example: NK fitness landscape by Kauffman & Levin

Fitness contributions are randomly drawn from [0,1]

Lookup table with $N \times 2^{K+1}$ possible values

K=0

	0	1
f_0	0.3	0.2
f_1	0.7	0.5
f_2	0.6	0.1
f_3	0.3	0.8

$$f(0110) = \frac{1}{4}\left(f_0(0) + f_1(1) + f_2(1) + f_3(0)\right) = 0.3$$

$$f(1110) = \frac{1}{4}\left(f_0(1) + f_1(1) + f_2(1) + f_3(0)\right) = 0.275$$

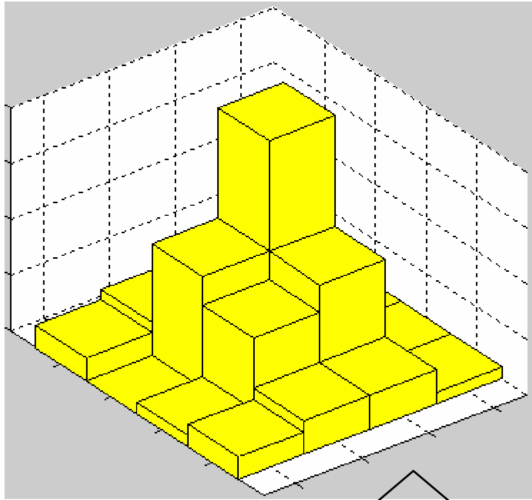
K=2

	000	001	011	010	100	101	111	110
f_0	0.7	0.9	0.6	0.6	0.1	0.9	0.4	0.3
f_1	0.4	0.3	0.5	0.8	0.6	0.7	0.5	0.2
f_2	0.2	0.5	0.8	0.2	0.3	0.4	0.7	0.7
f_3	0.6	0.4	0.1	0.9	0.1	0.3	0.2	0.8

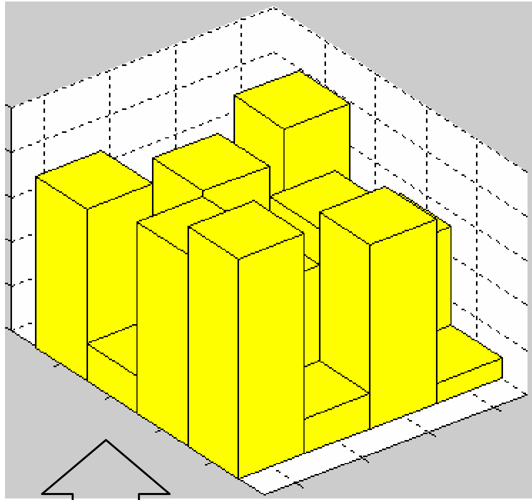
$$f(0110) = \frac{1}{4}\left(f_0(001) + f_1(011) + f_2(110) + f_3(100)\right) = 0.55$$

$$f(1110) = \frac{1}{4}\left(f_0(011) + f_1(111) + f_2(110) + f_3(101)\right) = 0.525$$

Example: NK fitness landscape by Kauffman & Levin



0000	0001	0011	0010
0100	0101	0111	0110
1100	1101	1111	1110
1000	1001	1011	1010



K=0

Smooth landscape
(Mt. Fuji)

K=N-1

Rugged landscape
(House-of-cards)

Landscapes over binary configuration space

Rough Mt. Fuji model (Aita & Husimi, 2000) $f(x) = \eta(x) - c \cdot d_H(x, x_0)$

$\eta(x)$ Realization of a random variable

$d_H(x, x_0)$ Hamming distance to reference genotype x_0

$c = 0$ House-of-cards → maximally rugged, maximally random maximally uncorrelated to neighbor's fitness

$c > 0$ Decreasing dependence on random → Single peak (Mt.Fuji) at $x = x_0$

Landscapes over real-valued configuration space

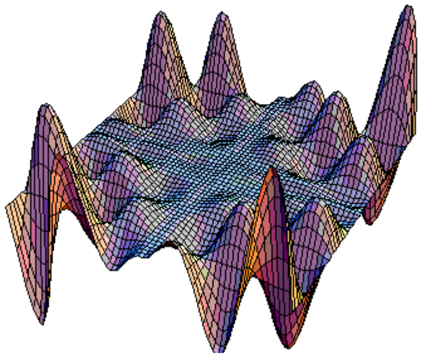
Every function $f(x) : \Re^n \to \Re$ can be seen as a landscape
Benchmark problems (Schwefel, Griewank, Rastrigin, Rosenbrock, etc. etc.)

What is fitness landscape? Limitation

Two types of criticisms

- **Dimensionality issue**
(biological genotype, but also search spaces are more than 2D)

    - landscape measures
    - landscape visualization tools



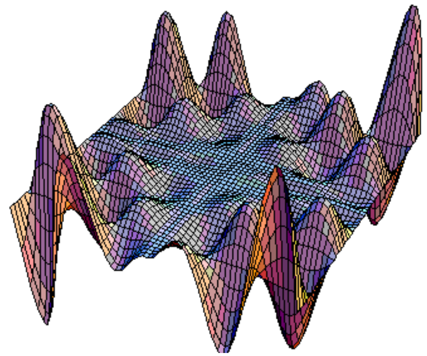
Source of figure: http://classes.yale.edu/fractals/CA/GA/Fitness/Fitness.html

- **Dynamics issue**
(the fitness of a genotype, but also the quality of a candidate solution of an optimization problem are not static)

    - dynamic fitness landscapes

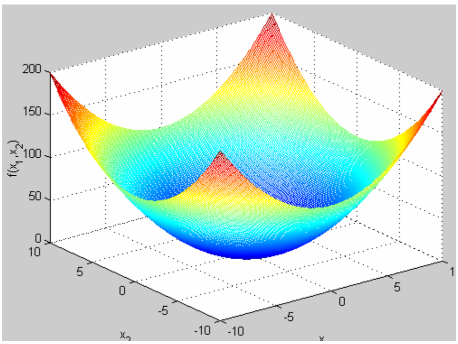
3. Measures & problem hardness

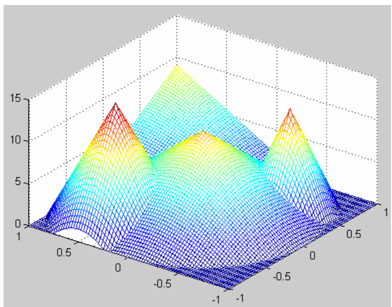
Landscape measures: How difficult is a certain optimization problem ?

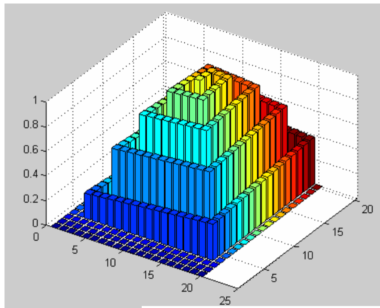


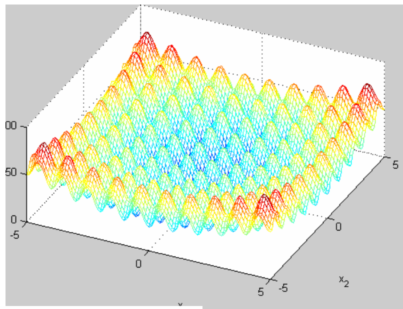
Evaluation of performance of an evolutionary algorithm
→ Landscape measures (= defining metrics)

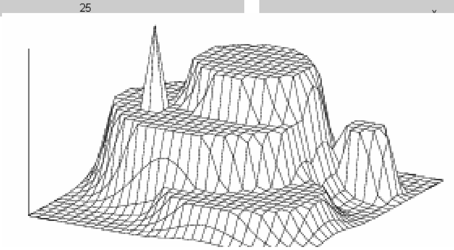
No simple answer to how difficult a certain landscape is to optimize in!









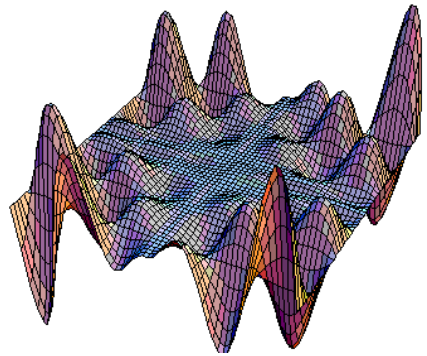


Easy

More difficult

Source of figure: http://www-lisic.univ-littoral.fr/~verel/recherche.html

Landscape measures: How difficult is a certain optimization problem ?



Generally: Fitness landscape

Geometrical object with features as

→ Number
→ Size
→ Form
→ Scattering

In terms of an optimization problem

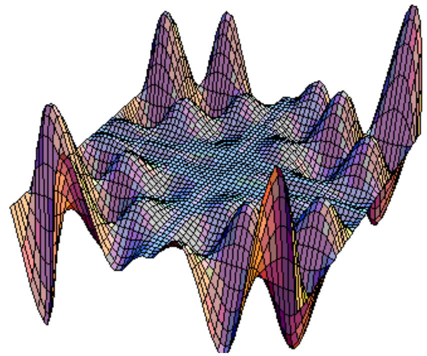
→ Number of optima — Easy to measure

→ Distribution in search space
→ Nature of the space between them — Rather difficult to measure

Landscape measures: How difficult is a certain optimization problem ?

Features contributing to problem hardness

→ Number of optima
→ Distribution in search space
→ Nature of the space between them



Another question: How do these features balance each other in terms of problem hardness ?

Landscape measures

→ Modality → Number of (local) optima
→ Ruggedness → Correlation structure
→ Information content → An entropic measure
→ Epistasis → Nonlinear interaction (Walsh measure)

Compressing complex landscape features in (hopefully meaningful) numbers

Landscape measures: How difficult is a certain optimization problem ?

Some rather bad news:

- no simple measure(s) to predict problem hardness generally for a given problem
- no direct prediction of algorithmic performance
- computationally feasible hardness measures futile (most likely)

Predictive version of hardness measures (for instance epistasis, fitness-distance correlation, etc.) that run in polynomial time do not exist unless P = NP (He et al., 2007)

Possible ways forward

Landscape measures: How difficult is a certain optimization problem ?

Recall No Free Lunch theorem

→ No search algorithm outperforms all others over all optimization problems

But: there is a correlation between an algorithmic design, a given problem, the expected behaviour and a prediction of performance

→ Matching of algorithms to problems (Malan & Engelbrecht, 2013)

- extract features of problems by landscape analysis
- select algorithm by the features
- predict performance based on the features and the algorithm
- apply algorithm to problem solving
- update base of prediction

4. Dynamic problems & coevolution

What is a dynamic optimization problem?

What exactly is a static optimization problem?

1. Search space S
2. Neighborhood structure $n(s) \quad \forall s \in S$
3. Fitness function $f(s)$

Find the search space point with a fitness higher (or at least equal) than the points in the neighborhood

$$f_S = \max_{s \in S} f(s)$$

Modify 1. 2. or 3. → relevant "modify 1" → most relevant "modify 3"

$$f_S^* = \max_{s \in S} f^*(s)$$

dynamic constraints

$$f(s,0) = f(s) \quad f(s,1) = f^*(s)$$

Rewrite

$$f_S(k) = \max_{s \in S} f(s,k), \forall k > 0$$

Introduce a time variable and obtain a dynamic optimization problem

What is a dynamic optimization problem?

When is it sensible to tackle a dynamic optimization problem?
Or asked differently: when is one problem dynamically originating from another

Series of static problems vs. One dynamic problem

$$f_S = \max_{s \in S} f(s) \quad f_S^* = \max_{s \in S} f^*(s)$$

$$f(s,0) = f(s) \quad f(s,1) = f^*(s)$$

$$f_S(k) = \max_{s \in S} f(s,k), \forall k > 0$$

In some ways a question of modelling but for evolutionary computation it makes sense if:

- some relation and alikeness between problems $f(s,k)$ and $f(s,k+1)$
- solving one problems gives information to solve the next one more efficiently (assumption similar problems best solved by similar algorithms)
- information usable to equip the evolutionary search with favorable settings (e.g. general parameters or such for diversity management)

Eco-evolutionary dynamics

• Conventional population genetics model → Genetic changes within a single population disregarding changes in the ecological context

• However: Changes in the genetic deposition in one species may affect fitness distribution of other species

→ particularly if species with well-defined ecological roles, e.g.

- predator – pray
- parasite – host

→ coevolutionary dynamics

• Selective pressure generated by the biotic and abiotic environment also affects fitness distribution

→ interaction of ecological and evolutionary dynamics

→ eco-evolutionary dynamics

Models of eco-evolutionary dynamics

• Conventional population genetics model → Genetic changes within a single population disregarding changes in the ecological context

• Not a very realistic view:
- Biotic and abiotic environment shapes evolutionary changes
- But also: genetic changes (and different traits) may alter ecological equilibria
- Result: ecological and evolutionary dynamics have overlapping timescales

→ Fitness landscape is dynamic → Model of dynamic fitness landscape

$$\Lambda_S = (S, n, f)$$

Configuration space
Neighborhood structure
Fitness function

$$\Lambda_D = (S, n, K, F, \Phi_F)$$

Time set

Set of fitness functions

Transition map
of fitness functions

Models of eco-evolutionary dynamics

$$\Lambda_D = \left(S, n, K, F, \Phi_F\right)$$

Time set: measuring and ordering scale for changes in the landscape $k \in K$

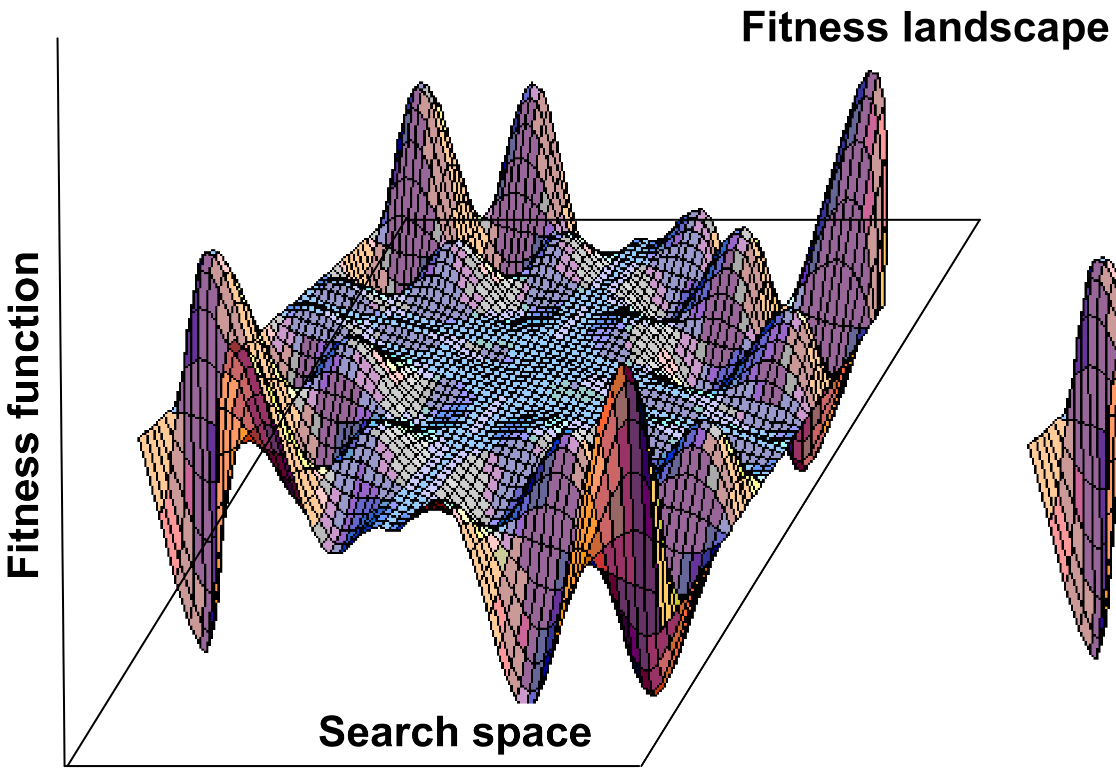
Set of fitness functions: Generalizes fitness function in time $f \in F$
and $f : S \times K \to \Re$

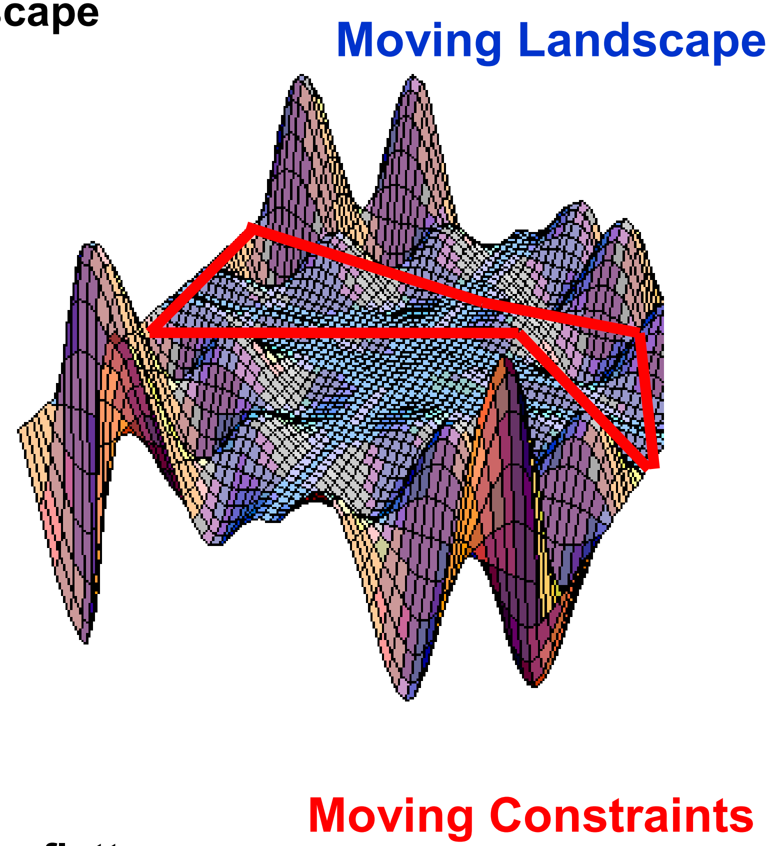
Transition map of fitness functions: defines how the fitness function changes over time

$$\Phi_F : F \times S \times K \to F$$

Definition allows to model continuous and discrete configuration spaces and time regimes
Changes in fitness landscape may happen (or may come into effect) at discrete points in time scale → sensible for computational approach
Continuous changes → PDE (or lattices of ODE) → numerical solution requires discretization

Dynamic Optimization





Fitness landscape metaphor remains valid:
But hills grow, shrink or move, valleys deepen or flatten,
Landscape completely or partially turns inside out

Modelling dynamics in dynamic fitness landscapes

How to specify the transitions between fitness functions?

Three answers:

Landscape model includes topology and dynamics → internal dynamics

Landscape model includes only topology and dynamics of topological features is specified outside the landscape model by an external driving systems → external dynamics

Landscape model is subject to feedback and changes from the population that evolves on it → population-based dynamics → coevolution → deformable landscapes

Modelling dynamics in dynamic fitness landscapes: Internal dynamics

Landscape model includes topology and dynamics → internal dynamics

Implies: Equations that specifies the timely evolution of every points in the landscape and changes may happen infinitely close to each other

Changes continuously in both space and time (infinite number of genotypes and change points)
→ Modeled as nonlinear Partial Differential Equation (PDE), for instance in 2D

$$\frac{\partial f}{\partial t} = \varphi\left(f, \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial^2 f}{\partial x_1{}^2}, \frac{\partial^2 f}{\partial x_2{}^2}, \ldots \right)$$

Spatially extended dynamical system

Timely evolution of fitness of any point depends on the actual fitness value and the partial derivatives of first and higher order
Geometrically interpreted: differences in fitness infinitesimally around the point
→ Spatially local (neighborhood-wise) deformation of the landscape

Modelling dynamics in dynamic fitness landscapes: Internal dynamics

For finite number of points in configuration space

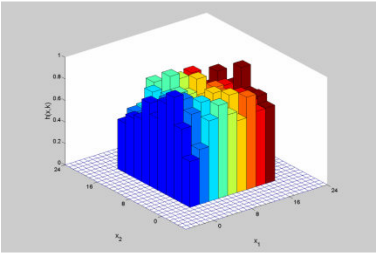
$$\frac{df(x_i,t)}{dt} = \varphi\big(f(x_1,t),\ldots f(x_i,t),\ldots, f(x_\mu,t)\big)$$

Time evolution of a spatially extended system: Lattice of (nonlinear) ODEs

Possibly the time evolution of fitness of x_i only depends on the fitness of its neighborhood

$$\frac{df(x_i,t)}{dt} = \varphi\big(f(x_i,t), f(n(x_i),t)\big)$$

However, any computational approach requires timely (and spatially) discretization



Numerical solution of PDE: Discretization of space and time
→ One way of doing this: Coupled Map Lattices (CML)

Coupled map lattices (CML), Richter, 2008

Spatial Extension **Temporal Evolution**

$$h(x_1,x_2,k+1)=(1-\varepsilon)g(h(x_1,x_2,k)+\frac{\varepsilon}{4}$$
$$[g(h(x_1-1,x_2,k)+$$
$$g(h(x_1+1,x_2,k)+$$
$$g(h(x_1,x_2-1,k)+$$
$$g(h(x_1,x_2+1,k)]$$

$$h(x_1,x_2,k),\ \ x_1=1,2,\dots,I_1, x_2=1,2,\dots,I_2$$

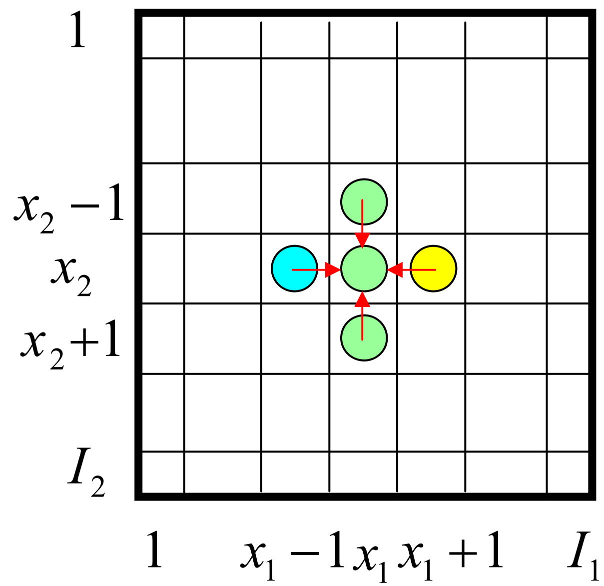
$$h(I_1+1,x_2,k)=h(1,x_2,k)$$
$$h(x_1,I_2+1,k)=h(x_1,1,k)$$

Periodic boundary condition

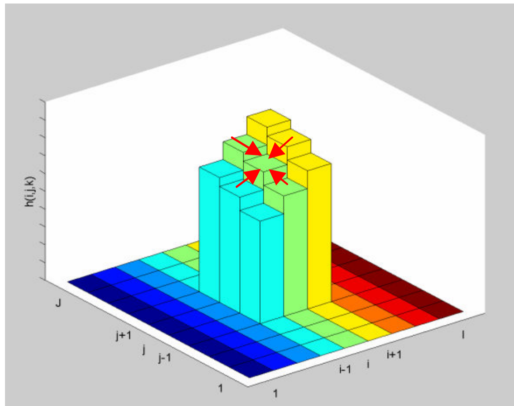
$$g(h(x_1,x_2,k))=\alpha\cdot h(x_1,x_2,k)(1-h(x_1,x_2,k))$$

Mapping function
Logistic equation: Complex and chaotic behavior

Space **Time**



k



k+1

Video

Modelling dynamics in dynamic fitness landscapes: External dynamics

Landscape model includes only topology and dynamics of topological features is specified outside the landscape model by an external driving systems → external dynamics

Only certain topological features in the landscape undergo changes

Requires: External driving system

Definition of moving sequence $z = \big(z(0), z(1), \ldots, z(k), z(k+1), \ldots\big)$

Specification of the type of dynamics

- Cyclic dynamics $z(k) = \sin(\varpi k + \delta)$
- Chaotic dynamics $z(k+1) = g\big(z(k)\big)$
- Random dynamics $z(k) = rand(k)$

Selection of the topological features that undergo changes

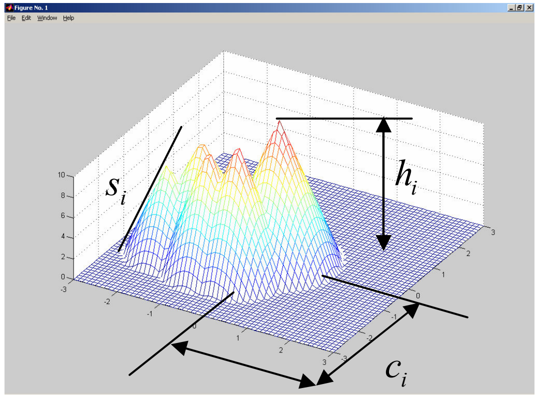
Alternatively: Selection of changing features of the objective function → Implicitly changes topological features of the landscape

Examples: Moving peaks

Dynamic fitness function: Moving peaks

Dynamic fitness function → **Fitness values over the search space change with time**

Benchmark: Moving peaks
Dynamic coordinates



Dynamic fitness function

$$c_i(k), k = 0,1,2,...$$

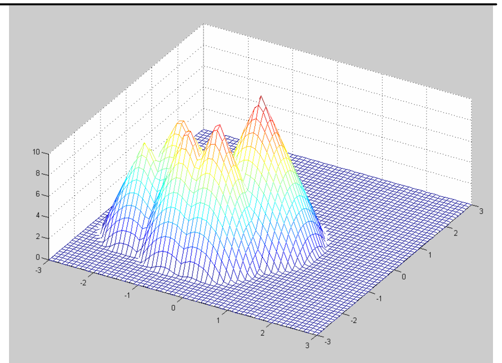
Solution trajectory

$$x_S(k) = \max_{x \in R^m} f(x,k) = \max\left\{0, \max_{1 \le i \le N}\left[h_i - s_i \left\| x - c_i(k) \right\|\right]\right\}$$

Parameters (and potential topological features to change): Number of peaks, shape and height of the peaks, (initial distribution, dimension of the fitness function)

Video

Dynamic fitness function: Benchmarks

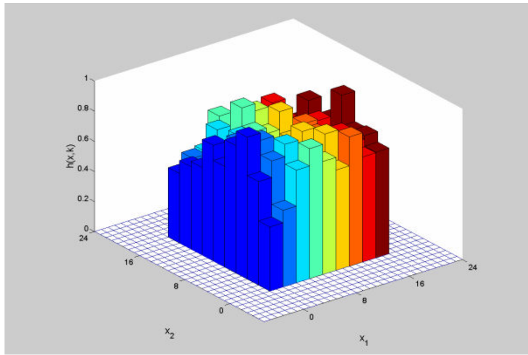


- **Moving peaks (Branke, Morrison & De Jong, 1999)**

- **XOR problem generator (Yang & Yao, 2005)**
Binary encoded static problem + bit-wise XOR change law

- **Dynamic knapsack problem (Mori et al., 1996)**
Static knapsack problem with dynamic knapsack size

- **Dynamic bit-matching (Stanhope & Diada, 1999)**
Use a binary template, and find the number of – matching bit + change the template



- **CML-based dynamic fitness landscapes (Richter, 2008)**

- **Dynamic NK landscape (Wilke & Martinetz, 1999, Barbazon et al., 2004)**

Population-based dynamics and coevolution

Landscape model is subject to feedback and changes from the population that evolves on it → population-based dynamics → coevolution → deformable landscapes

Coevolution: Evolutionary dynamics in one population (inter-)depends with dynamics in another population

Coevolution requires at least two populations

May be cooperative or competitive

Coevolutionary algorithms: Multi-population search algorithms employing coevolutionary dynamics

Features and issues of coevolutionary algorithms

Two coevolving populations
Fitness of (evolving) individuals in one population defined via fitness (or search point features) of (evolving) individuals of the other population

Interactive domain → Rules of reciprocal actions between (samples of) individuals of both populations
Solution concept → Translates interaction to individual’s fitness

Consequence: fitness is not longer a static property of a search space point
Coevolutionary fitness is subjective, depends on generational time, and on the selection of the individuals and their current fitness

Consequence: What is coevolutionary progress?

Fitness landscape view

Coevolution: Codependent and codynamic fitness landscapes

For each population, there is a specific fitness landscape

Consequences: coevolutionary dynamics interacts with the topological features of the landscape

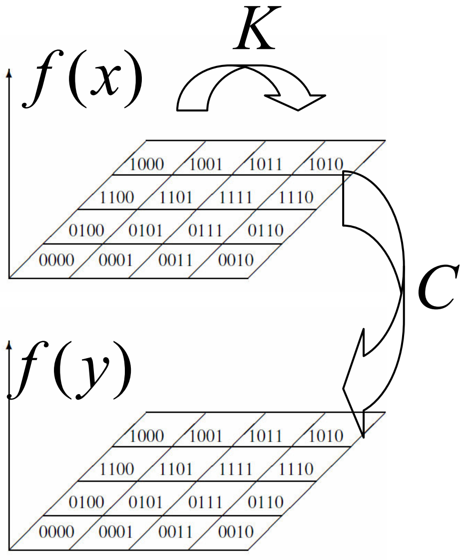
→ (co-)dynamic and deformable landscapes

Example: NKC fitness landscape by Kauffman & Johnsen

Two coevolving landscapes

$$f(x) = \frac{1}{N}\sum_{i=0}^{N-1} f_i(x_i; n(x_i, K, C)) \qquad x = x_0 x_1 x_2 \ldots x_{N-1}$$

$$f(y) = \frac{1}{N}\sum_{i=0}^{N-1} f_i(y_i; n(y_i, K, C)) \qquad y = y_0 y_1 y_2 \ldots y_{N-1}$$



K internal (epistatic) interactions (with each genotype)
C external (epistatic) interactions (from one genotype to the other)

$$n(x_i, K, C) = n(x_i, K) \mid n(x_i, y_i, C)$$ Concatenated neighborhood function

$n(x_i, K)$ Internal neighborhood function (usually identical)

$n(x_i, y_i, C)$ External neighborhood function (usually symmetric)

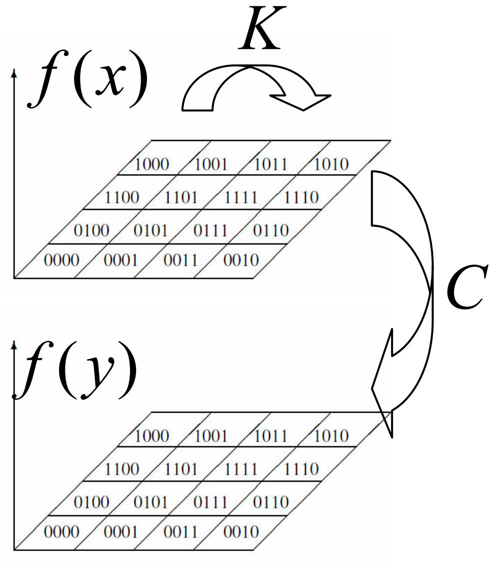
Example: NKC fitness landscape by Kauffman & Johnsen

Lookup table with $N \times 2^{K+C+1}$ possible values
(for each landscape, not identical, not symmetric)

N=4, K=1, C=1

	00|0	00|1	01|1	01|0	10|0	10|1	11|1	11|0
f_0^x	0.7	0.9	0.6	0.6	0.1	0.9	0.4	0.3
f_1^x	0.4	0.3	0.5	0.8	0.6	0.7	0.5	0.2
f_2^x	0.2	0.5	0.8	0.2	0.3	0.4	0.7	0.7
f_3^x	0.6	0.4	0.1	0.9	0.1	0.3	0.2	0.8

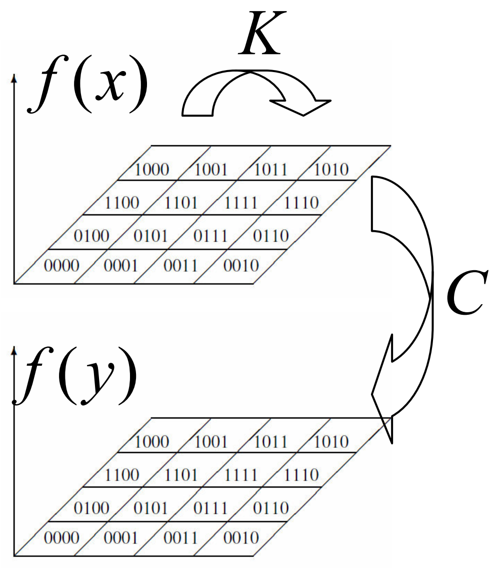
	00|0	00|1	01|1	01|0	10|0	10|1	11|1	11|0
f_0^y	0.3	0.4	0.1	0.1	0.7	0.1	0.6	0.8
f_1^y	0.7	0.1	0.7	0.8	0.6	0.2	0.5	0.9
f_2^y	0.2	0.4	0.3	0.9	0.5	0.5	0.8	0.1
f_3^y	0.5	0.9	0.2	0.2	0.1	0.4	0.2	0.2



Two coevolving landscapes

Example: NKC fitness landscape by Kauffman & Johnsen

Lookup table with $N \times 2^{K+C+1}$ possible values
(for each landscape, not identical, not symmetric)



N=4, K=1, C=2

$$x = x_0 x_1 x_2 x_3 = 0110$$

$$y = y_0 y_1 y_2 y_3 = 1001$$

$$f^x(0110) = \frac{1}{4}\left(f_0^x(01|10) + f_1^x(11|00) + f_2^x(10|01) + f_3^x(00|11)\right)$$

$$f^y(1001) = \frac{1}{4}\left(f_0^y(10|01) + f_1^y(00|11) + f_2^y(01|10) + f_3^y(11|00)\right)$$

Two coevolving landscapes

$$x = x_0 x_1 x_2 x_3 = 1110$$

$$y = y_0 y_1 y_2 y_3 = 1001$$

$$f^x(1110) = \frac{1}{4}\left(f_0^x(11|10) + f_1^x(11|00) + f_2^x(10|01) + f_3^x(01|11)\right)$$

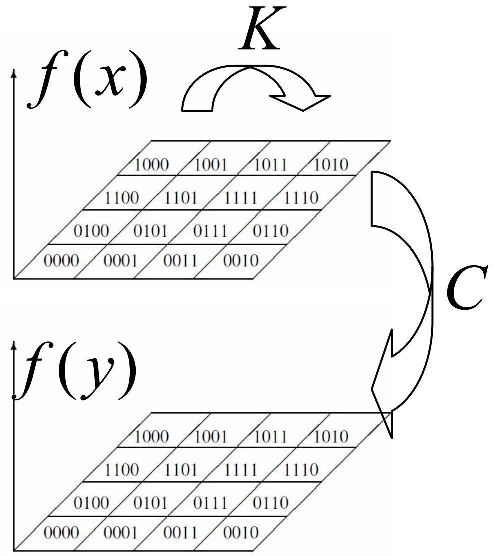
$$f^y(1001) = \frac{1}{4}\left(f_0^y(10|11) + f_1^y(00|11) + f_2^y(01|10) + f_3^y(11|01)\right)$$

Change in genotype in one landscape alters fitness of genotype in the other landscape

Example: NKC fitness landscape by Kauffman & Johnsen

$$x = x_0x_1x_2x_3 = 0110$$
$$y = y_0y_1y_2y_3 = 1001$$
$$f^x(0110) = \frac{1}{4}\left(f_0^x(01|10) + f_1^x(11|00) + f_2^x(10|01) + f_3^x(00|11)\right)$$
$$f^y(1001) = \frac{1}{4}\left(f_0^y(10|01) + f_1^y(00|11) + f_2^y(01|10) + f_3^y(11|00)\right)$$

$$x = x_0x_1x_2x_3 = 0110$$
$$y = y_0y_1y_2y_3 = 1011$$
$$f^x(0110) = \frac{1}{4}\left(f_0^x(01|10) + f_1^x(11|01) + f_2^x(10|11) + f_3^x(00|11)\right)$$
$$f^y(1011) = \frac{1}{4}\left(f_0^y(10|11) + f_1^y(01|11) + f_2^y(11|10) + f_3^y(11|00)\right)$$



Fitness evaluation of a given genotype in one landscape *(x)* can only be done with respect to the other landscape *(y)*
For an ongoing fitness evaluation
→ deformation of the fitness landscape

Evolutionary dynamics: codependent vs. codynamic

Properties of the NKC model:

As fitness of a genotype in one landscape can only be assigned by specifying a co-genotype in the other landscape → fitness is codependent

Codepedence is in genotype, not in fitness

The same selection of codependent genotype gives a consistent fitness value

→Not codynamic
→Would become codynamic by look-up table entries that depend on evolutionary time

Codynamic and deformable landscape require other models

In particular, the fitness landscape must depend on (generational) time k

Deformable fitness landscape (Ebner & Watson)

Motivating thoughts: Population dynamics (succession of positions of individuals in the landscape) changes topology and hence deforms the landscape

Biological background: individuals utilize abilities and features connected to their phenotype → interaction, competition and cooperation with other individuals → adaption process of the interacting parties → diminishing of fitness attached to genotype → bulging the landscape

Closely related to Red Queen effect (moving in genotypic space to maintain the same level of fitness)

Consequence of these thinking: Coevolution and codynamic can be observed in a single landscape

Deformable Landscapes (Ebner & Watson)

Deformable fitness landscape (Ebner & Watson)

Deformable Landscapes (Ebner & Watson)

Model: Each individual $p_i(k)$ of a population $P(k)$ bulges a given constant landscape

Bulging is modelled as negative Gaussian hills

$$-\exp\left(-\frac{1}{2}(x-P(k))^T A(x-P(k))\right)$$

A Dilation of spatial deformation

Bulging is temporally smoothened (also by a Gaussian function)

$$\exp\left(\frac{(i-k-\tau_{Lat})^2}{2\sigma^2}\right)$$

i temporal counter: maximal deformation $i = k + \tau_{Lat}$

τ_{Lat} latency: move maximal deformation in time k

σ dilation of timely deformation

Deformable fitness landscape (Ebner & Watson)

Static landscape: zero plane

$$f(x)$$

Dynamic population-based landscape

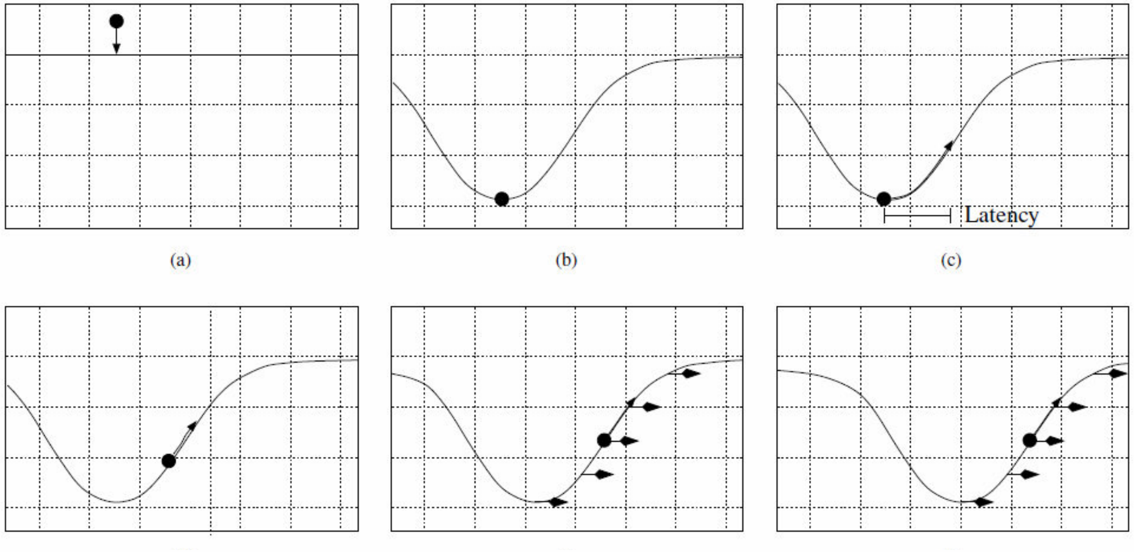
$$f(x,k) = f(x) - \sum_{i=1}^{I_{end}} \exp\left(\frac{\left(i-k-\tau_{Lat}\right)^2}{2\sigma^2}\right) \exp\left(-\frac{1}{2}\left(x-P(k)\right)^T A\left(x-P(k)\right)\right)$$

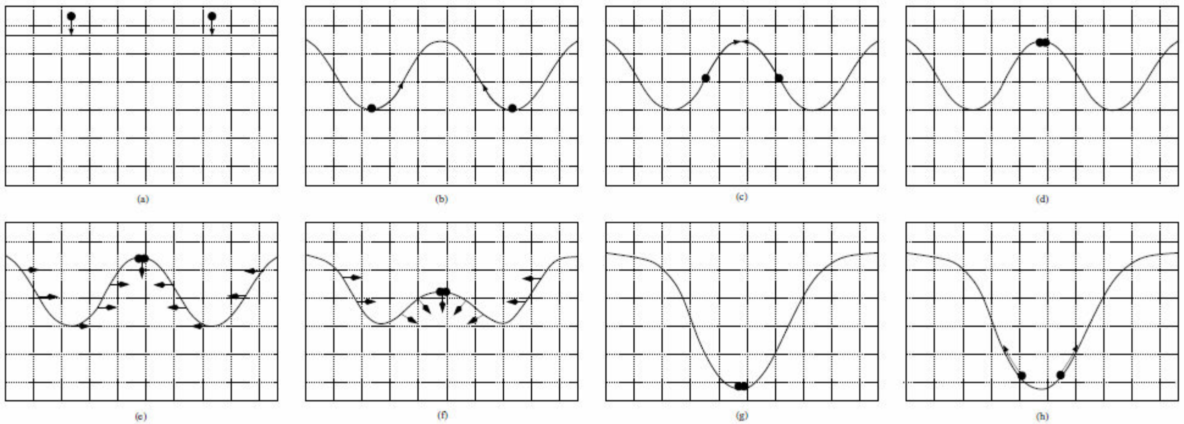
Population dynamics: Hill-climbing → feedback from landscape's topology

Different types of hill-climbing dynamics (= different update rules), for instance

$$P(k+1) = \alpha\left(\left.\frac{\partial f(x,k)}{\partial x}\right|_{x=P(k)}\right) \cdot P(k)$$

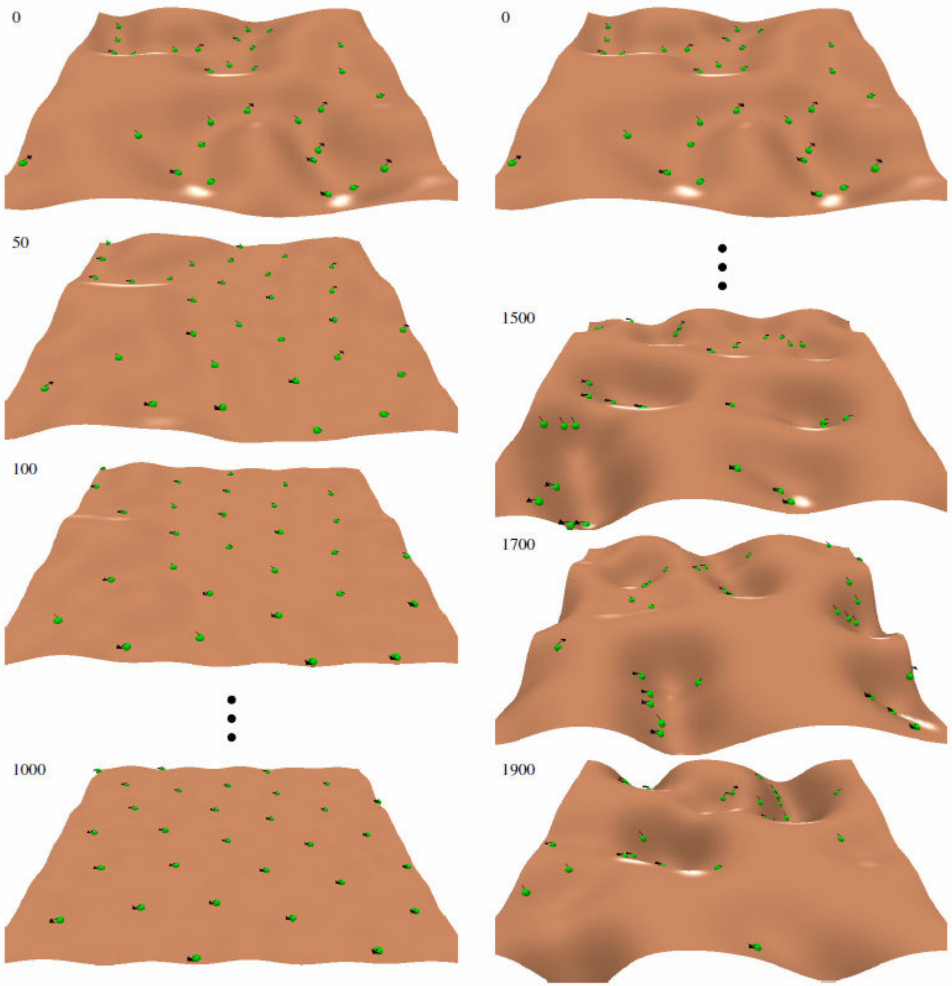
Deformable fitness landscape (Ebner & Watson)





From: Watson & Ebner: Eco-evolutionary dynamics on deformable fitness landscapes. In: H. Richter, A. P. Engelbrecht. Recent Advances in the Theory and Application of Fitness Landscapes. Springer, 2014, 339-368

Deformable fitness landscape (Ebner & Watson)



Video

From: Watson & Ebner: Eco-evolutionary dynamics on deformable fitness landscapes. In: H. Richter, A. P. Engelbrecht. Recent Advances in the Theory and Application of Fitness Landscapes. Springer, 2014, 339-368

Further issues in dynamic landscapes

So far: mainly pictorial and qualitative description

Required extension: dynamic landscape measures and possible links to evolutionary search behavior and performance (Richter, 2010)

Proposed dynamic landscapes measures:

- dynamic severity (how far away is the next optimum)
- Lyapunov exponent (how predictable is it)

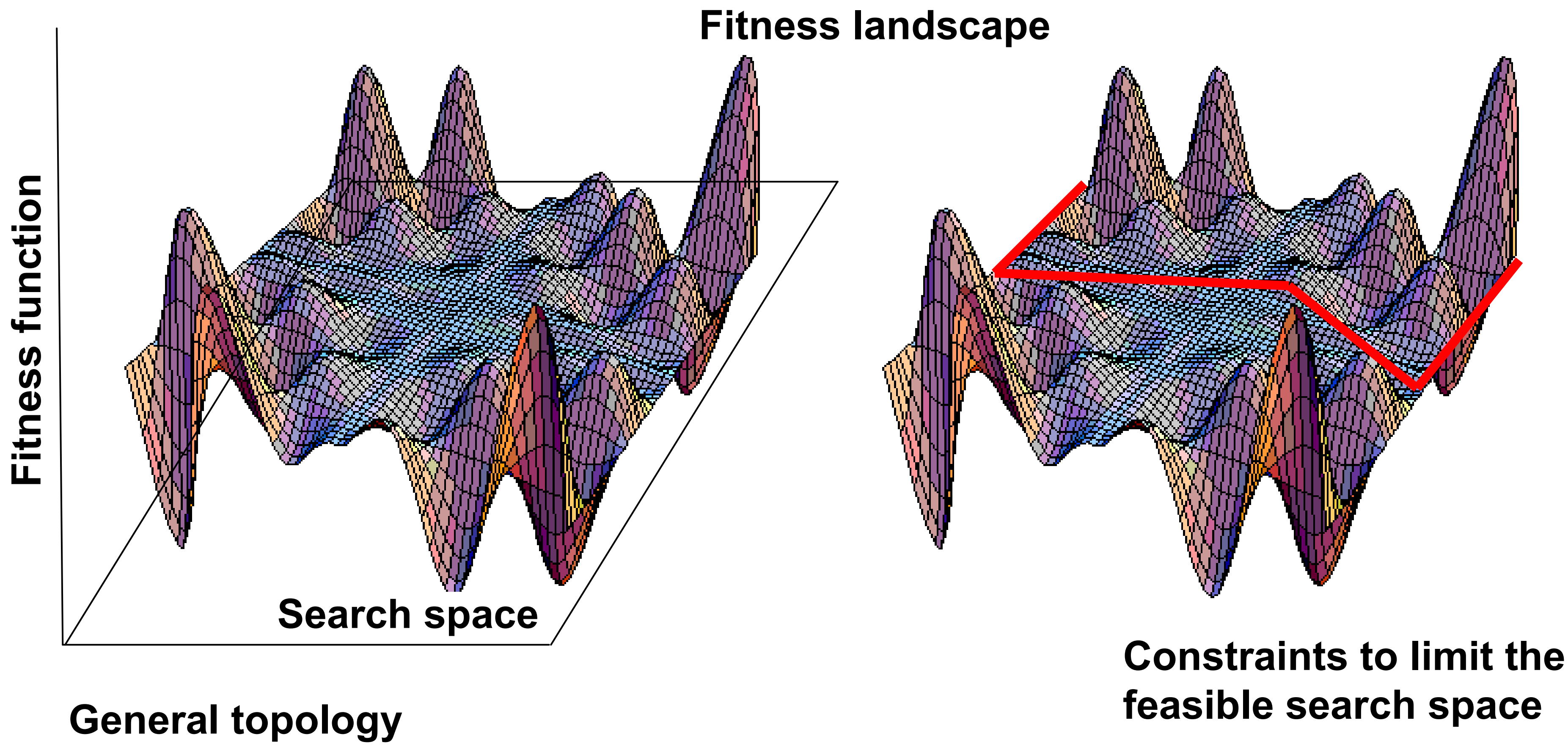
But: main factor in performance is relative speed (change frequency)
Only definable relative to the timescale of evolutionary search algorithm

Moreover: coevolutionary (codynamic) landscape change every generation

Possible approach: similarity measures (later in conference, Richter, 2014b)

5. Summary, challenges & open questions

Structure and geometry of configuration space and fitness function



Immediate 'geographic' consequences: mountains and valleys → topology
Further consequences: lakes and flows → evolutionary dynamics
→ evolutionary paths

What is fitness landscape? The ideas of evolutionary paths

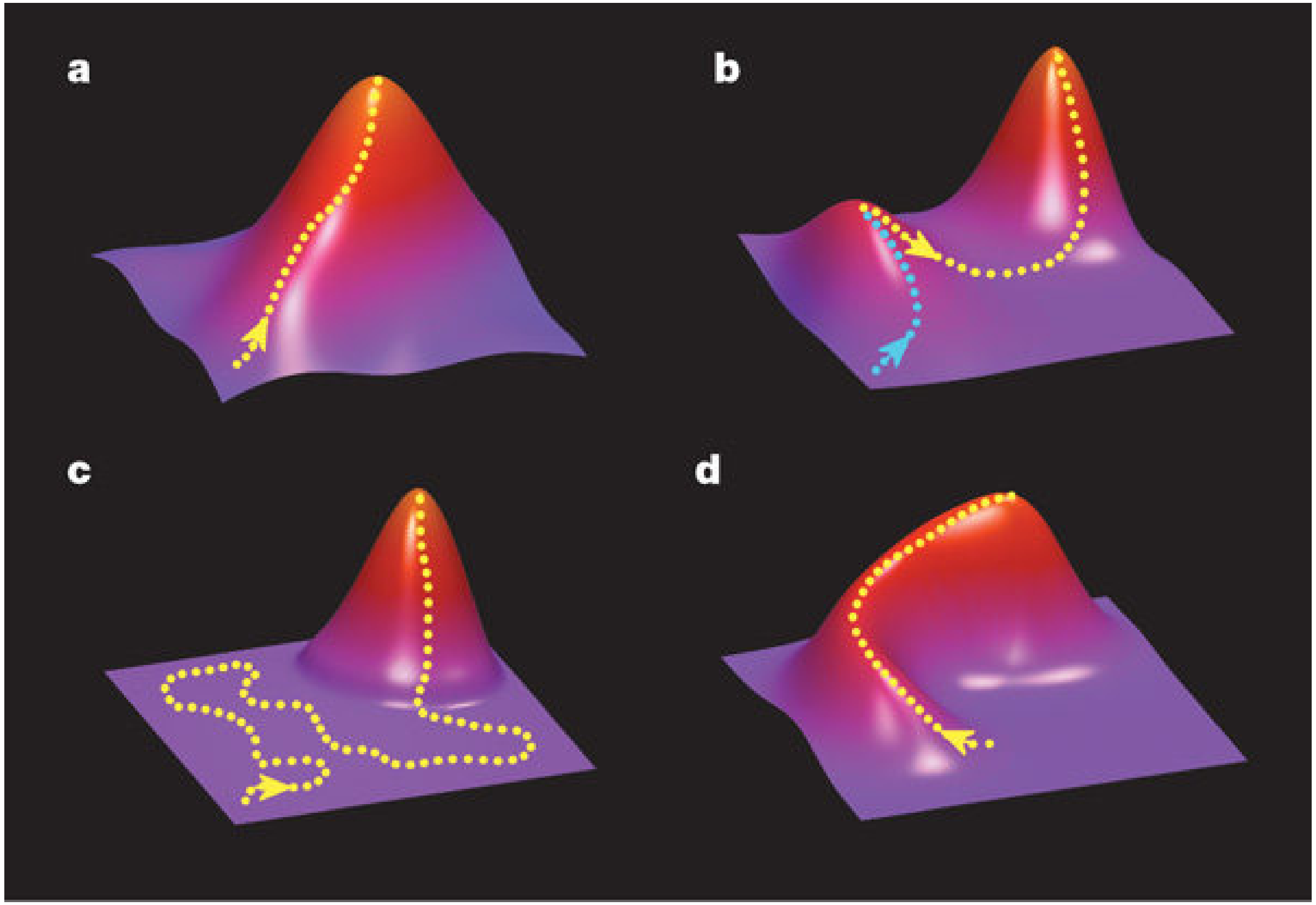
How likely are the paths?

Smooth single peak

rugged multimodal

neutrality

detour



From Poelwijk et al. 2007

Challenges to fitness landscape thinking

Challenges to our understanding of genetic moves on the landscape:

Current fitness landscape research mainly focuses on mutational landscapes strong selection/weak mutation (SSWM)- model

→ Information transfer and inheritance much more complicated as „classically" assumed

„Central dogma of molecular biology" (Crick) = Information flow: DNA → RNA → protein → upper organizational levels of complex biological systems (RNA networks, sub-cellular processes, cells, tissues, organs, organisms) → Vertical gene transfer
In short: DNA uniquely specifies a phenotypic realization

Most likely an oversimplification: multitude of additional information processing activities that specify the traits and abilities of living matter, e.g. horizontal gene transfer (HGT): nonhereditary transfer of genetic material

Challenges to fitness landscape thinking

Nonhereditary transfer: moves in genotypic space during the lifetime of a phenotype

Also: Transfer of genetic information from DNA to RNA to protein → not exclusive and unidirectional

Translation process:
Not only reading, decoding and converting information from DNA to protein
But: RNA performs regulation processes depending on the molecular environment in which the translation process takes place

- determines which genomic region is accessible
- activates and/or silences promoters to control activity of genes
- repairs the genome using compensatory mutations

Molecular environment: stimulated by the organism it belongs to and ecological processes the organism is exposed to
In short: Phenotypic realization depends not only on the code but on how the code is processes → merger of ecological and evolutionary timescales

Challenges to fitness landscape thinking

What is fitness signifying?

Generally: phenotypic traits aggregate to individual quality in surviving and reproduction → correlates positively with fitness

But: assigning fitness to actual biological entity is highly debatable

Reason: fitness is scalar abstraction of multiple phenotypic traits and their significance in a given environment

Interpretation: Fitness = metric on a highly complex data set
= result of abstraction, aggregation and interpretation
= depends on the parameter of these processes

However: no sensible alternative

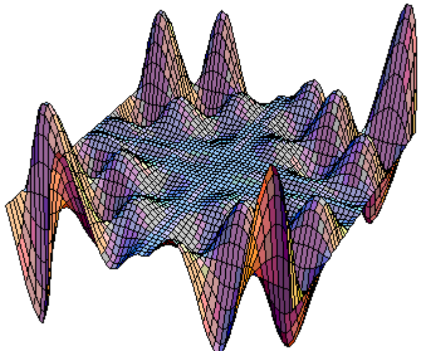
How do different metrics interfere with resulting topology (and evolutionary dynamics)?

Open questions

- predictions about evolutionary paths from landscape‘s topology
- genetic moves other than mutation
- landscape models of coevolutionary processes
- empirical landscapes
- meaning and significance of fitness

Concluding remarks

Fitness landscape	→ Relationships between the genetically possible (genotype), the actually realized (phenotype) and its quality (fitness) → Structure and topology of the landscape define possible evolutionary paths → Possible evolutionary paths open up possibilities for predicting evolutionary outcome
Thesis:	Evolutionary dynamics will become a own field of research within evolutionary computation → framework of fitness landscapes is an integral part of evolutionary dynamics
Fitness landscape	Not only a tool for supporting the design of evolutionary search algorithms
Recent challenges	Substantial demand for further research

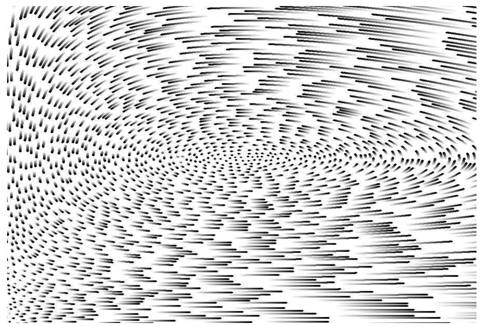


Thank you !

Hendrik Richter

HTWK Leipzig University of Appiled Sciences
Faculty of Electrical Engineering & Information Technology

richter@eit.htwk-leipzig
http://www.msr.eit.htwk-leipzig.de/institut/personen/



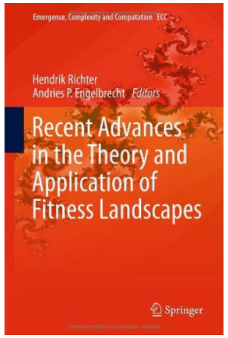
References

The tutorial is mainly based on:

Hendrik Richter. Fitness landscapes: from evolutionary biology to evolutionary computation. In: Recent Advances in the Theory and Application of Fitness Landscapes, (Eds.: H. Richter, A. P. Engelbrecht), Springer-Verlag, Berlin, 2014, 3-31.

Hendrik Richter. Fitness landscapes that depend on time. In: Recent Advances in the Theory and Application of Fitness Landscapes, (Eds.: H. Richter, A. P. Engelbrecht), Springer-Verlag, Berlin, 2014, 265-299.

Hendrik Richter. Frontiers of fitness landscapes: a summary of open questions. In: Recent Advances in the Theory and Application of Fitness Landscapes, (Eds.: H. Richter, A. P. Engelbrecht), Springer-Verlag, Berlin, 2014, 529-544.



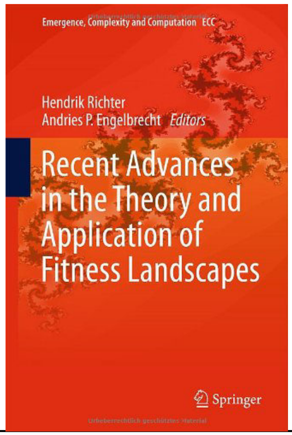
References

The tutorial is mainly based on:

Hendrik Richter. Fitness landscapes: from evolutionary biology to evolutionary computation. In: Recent Advances in the Theory and Application of Fitness Landscapes, (Eds.: H. Richter, A. P. Engelbrecht), Springer-Verlag, Berlin, 2014, 3-31.

Hendrik Richter. Fitness landscapes that depend on time. In: Recent Advances in the Theory and Application of Fitness Landscapes, (Eds.: H. Richter, A. P. Engelbrecht), Springer-Verlag, Berlin, 2014, 265-299.

Hendrik Richter. Frontiers of fitness landscapes: a summary of open questions. In: Recent Advances in the Theory and Application of Fitness Landscapes, (Eds.: H. Richter, A. P. Engelbrecht), Springer-Verlag, Berlin, 2014, 529-544.



Source of figure: www.amazon.com

References

T. Aita, Y. Husimi. Adaptive walks by the fittest among finite random mutants on a Mt. Fuji-type fitness landscape. II. Effect of small non–additivity. J. Math. Biol. 41: 207–231, 2000.

A. Brabazon, A. Silva, T. F. de Sousa, M. O'Neill, R. Matthews, E. Costa, A particle swarm model of organizational adaptation. In: K. Deb (ed.), Proc. Genetic and Evolutionary Computation Conference, GECCO 2004, Springer– Verlag, Berlin, 12–23, 2004.

J. Branke, Memory enhanced evolutionary algorithms for changing optimization problems. In: P. J. Angeline, Z. Michalewicz, M. Schoenauer, X. Yao, A. Zalzala (eds.), Proc. Congress on Evolutionary Computation, IEEE CEC 1999, IEEE Press, Piscataway, NJ, 1875–1882, 1999.

J. H. Gillespie, Molecular evolution over the mutational landscape. Evolution 38: 1116–1129, 1984.

M. Ebner, R. A. Watson, J. Alexander, Co–evolutionary dynamics on a deformable landscape. In: A. Zalzala, C. Fonseca, J. H. Kim, A. Smith, X. Yao (eds.), Proc. Congress on Evolutionary Computation, IEEE CEC 2000, IEEE Press, Piscataway, NJ, 1284–1291, 2000.

M. Ebner, R. A.Watson, J. Alexander, Coevolutionary dynamics of interacting species. In: C. Di Chio et al. (eds.), Applications of Evolutionary Computation - EvoApplications 2010, Springer–Verlag, Berlin Heidelberg New York, 1–10, 2010.

J. He, C. Reeves, C. Witt, X. Yao, A note on problem difficulty measures in black-box optimization: Classification, realizations and predictability. Evol. Comp. 15: 435-443, 2007.

References

T. Jones, Evolutionary algorithms, fitness landscape and search. PhD thesis, The University of New Mexico, Albuquerque, 1995.

S. A. Kauffman, S. Levin, Towards a general theory of adaptive walks on rugged landscapes. J. Theor. Biology 128: 11–45, 1987.

S. A. Kauffman, S. Johnsen, Coevolution to the edge of chaos: Coupled fitness landscapes, poised states, and coevoluti onary avalanches. J. Theor. Biol. 149: 467–505, 1991.

A. E. Lobkovsky, Y I. Wolf, E. V. Koonin, Predictability of evolutionary trajectories in fitness landscapes, PLoS Comput Biol 7(12): e1002302. doi:10.1371/journal.pcbi.1002302, 2011.

K. M. Malan, A. P. Engelbrecht, A survey of techniques for characterising fitness landscapes and some possible ways forward. Information Science 241:148–163, 2013.

B. Manderick, M. deWeger, P. Spiessens, The genetic algorithm and the structure of the fitness landscape. In: R. K. Belew, L. B. Booker (eds.) Proceedings of the Fourth International Conference on Genetic Algorithms, Morgan Kauffman, San Mateo, CA, 143–150, 1991.

N. Mori, H. Kita, Y. Nishikawa, Adaption to a changing environment by means of the thermodynamical genetic algorithm. PPSN IV, 513-522 ,1996

References

R. W. Morrison, K. A. De Jong, A test problem generator for non–stationary environments. In: P. J. Angeline, Z. Michalewicz, M. Schoenauer, X. Yao, A. Zalzala (eds.), Proc. Congress on Evolutionary Computation, IEEE CEC 1999, IEEE Press, Piscataway, NJ, 2047–2053, 1999.

F. J. Poelwijk, D. J. Kiviet, D. M. Weinreich, S. J. Tans, Empirical fitness landscapes reveal accessible evolutionary paths. Nature 445: 383–386, 2007.

H. Richter, Coupled map lattices as spatio–temporal fitness functions: Landscape measures and evolutionary optimization. Physica D237: 167–186, 2008.

H. Richter, Evolutionary optimization and dynamic fitness landscapes: From reaction-diffusion systems to chaotic CML. In: Evolutionary algorithms and chaotic systems, (Eds.: I. Zelinka, S. Celikovsky, H. Richter, G. Chen), Springer-Verlag, Berlin, 409-446, 2010.

H. Richter, A. P. Engelbrecht, Recent Advances in the Theory and Application of Fitness Landscapes. Springer, 2014a.

H. Richter, Codynamic fitness landscapes of coevolutionary minimal substrates. In: Proc. IEEE Congress on Evolutionary Computation, IEEE CEC 2014, (Ed.: C. A. Coello, Coello), IEEE Press, Piscataway, NJ, 2014, 2692-2699, 2014b

S. A. Stanhope, J.M. Daida, (1+1) Genetic Algorithm Fitness Dynamics in a Changing Environment, In: P. J. Angeline, Z. Michalewicz, M. Schoenauer, X. Yao, A. Zalzala (eds.), Proc. Congress on Evolutionary Computation, IEEE CEC 1999, IEEE Press, Piscataway, NJ, 1851-1858, 1999

References

M. Shpak, G. P. Wagner. Asymmetry of configuration space induced by unequal crossover: implications for a mathematical theory of evolutionary innovation. Artificial Life 6: 25–43, 2000

B. M. R. Stadler, P. F. Stadler, G. P.Wagner, W. Fontana, The topology of the possible: Formal spaces underlying patterns of evolutionary change. J. Theor. Biology 213: 241–274, 2001.

C. O. Wilke, T. Martinetz, Adaptive walks on time-dependent fitness landscapes. Phys. Rev. E 60: 2154–2159, 1999.

S. Wright, The roles of mutation, inbreeding, crossbreeding and selection in evolution. In: D. F. Jones (ed.), Proc. of the Sixth International Congress on Genetics, 356–366, 1932.

S. Yang, X. Yao, Experimental study on population-based incremental learning algorithms for dynamic optimization problems. Soft Computing 9: 815–834, 2005.