



VERTEILTE SYSTEME

Thomas Weise, Stefan Niemczyk, Hendrik Skubch, Roland Reichle, Kurt Geihs
{weise|skubch|reichle|geihs}@vs.uni-kassel.de
Distributed Systems Group, FB16, University of Kassel
Wilhelmshöher Allee 73
D-34121 Kassel
Germany

A Tunable Model for Multi-Objective, Epistatic, Rugged, and Neutral Fitness Landscapes

GECCO 2008

Genetic and Evolutionary Computation Conference

Theory II Track, Roswell Room, 2008-07-16 11:30-11:55, Chair: Thomas Jansen

Renaissance Atlanta Hotel Downtown, 590 West Peachtree Street NW, Atlanta, Georgia 30308, USA

Contents

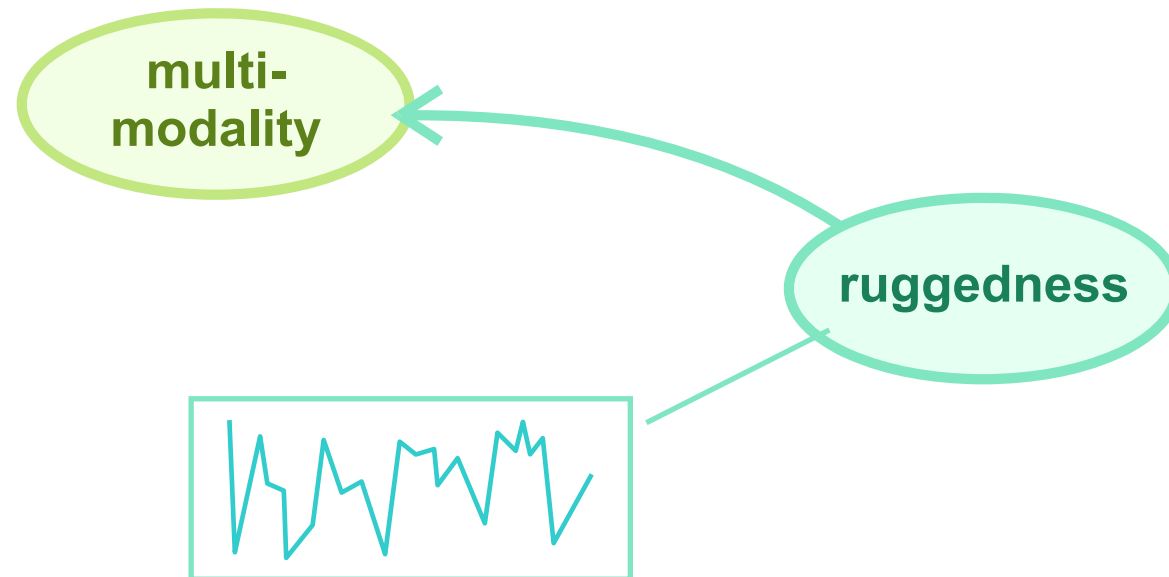
- Problems in Optimization
- Motivation
- Model Problem
 - Neutrality
 - Epistasis
 - Multi-Objectivity
 - Overfitting/Oversimplification
 - Ruggedness
- Filter Combinations
- Conclusions

Problems in Optimization

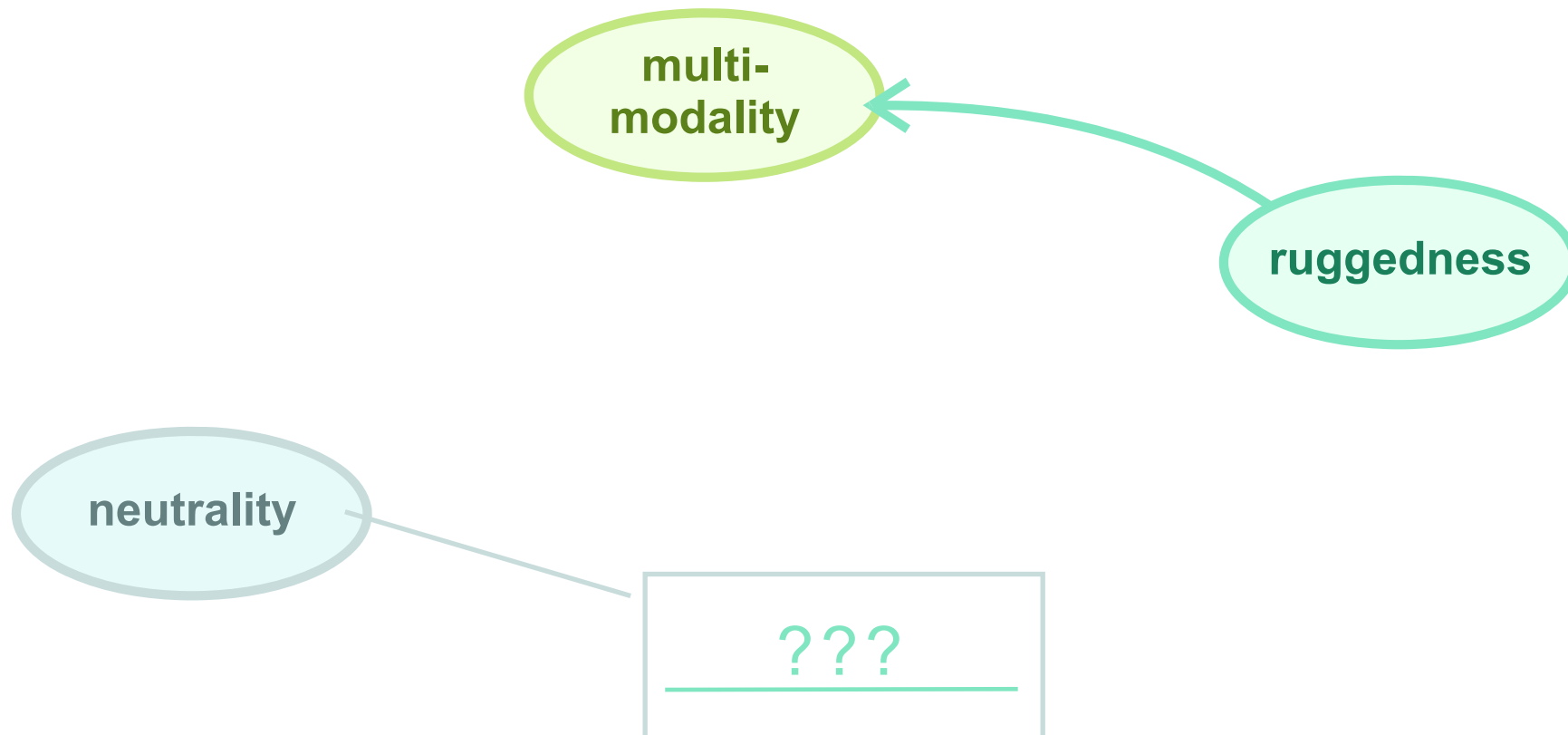
multi-
modality

multiple (local) optima,
danger of premature
convergence

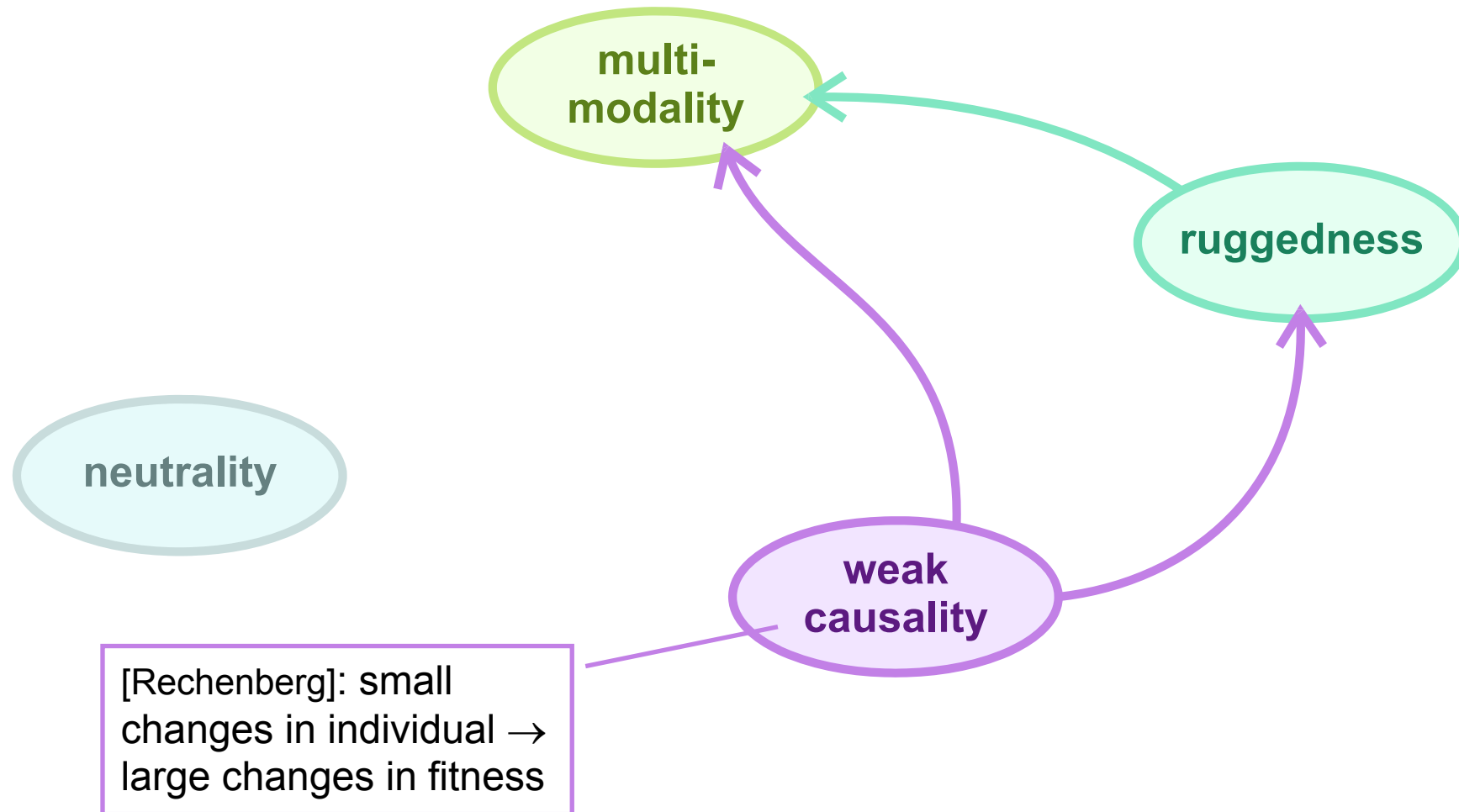
Problems in Optimization



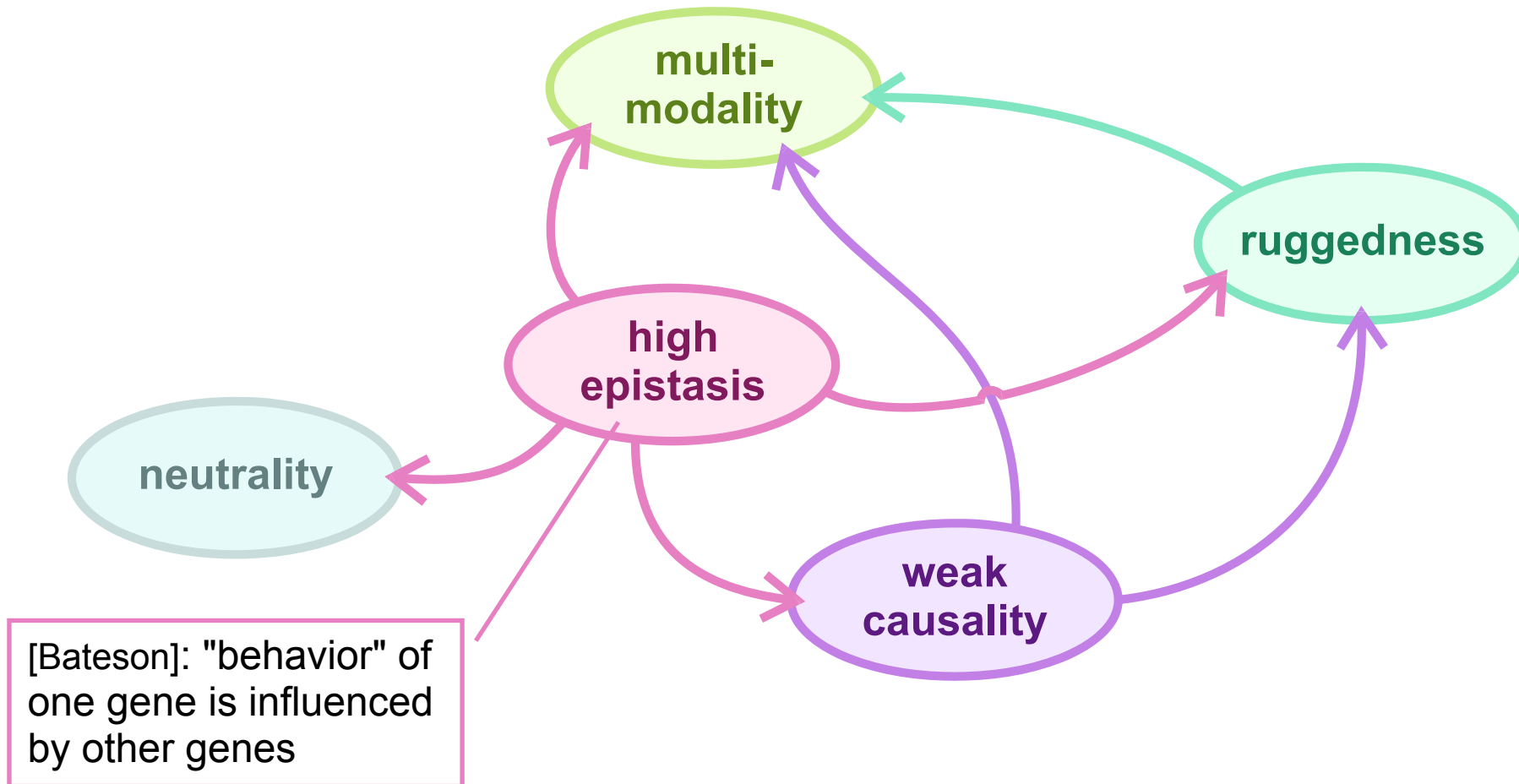
Problems in Optimization



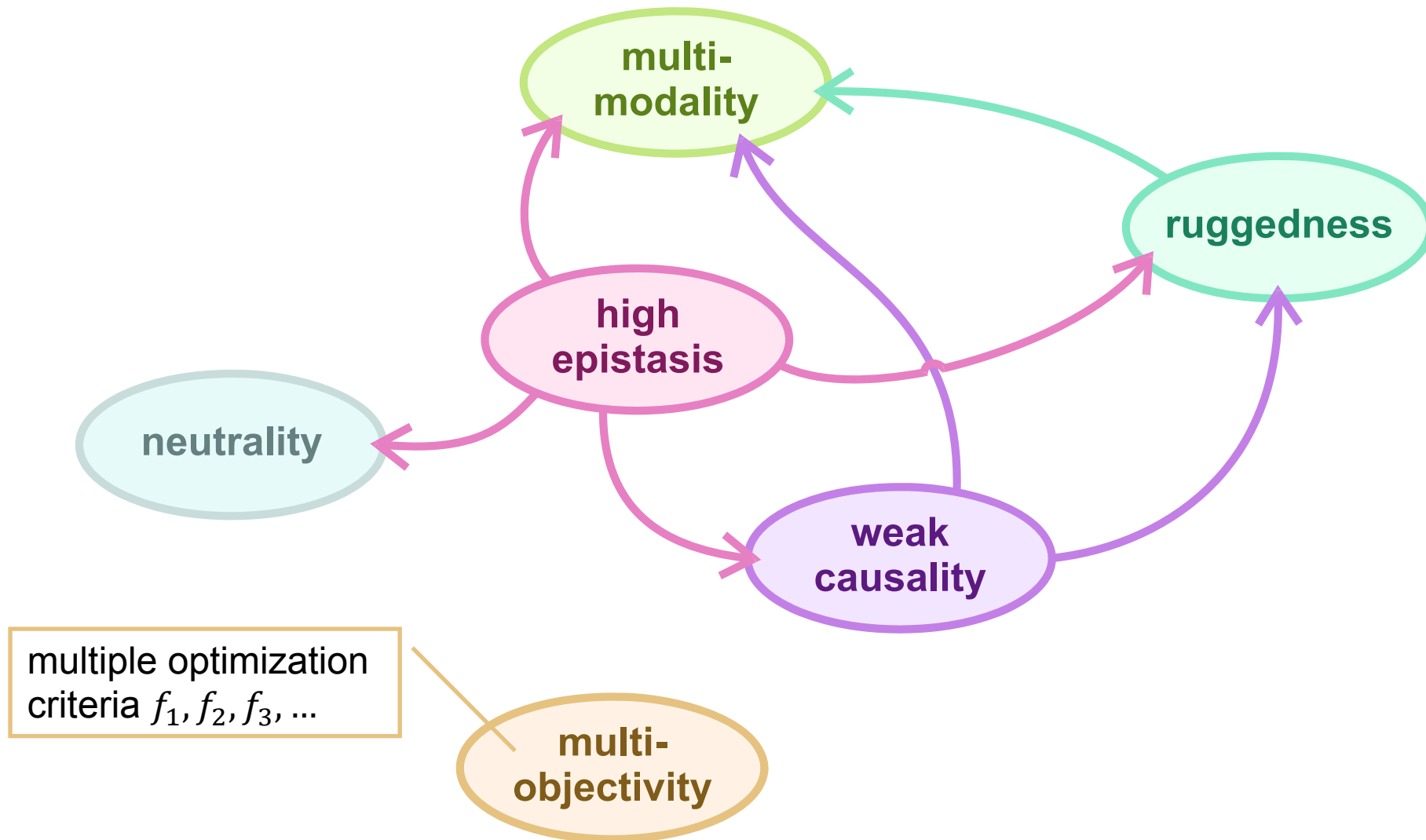
Problems in Optimization



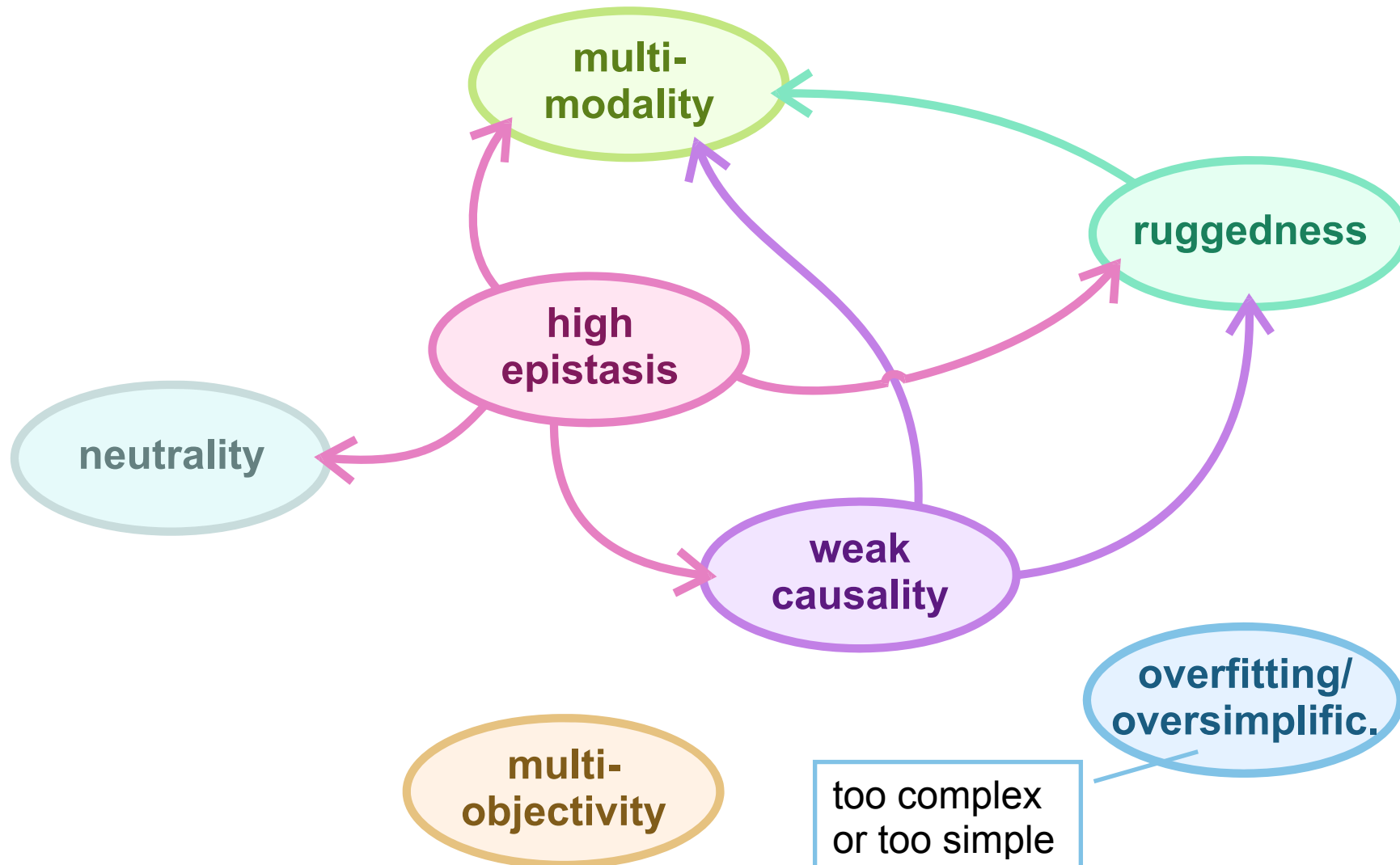
Problems in Optimization



Problems in Optimization



Problems in Optimization



Motivation

- optimization tasks exhibit these features to different degrees
- which settings of an EA perform best?
- experiments with "real" tasks often take very long
 - can test only very few settings
- want: model problem
 - fast
 - all features tunable

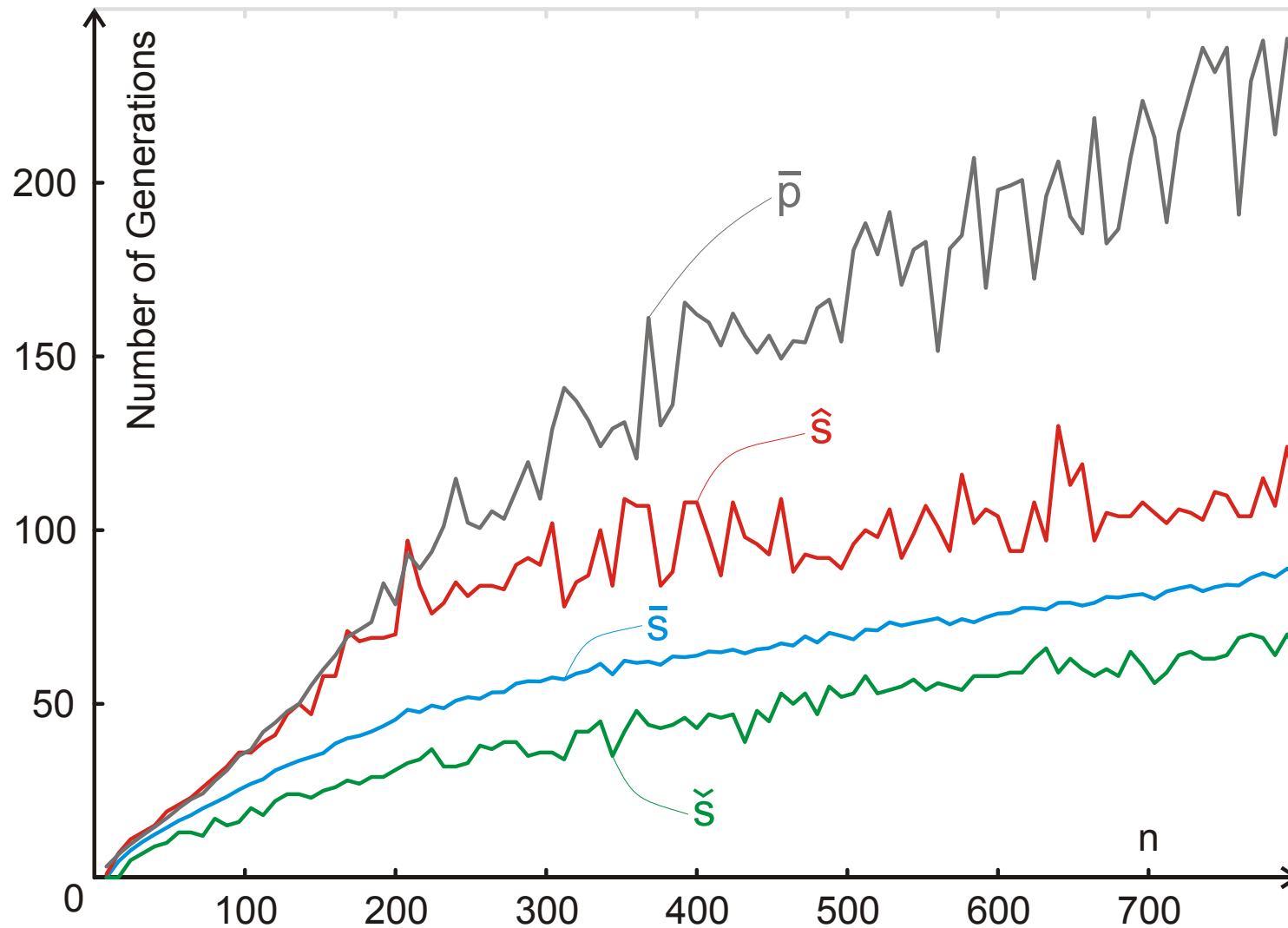
Model Problem – Basics

- each feature introduced as tunable filter
- search space: variable length bit strings (or fixed length)
- basic problem: find "optimal" bit string x^* of length n

$$x^* = 0101010101 \dots 01$$

- objective functions
 - Hamming distance (first n bits) to x^*
 - length

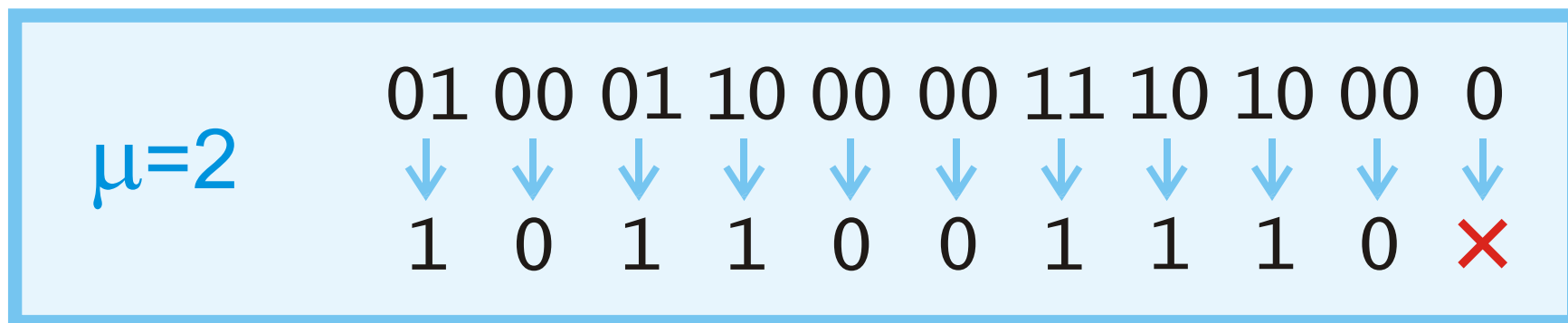
Model Problem – Basics



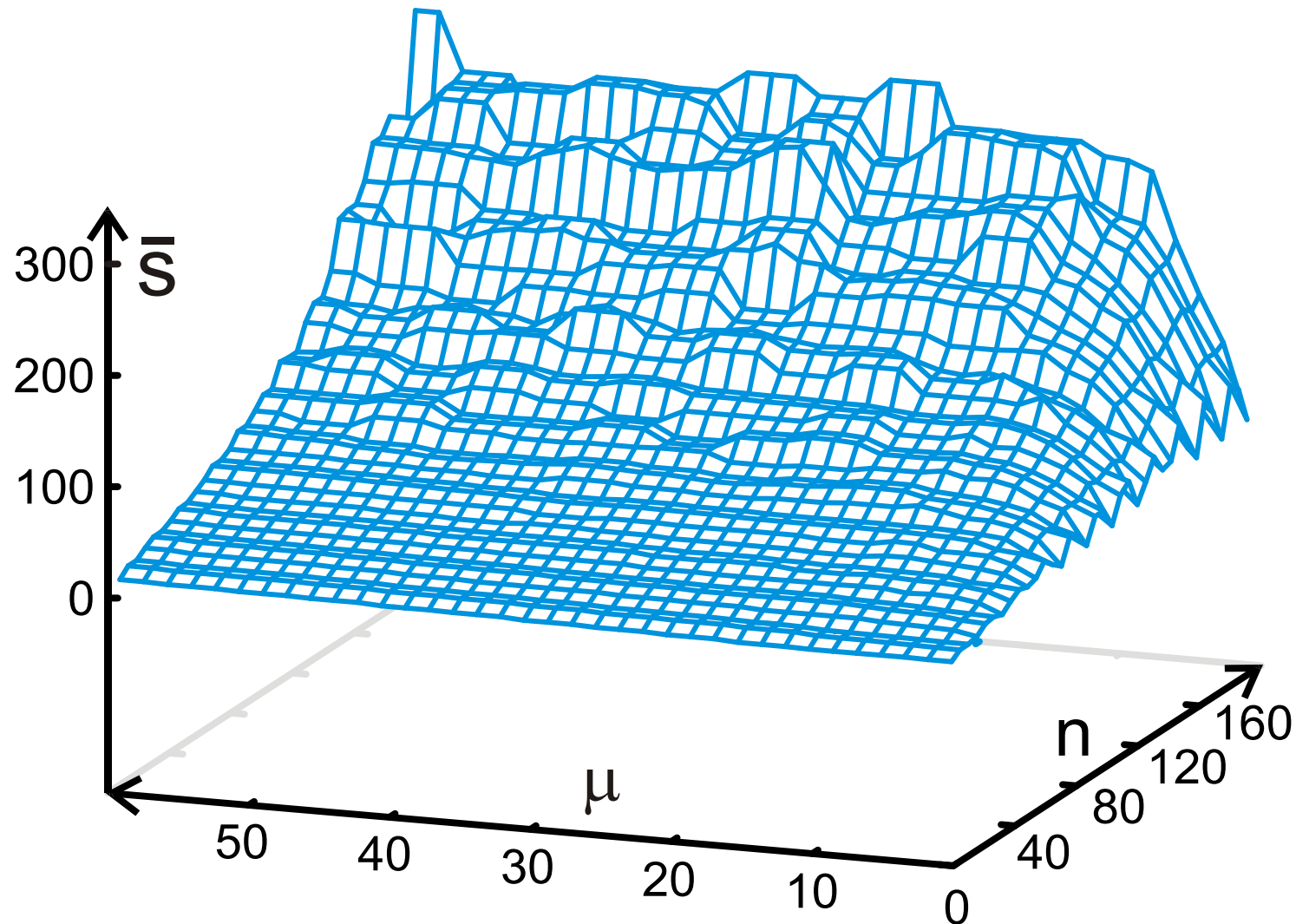
- population: 1000
- tournament selection (size 5)
- Pareto-ranking
- 1000 generation limit
- single-point crossover (80%)
- one-bit-mutation (20%)

Model Problem – Neutrality

- generate neutrality via redundancy
- combine blocks of μ bits to one bit via majority
- use 1 in draw situations
- genotype length l no multiple of $\mu \rightarrow$ ignore last $l \bmod \mu$ bits



Model Problem – Neutrality



Model Problem – Epistasis

- bijective mapping e_η of blocks z of the fixed length η
- modify the Hamming distance of the blocks
- $h(z_1, z_2) = 1 \Rightarrow h(e_\eta(z_1), e_\eta(z_2)) \geq \eta - 1 \quad \forall z_1, z_2 \in \{0,1\}^\eta$

	z	\rightarrow	$e_4(z)$		z	\rightarrow	$e_4(z)$		z	\rightarrow	$e_4(z)$	
} $h=1$	0000	\rightarrow	0000	} $h \geq 3$	1111	\rightarrow	1110		0011	\rightarrow	0110	
	0001	\rightarrow	1101			0111	\rightarrow	0001		0101	\rightarrow	1010
	0010	\rightarrow	1011			1011	\rightarrow	1001		0110	\rightarrow	1100
	0100	\rightarrow	0111			1101	\rightarrow	0101		1001	\rightarrow	0010
	1000	\rightarrow	1111			1110	\rightarrow	0011		1010	\rightarrow	0100
									1100	\rightarrow	1000	

Model Problem – Epistasis

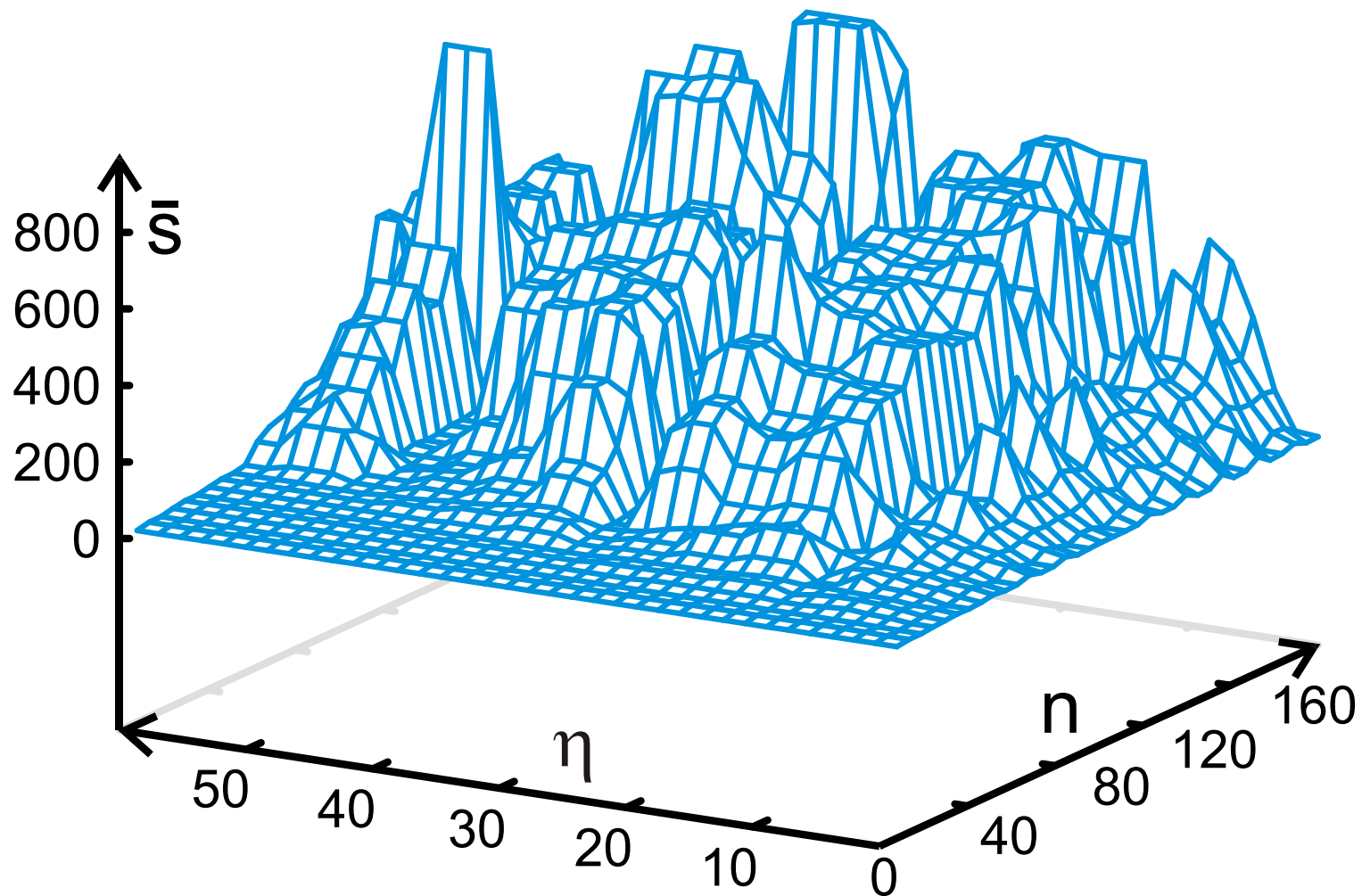
$$\bullet \quad e_{\eta}(z)_{[k]} = \begin{cases} \text{xor } z_{[i]} & \forall z: 0 \leq z < 2^{\eta-1} \\ \forall i: 0 \leq i < \eta, \\ i \neq (k-1)\% \eta \\ \hline e_{\eta}(z - 2^{\eta-1})_{[k]} & \textit{otherwise} \end{cases}$$

$\eta=4$

1011	0011	10
$e_4 \downarrow$	$e_4 \downarrow$	$e_2 \downarrow$
1001	0110	11

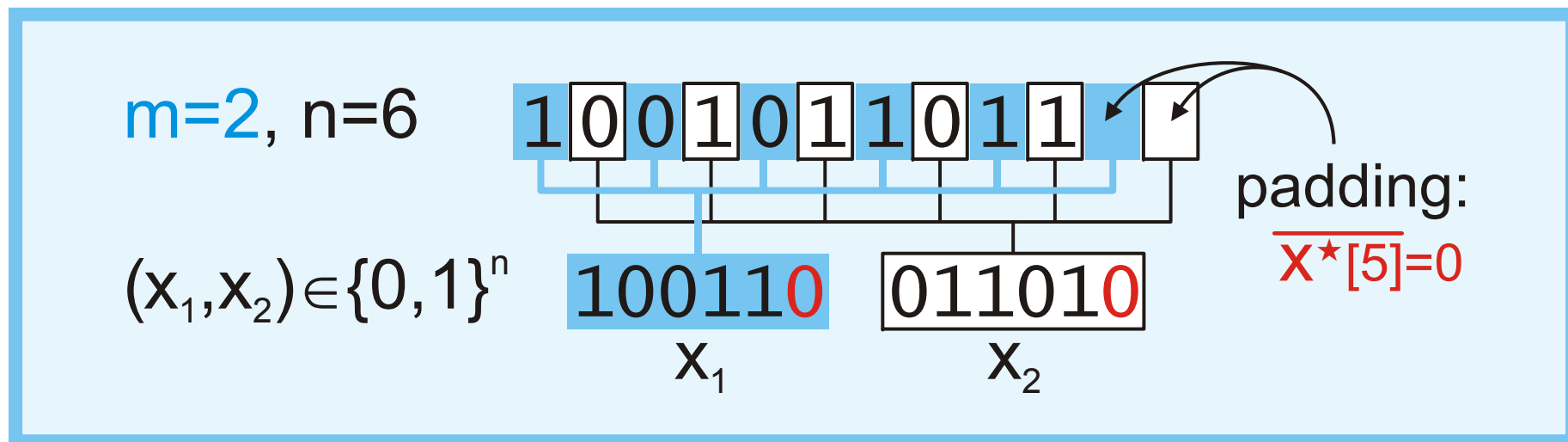
insufficient bits,
at the end, use
 $\eta=2$ instead of
 $\eta=4$

Model Problem – Epistasis



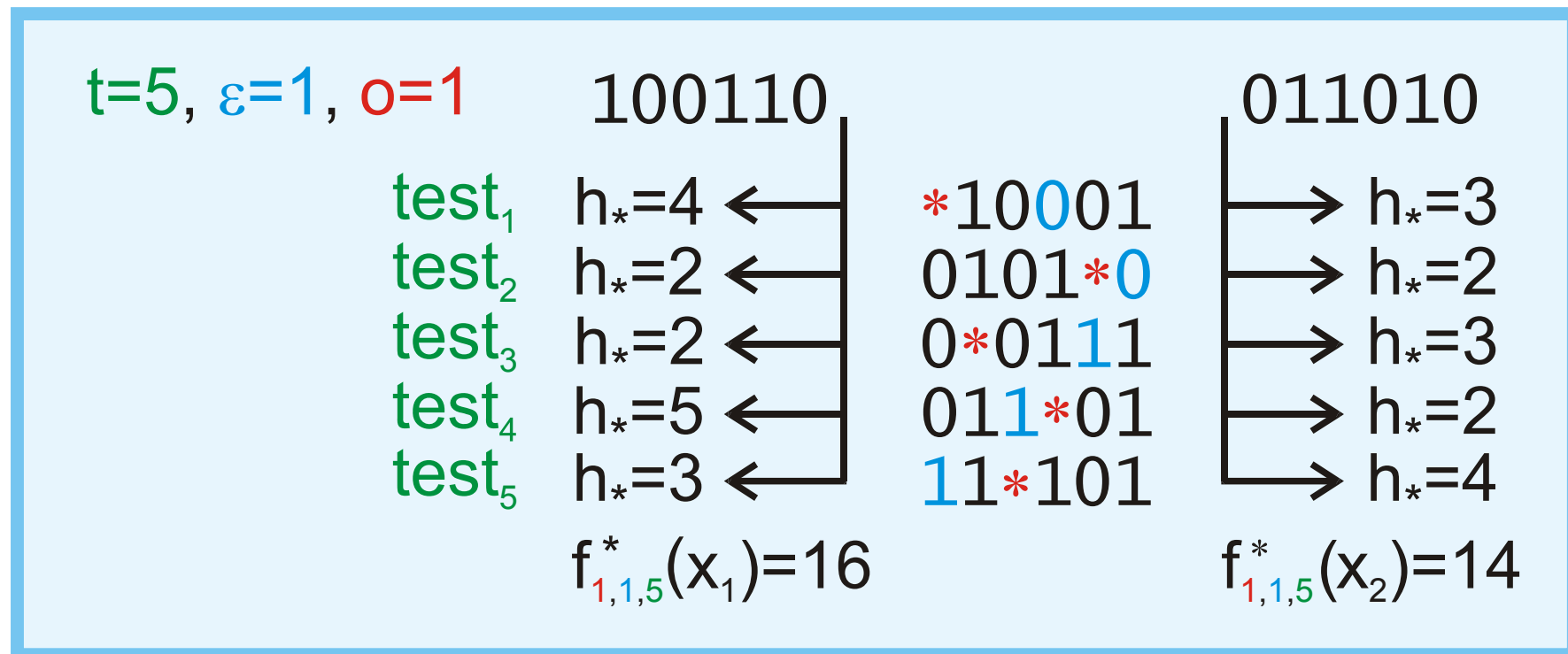
Model Problem – Multi-Objectivity

- using the functional objective function m times
- $x_i = (g_{i-1}, g_{m+i-1}, g_{2m+i-1}, \dots, g_{(n-1)m+i-1})$
- conflicts will ensue if epistasis is turned on
- non-functional objective: length of string (in the VL case)



Model Problem – Overfitting/~simplification

- instead of direct comparison with $x^* = (0101 \dots)$ use modified Hamming h_* distance to t test cases



Model Problem – Ruggedness

- introduce ruggedness via weak causality
- objective values: 0 is best, $q \approx (n - o)t$ is worst

0 ← 1 ← 2 ← 3 ← ... ← $q - 2$ ← $q - 1$ ← q

Model Problem – Ruggedness

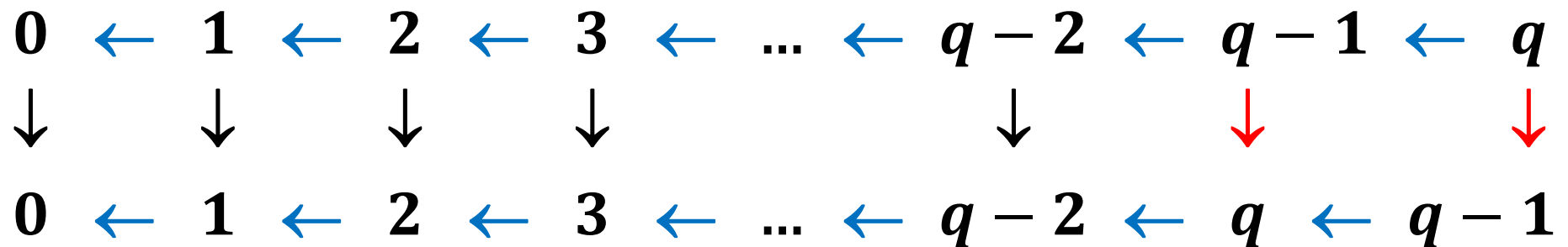
- introduce ruggedness via weak causality
- objective values: 0 is best, $q \approx (n - o)t$ is worst

$$\begin{array}{cccccccccccc}
 0 & \leftarrow & 1 & \leftarrow & 2 & \leftarrow & 3 & \leftarrow & \dots & \leftarrow & q - 2 & \leftarrow & q - 1 & \leftarrow & q \\
 & & 1 & + & 1 & + & 1 & + & 1 & + & 1 & + & 1 & + & 1
 \end{array}$$

- $\gamma = 0 \implies \Delta(r_\gamma) = q$

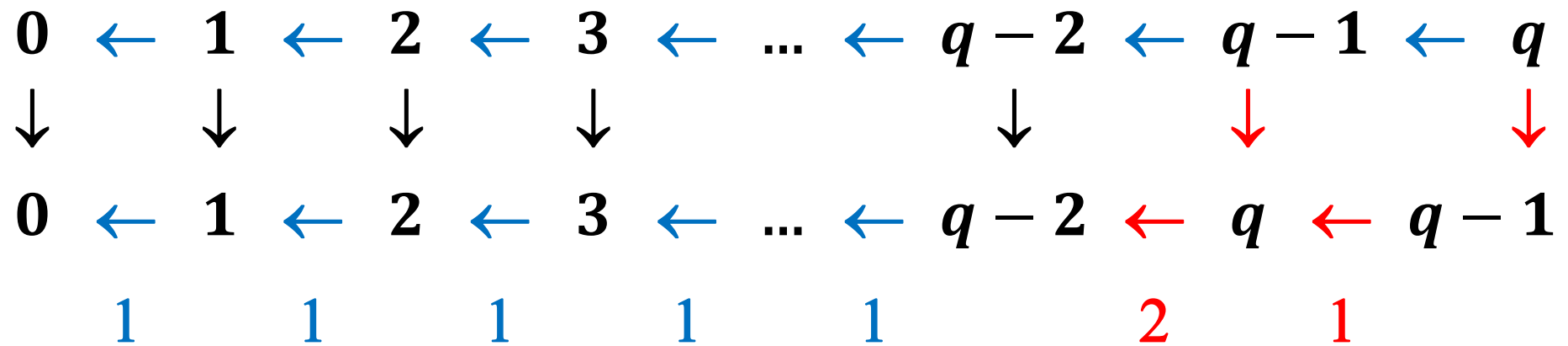
Model Problem – Ruggedness

- introduce ruggedness via weak causality
- objective values: 0 is best, $q \approx (n - o)t$ is worst



Model Problem – Ruggedness

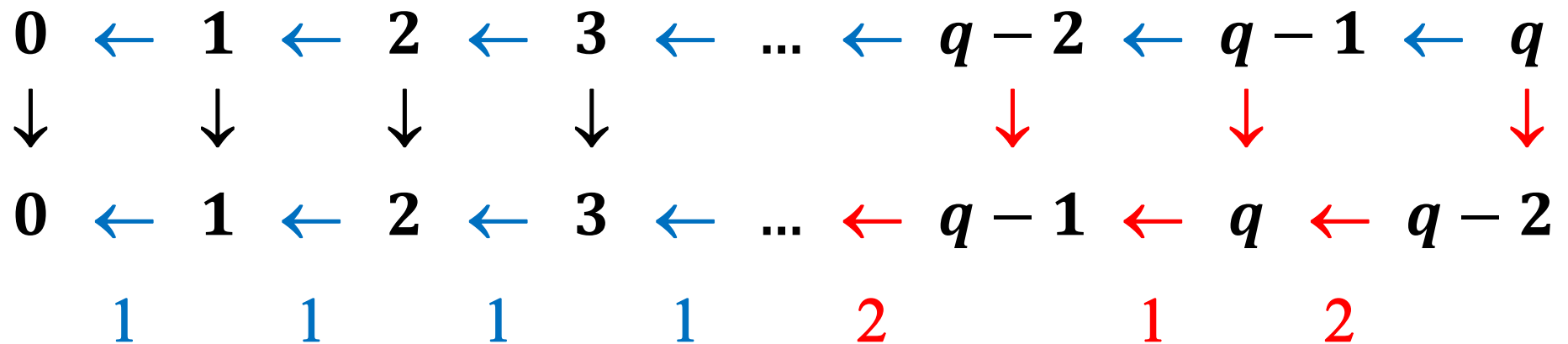
- introduce ruggedness via weak causality
- objective values: 0 is best, $q \approx (n - o)t$ is worst



- $\gamma = 1 \implies \Delta(r_\gamma) = q + 1$

Model Problem – Ruggedness

- introduce ruggedness via weak causality
- objective values: 0 is best, $q \approx (n - o)t$ is worst



- $\gamma = 2 \implies \Delta(r_\gamma) = q + 2$

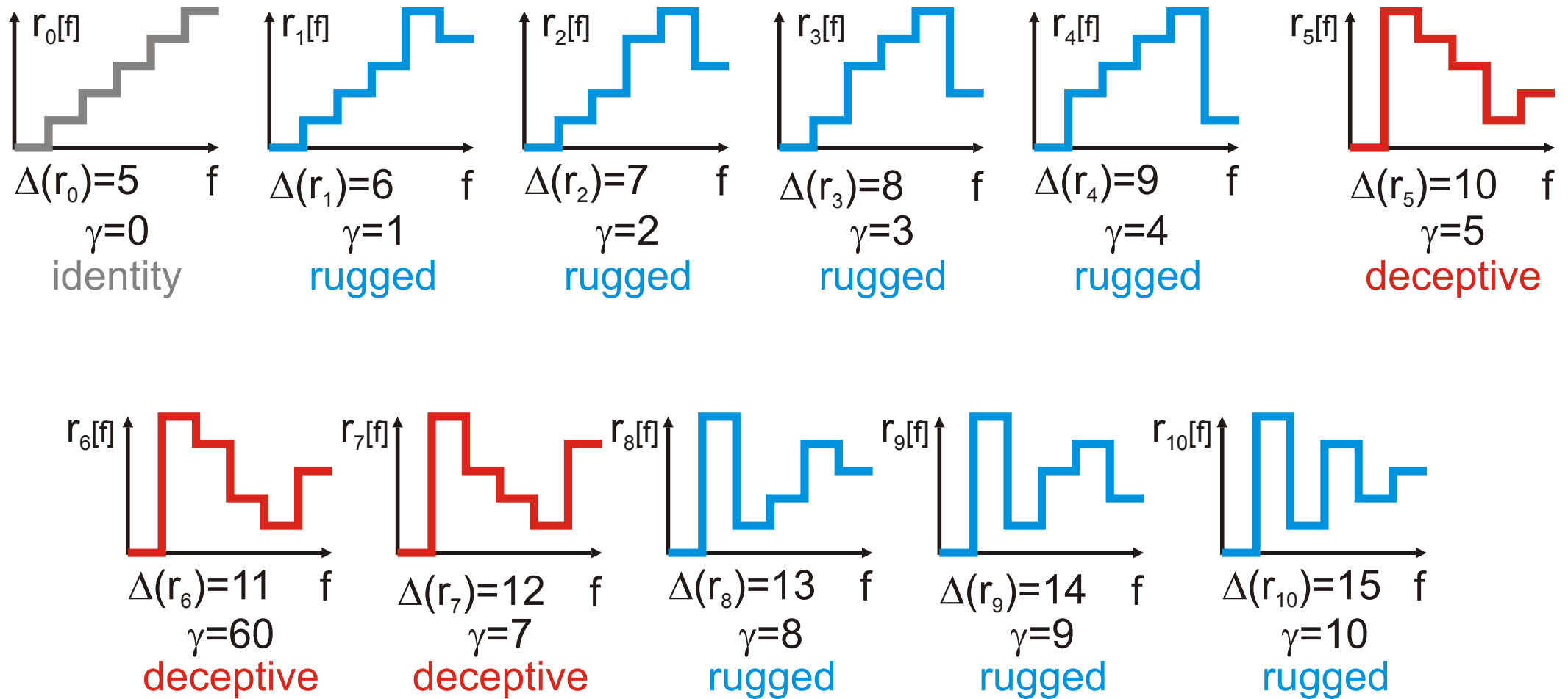
Model Problem – Ruggedness

- introduce ruggedness via weak causality
- objective values: 0 is best, $q \approx (n - o)t$ is worst

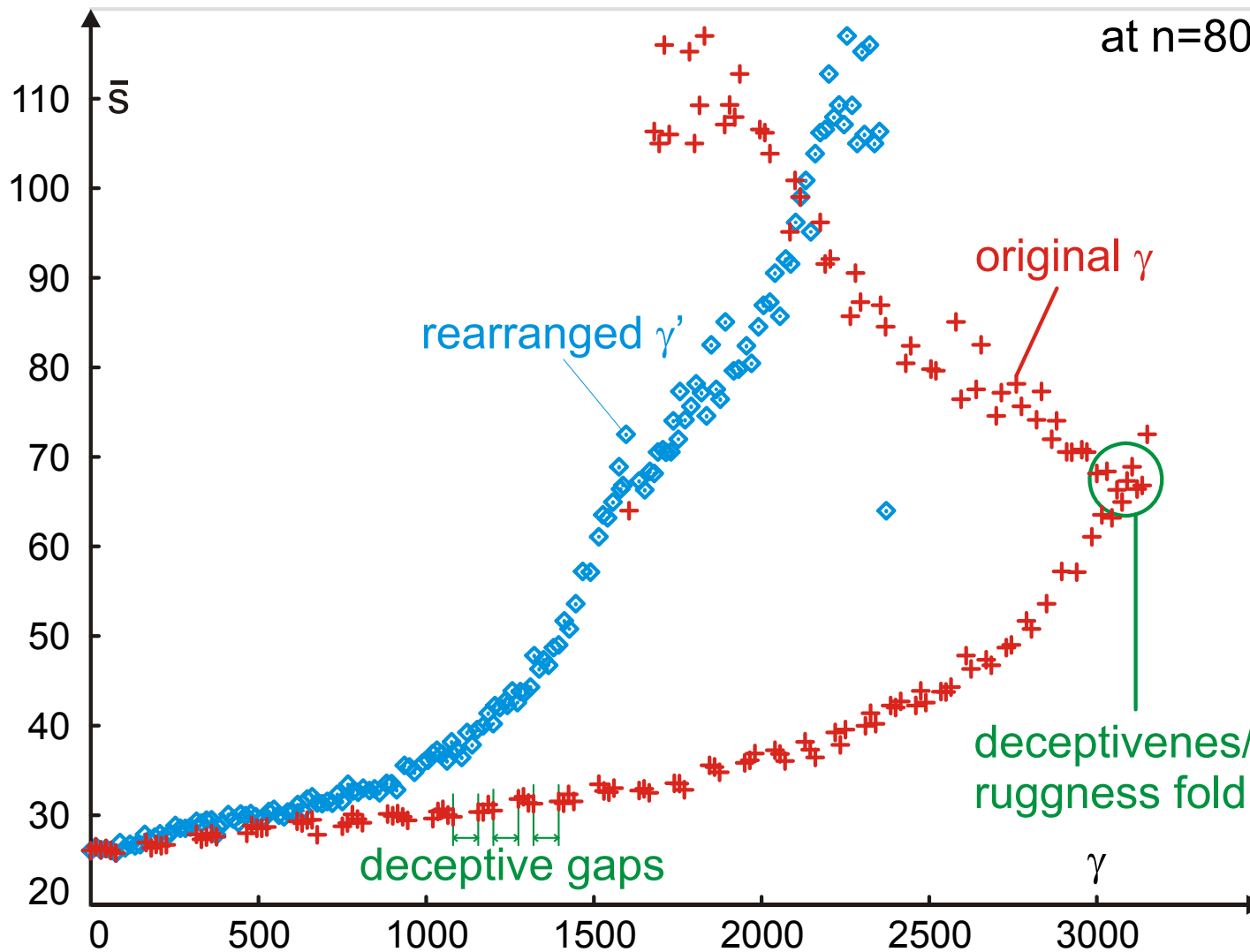


- $$\gamma = \frac{q(q-1)}{2} \implies \Delta(r_\gamma) = q + \frac{q(q-1)}{2} = \frac{q(q+1)}{2}$$

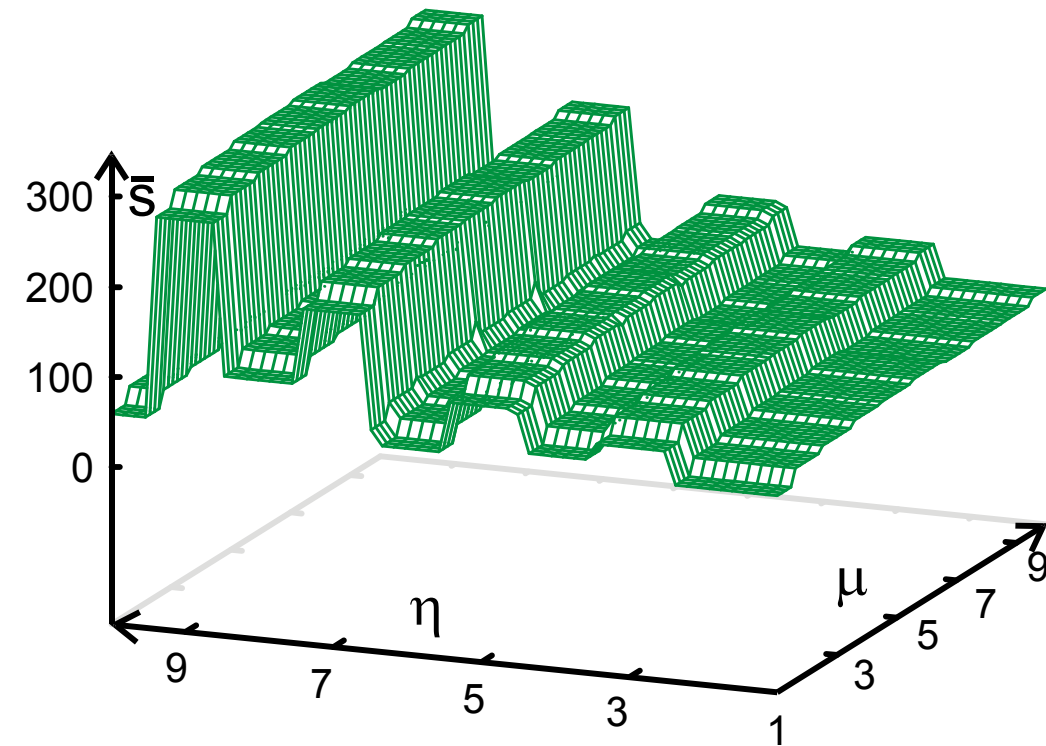
Model Problem – Ruggedness



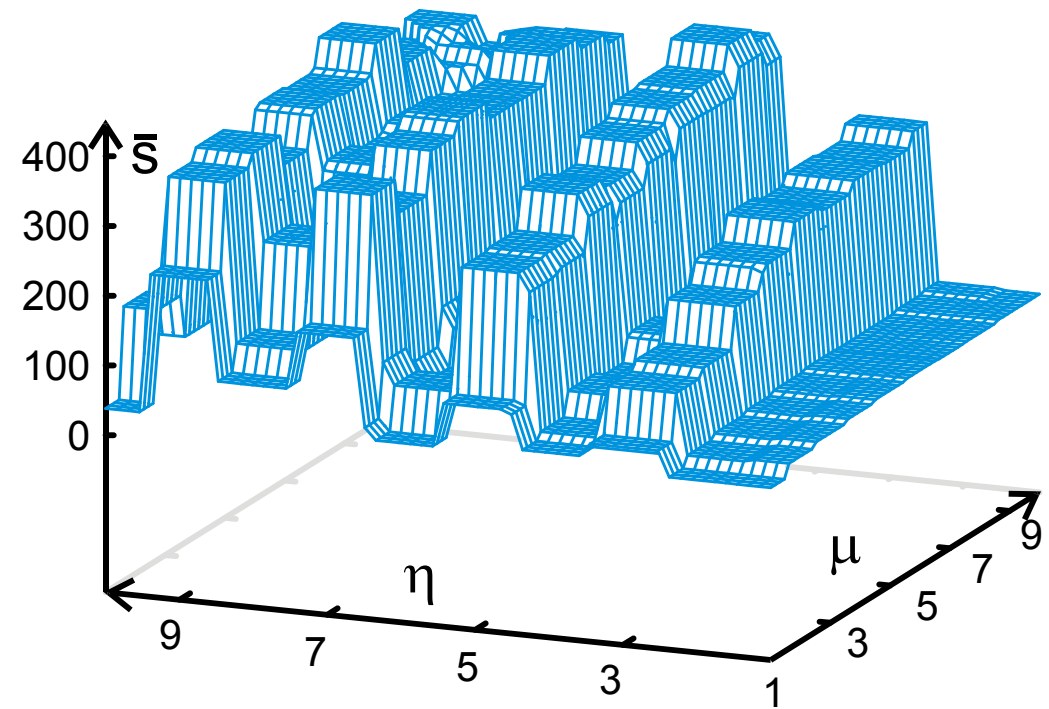
Model Problem – Ruggedness



Combinations – Epistasis/Neutrality

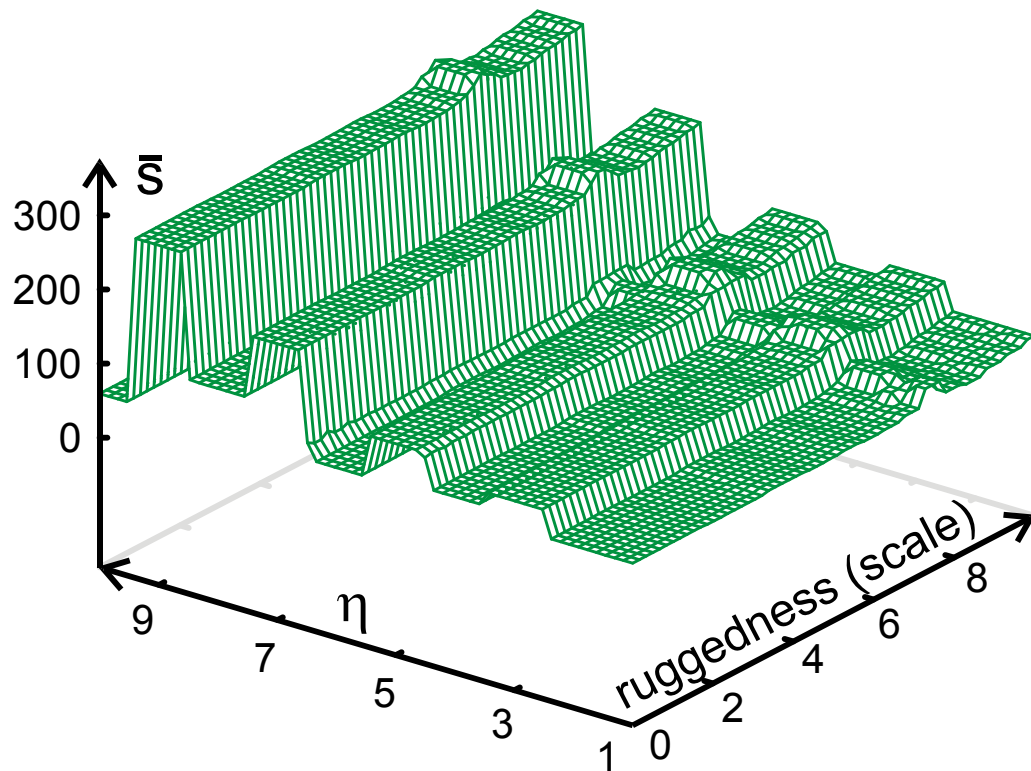


expected outcome

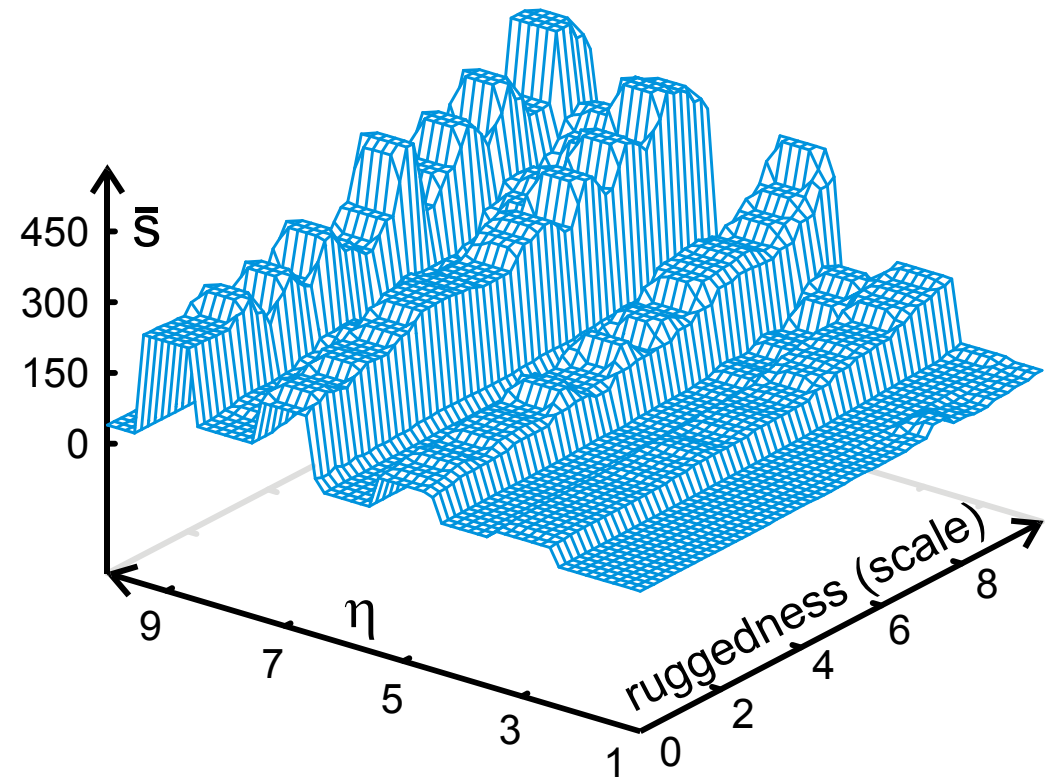


experimental outcome

Combinations – Epistasis/Ruggedness



expected outcome



experimental outcome

Conclusions

- model problem with tunable fitness landscape features
- additionally: scale from 0 to 10 for all features
- shown that the presented approaches are viable
- extremely fast evaluation ($ps=1000$, $0.1s < t/gen < 0.3s$,)
- improvements required, e.g. for epistasis/ruggedness
investigation on EA parameter influence on search
- Online available:
<http://www.it-weise.de/documents/files/TunableModel.java> and
<http://www.sigoa.org/documents/>

Thanks for your kind attention!

You may ask questions now.

<http://www.vs.uni-kassel.de/>
weise@vs.uni-kassel.de
<http://www.it-weise.de/>
tweise@gmx.de

Model Problem

Genotype

$g \in$ 010001100000111010000

Introduction of Neutrality

$\mu=2$ 01 00 01 10 00 00 11 10 10 00 0
 $u_2(g) \in$ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 1 0 1 1 0 0 1 1 1 0 **X**

Introduction of Epistasis

$\eta=4$ 1011 0011 10 insufficient bits,
 $e_4 \downarrow$ $e_4 \downarrow$ $e_2 \downarrow$ at the end, use
 $\eta=2$ instead of
 1001 0110 11 $\eta=4$

Multi-Objectivity and Phenotype

$m=2, n=6$ 1001011011 padding: $\overline{x^*[5]}=0$
 $(x_1, x_2) \in$ 100110 011010
 x_1 x_2

Introduction of Overfitting

$t=5, \epsilon=1, o=1$ 100110 011010
 test₁ $h_*=4 \leftarrow$ *10001 $\rightarrow h_*=3$
 test₂ $h_*=2 \leftarrow$ 0101*0 $\rightarrow h_*=2$
 test₃ $h_*=2 \leftarrow$ 0*0111 $\rightarrow h_*=3$
 test₄ $h_*=5 \leftarrow$ 011*01 $\rightarrow h_*=2$
 test₅ $h_*=3 \leftarrow$ 11*101 $\rightarrow h_*=4$
 $f_{1,1,5}^*(x_1)=16$ $f_{1,1,5}^*(x_2)=14$

Introduction of Ruggedness

$\gamma'=34$
 $\gamma=57, q=25$ $f_{1,1,5}(x_1)=16$ $f_{1,1,5}(x_2)=14$
 $(\Delta(r_{\gamma=57}))=82$ ↓ ↓
 $r_{57}[f_{1,1,5}(x_1)]=17$ $r_{57}[f_{1,1,5}(x_2)]=15$

Related Work I

- Kauffman's NK fitness landscapes (Rugged/Epistatic)
- N bits, each contributes to total fitness (sum of contribs)
- each contribution influenced by K neighboring genes
- S.A. Kauffman. *Adaptation on rugged fitness landscapes*. Lectures in the Sciences of Complexity: 1988 Compl. Sys. Summer School, ISBN 978-0201510157, 0201510154.
- S.A. Kauffman and S. Asher. *Towards a general theory of adaptive walks on rugged landscapes*. Journal of Theoretical Biology, 128(1):11–45, 1987
- S.A. Kauffman and E.D. Weinberger. *The NK model of rugged fitness landscapes and its application to maturation of the immune response*. Journal of Theoretical Biology, 141:211–245, 1989. [http://dx.doi.org/10.1016/S0022-5193\(89\)80019-0](http://dx.doi.org/10.1016/S0022-5193(89)80019-0)
- L. Altenberg. NK Fitness Landscapes. In Handbook of Evolutionary Computation, 1997. ISBN 0750303921. <http://citeseer.ist.psu.edu/704814.html> and <http://dynamics.org/Altenberg/FILES/LeeNKFL.pdf>

Related Work II

- NKp landscape: add neutrality to NK landscape by setting a fraction p of contribution to zero
- L. Barnett. *Ruggedness and neutrality – the NKp family of fitness landscapes*. Artificial Life VI: Proceedings of the sixth international conference on Artificial life, 1998, ISBN 0-262-51099-5. See <http://citeseer.ist.psu.edu/barnett98ruggedness.html> and ftp://ftp.informatics.sussex.ac.uk/pub/users/lionelb/publications/alife6_paper.pdf
- NKq landscape: add neutrality to NK landscape by discretizing contributions into interval $[0, q) : q \geq 1$
- M. E. J. Newman and R. Engelhardt. *Effect of neutral selection on the evolution of molecular species*. Proceedings of the Royal Society of London B, 256(1403):1333–1338, 1998. Online available at <http://citeseer.ist.psu.edu/202669.html> and <http://journals.royalsociety.org/content/5ttwagyubu88bang/fulltext.pdf>

Related Work III

- Technological landscape: add neutrality to NK landscape by discretizing total fitness into M bins
- J. Lobo, J.H. Miller, and W. Fontana. *Neutrality in technological landscapes*. Santa Fe working paper, 2004. Online available at <http://citeseer.ist.psu.edu/lobo04neutrality.html> and <http://fontana.med.harvard.edu/www/Documents/WF/Papers/tech.neut.pdf>
- p-Spin model: complete evolutionary model for epistasis and ruggedness with selection/mutation, N bits, **all** sets of K genes contribute to total fitness
- C. Amitrano, L. Peliti, and M. Saber. Population dynamics in a spin-glass model of chemical evolution. *Journal of Molecular Evolution*, 29(6):513–525, 1989. Online available at <http://www.springerlink.com/content/n172u715884w0632/fulltext.pdf>

Related Work IV

- ND fitness landscape:
 - neutrality: N bits, D is distribution of number of neutral neighbors of genotypes
 - deceptiveness: fitness of neutral networks set according to Trap Function
- W. Beaudoin, S. Verel, P. Collard, and C. Escazut. *Deceptiveness and neutrality the ND family of fitness landscapes*. In GECCO '06: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, 2006, ACM Press, ISBN 1-59593-186-4. Online available at <http://doi.acm.org/10.1145/1143997.1144091>
- K. Deb and D.E. Goldberg. *Analyzing deception in trap functions*. In Proceedings of the Second Workshop on Foundations of Genetic Algorithms, 1992, Morgan Kaufmann, ISBN 1-55860-263-1

Related Work V

- Royal Road: search for schemas of doubling order
- neutrality via don't cares
- M. Mitchell, S. Forrest, and J.H. Holland. The Royal Road for Genetic Algorithms: Fitness Landscapes and GA Performance. In *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life, 1991*, ISBN 0-262-72019-1. See <http://citeseer.ist.psu.edu/mitchell91royal.html> and <http://web.cecs.pdx.edu/~mm/ecal92.pdf>

Ruggedness-generating Algorithms

Algorithm 19.4: $r_\gamma \leftarrow \text{buildRPermutation}(\gamma, \hat{f})$

Input: γ : the γ value

Input: \hat{f} : the maximum objective value

Data: i, j, d, tmp : temporary variables

Data: $k, start, r$: parameters of the subalgorithm

Output: r_γ : the permutation r_γ

```

1 begin
2   Subalgorithm  $r \leftarrow \text{permutate}(k, r, start)$ 
3   begin
4     if  $k > 0$  then
5       if  $k \leq (\hat{f} - 1)$  then
6          $r \leftarrow \text{permutate}(k - 1, r, start)$ 
7          $tmp \leftarrow r[\hat{f}]$ 
8          $r[\hat{f}] \leftarrow r[\hat{f} - k]$ 
9          $r[\hat{f} - k] \leftarrow tmp$ 
10      else
11         $i \leftarrow \lfloor \frac{start + 1}{2} \rfloor$ 
12        if  $(start \bmod 2) = 0$  then
13           $i \leftarrow \hat{f} + 1 - i$ 
14           $d \leftarrow -1$ 
15        else
16           $d \leftarrow 1$ 
17        for  $j \leftarrow start$  up to  $\hat{f}$  do
18           $r[j] \leftarrow i$ 
19           $i \leftarrow i + d$ 
20         $r \leftarrow \text{permutate}(k - \hat{f} + start, r, start + 1)$ 
21      end
22     $r \leftarrow (0, 1, 2, \dots, \hat{f} - 1, \hat{f})$ 
23    return  $\text{permutate}(\gamma, r, 1)$ 
24  end
```

$\gamma \leftarrow \text{translate}(\gamma', \hat{f})$

Input: γ' : the raw γ value

Input: \hat{f} : the maximum value of $f_{\varepsilon, o, t}$

Data: i, j, k, l : some temporary variables

Output: γ : the translated γ value

```

1 begin
2    $l \leftarrow \frac{\hat{f}(\hat{f} - 1)}{2}$ 
3    $i \leftarrow \lfloor \frac{\hat{f}}{2} \rfloor * \lfloor \frac{\hat{f} + 1}{2} \rfloor$ 
4   if  $\gamma \leq \hat{f}i$  then
5      $j \leftarrow \lfloor \frac{\hat{f} + 2}{2} - \sqrt{\frac{\hat{f}^2}{4} + 1 - \gamma} \rfloor$ 
6      $k \leftarrow \gamma - j(\hat{f} + 2) + j^2 + \hat{f}$ 
7     return  $k + 2(j(\hat{f} + 2) - j^2 - \hat{f}) - j$ 
8   else
9      $j \leftarrow \lfloor \frac{(\hat{f} \bmod 2) + 1}{2} + \sqrt{\frac{1 - (\hat{f} \bmod 2)}{4} + \gamma - 1 - i} \rfloor$ 
10     $k \leftarrow \gamma - (j - (\hat{f} \bmod 2))(j - 1) - 1 - i$ 
11    return  $l - k - 2j^2 + j - (\hat{f} \bmod 2)(-2j + 1)$ 
12  end
```

Neutrality

- Neutrality can have positive and negative effects
- Negative:
 - uniform redundancy
 - neutral networks with very high proportion of neutral search operation outcomes
- Beaudoin et al. *Deceptiveness and neutrality the ND family of fitness landscapes.*
- Positive:
 - neutral networks with moderate proportion of neutral search operations
- M. Shackleton, R. Shipman, and M. Ebner. An investigation of redundant genotype-phenotype mappings and their role in evolutionary search. Congress on Evolutionary Computation, pages 493–500, 2000. <http://citeseer.ist.psu.edu/409243.html>
- M. Toussaint and C. Igel. Neutrality: A Necessity for Self-Adaptation. Congress on Evolutionary Computation, pages 1354–1359, 2002.

Strange Epistasis Behavior

- $$e_{\eta}(z)_{[k]} = \begin{cases} \text{xor } z_{[i]} & \forall z: 0 \leq z < 2^{\eta-1} \\ \forall i: 0 \leq i < \eta, \\ i \neq (k-1)\% \eta \\ \hline e_{\eta}(z - 2^{\eta-1})_{[k]} & \text{otherwise} \end{cases}$$
- e_{η} consists of two parts (second = negated first one)
- otherwise not bijective for even η
- mappings with $\eta = 2k + 2, k \in \mathbb{N}$ are much simpler
- then, there exist many pairs with

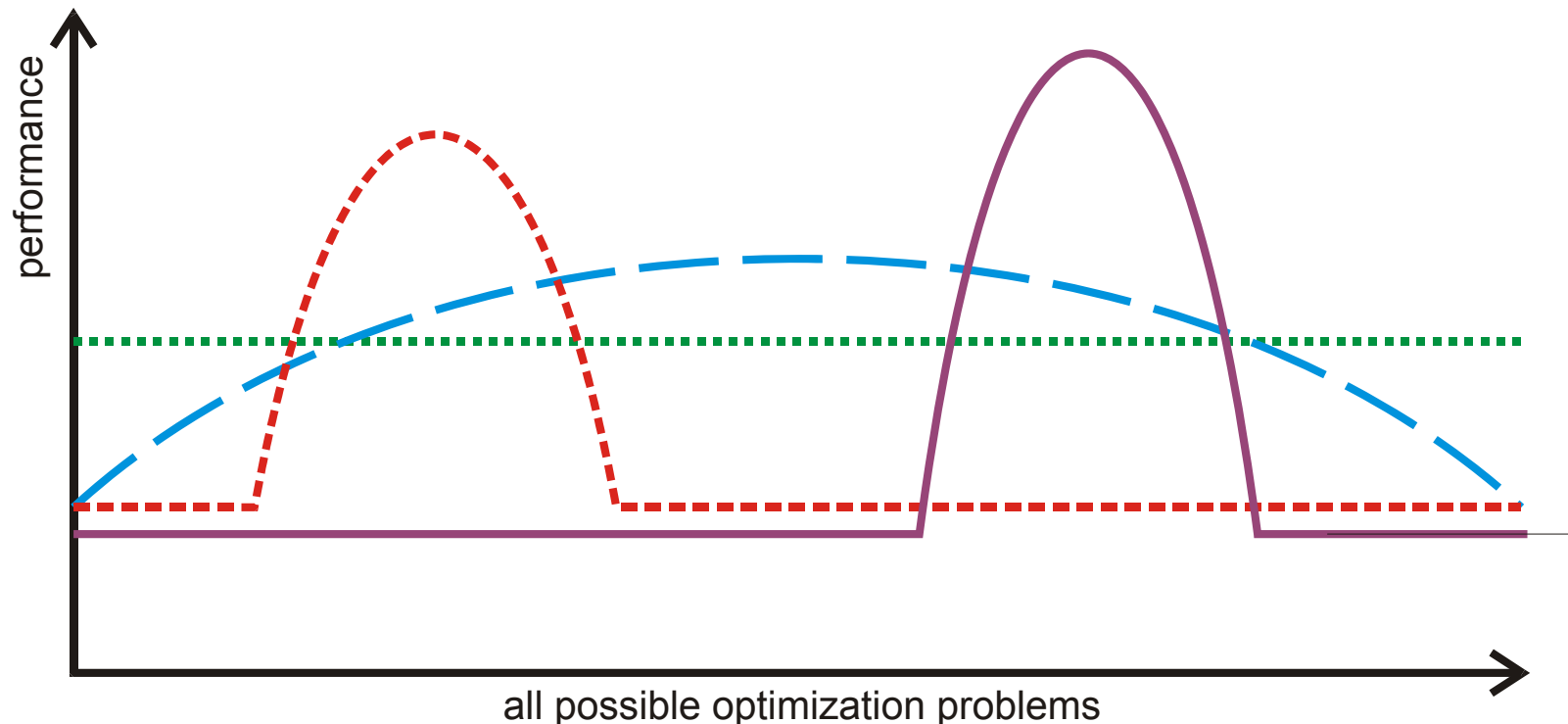
$$\left[h(z_1, z_2) = \frac{1}{2}\eta \right] \wedge \left[h(e_{\eta}(z_1), e_{\eta}(z_2)) \neq \frac{1}{2}\eta \right]$$

Epistasis vs. Ruggedness

- Epistasis
 - explicit dependency between *groups of genes*
 - one configuration of the group
 - ⇒ one distinct fitness contribution (tendency)
- Ruggedness permutations
 - influence independent from gene groups ($< n$), assume
 - + basic problem ($n = 10$) + ruggedness
 - + individual with: 6 bits wrong, 4 bits correct
 - + toggling *any* correct bit ⇒ the same result x
 - + toggling *any* wrong bit ⇒ same result y
- **Either** no epistasis **or** weak epistasis involving all n genes...

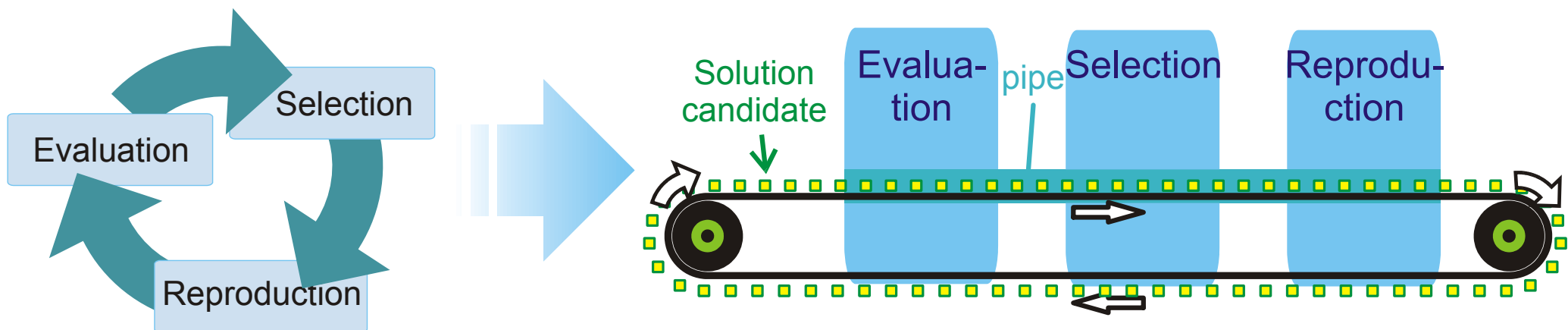
No Free Lunch (NFL)

- All optimization algorithms have the same performance averaged over all possible optimization problems.
- D.H. Wolpert and W.G. Macready. No Free Lunch Theorems for Optimization. IEEE Trans. on Ev. Comp., 1(1):67–82, 1997. <http://citeseer.ist.psu.edu/wolpert96no.html>



σgoa

- Simple Interface for Global Optimization
- versatile Java framework for global optimization
- especially suitable for problems involving complex simulations
- pipes and filters principle
- open source, <http://www.sigoa.org/>



Thomas Weise, Stefan Niemczyk, Hendrik Skubch,
Roland Reichle, and Kurt Geihs
{weise|skubch|reichle|geihs}@vs.uni-kassel.de

Distributed Systems Group, FB16, University of Kassel
Wilhelmshöher Allee 73
D-34121 Kassel
Germany