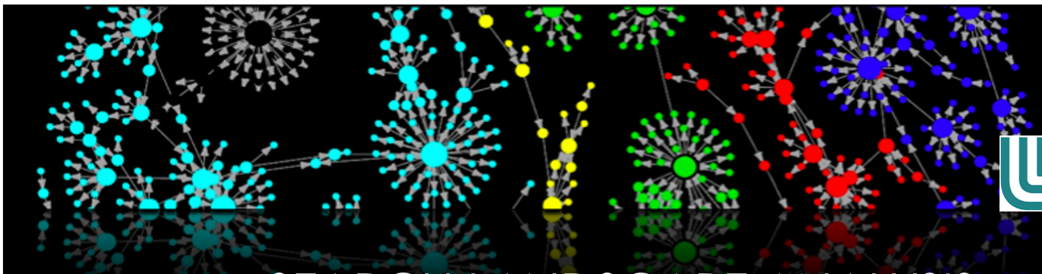


SSAAD 2023

Université de Lille

SEARCH LANDSCAPE ANALYSIS

PART 1 – INTRODUCTION TO LANDSCAPE ANALYSIS

Gabriela Ochoa

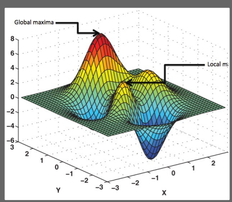
Thanks To
Katherine M. Malan
Nadarajen Veerapen

UNIVERSITY of STIRLING

UNIVERSITY of STIRLING

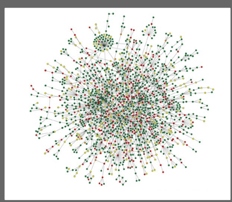
1

OUTLINE



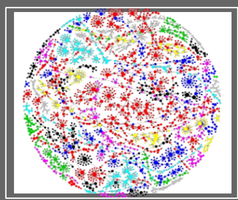
1932, ~2000

Landscape Analysis



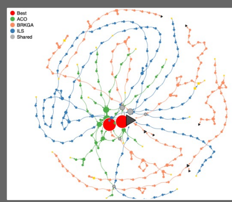
~2000

Complex Networks



2008

Local Optima Networks (LONs)

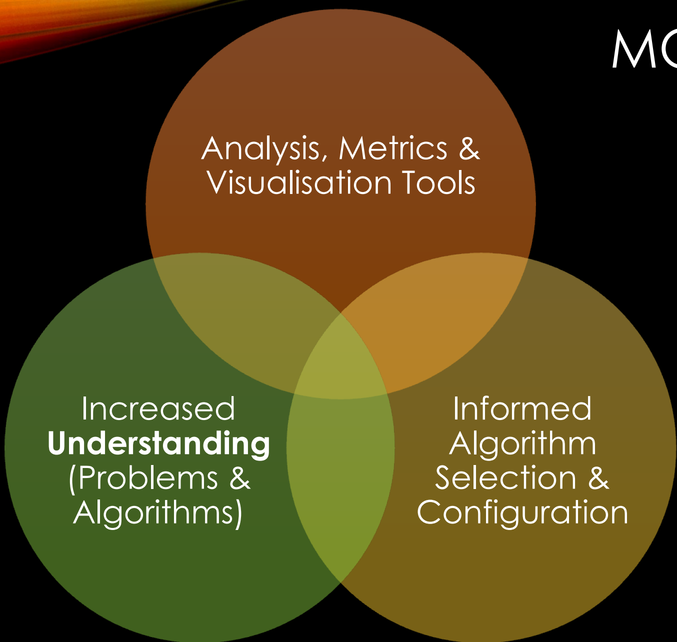


2020

Search Trajectory Networks (STNs)

2

MOTIVATION



3

ORIGINS OF FITNESS LANDSCAPES

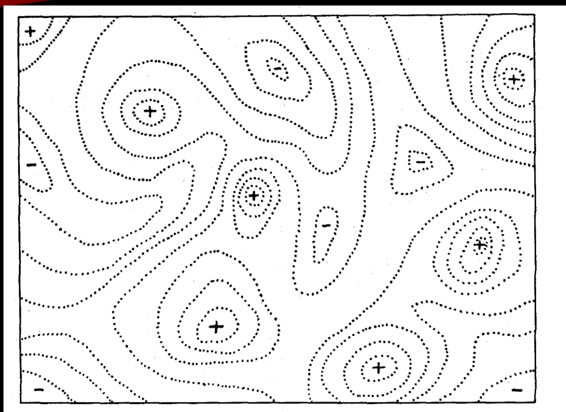


FIGURE 2.—Diagrammatic representation of the field of gene combinations in two dimensions instead of many thousands. Dotted lines represent contours with respect to adaptiveness.

Sewell Wright
(1889 – 1988)
American
geneticist.

"In a rugged field of this character, selection will easily carry the species to the nearest peak, but there may be innumerable other peaks which are higher but which are separated by "valleys." The problem of evolution as I see it is that of a mechanism by which the species may continually find its way from lower to higher peaks in such a field. "

S Wright (1932) The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution: Proc. of the Sixth International Congress on Genetics

4

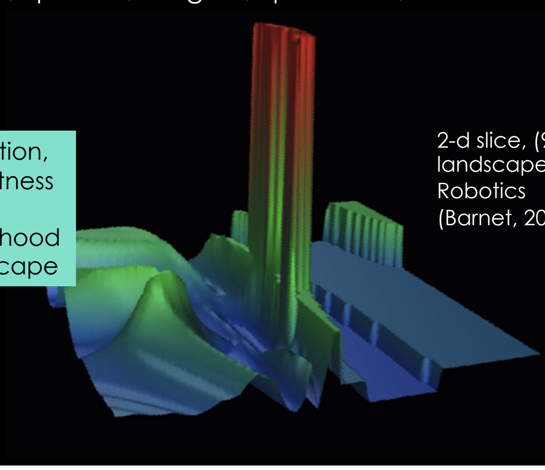
FITNESS LANDSCAPES TODAY

Mathematical Object (S, N, f)
Weinberger (1990),
Jones (1995), Stadler (2002)

- Search Space
- Neighbourhood Structure
- Fitness Function

Metaphor Visualisation of the terrain capturing how fitness changes between neighbouring solutions.
"valleys", "peaks", "ridges", "plateaus", "funnels"

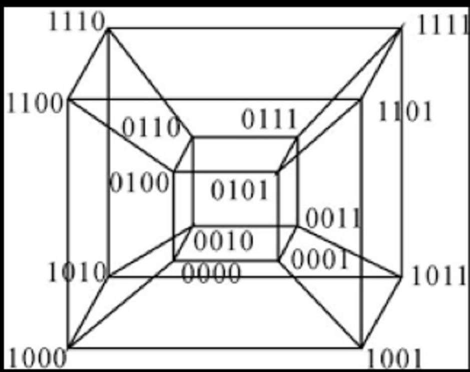
One fitness function, many possible fitness landscapes!
Each neighbourhood induces a landscape



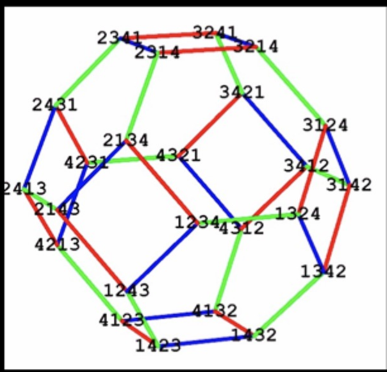
2-d slice, (9-d) landscape Evol. Robotics (Barnet, 2002)

5

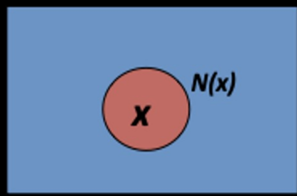
SEARCH SPACES & NEIGHBOURHOODS



Binary strings with 1-flip neighbourhood - Hamming distance



Permutations with swap neighbourhood
Swap distance, bond distance



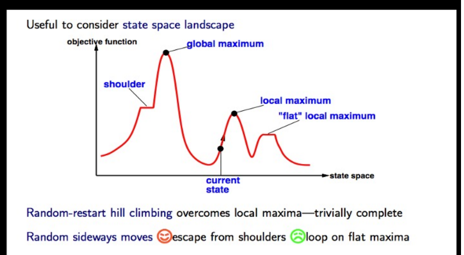
Vector of real numbers with Gaussian or uniform moves
Euclidean distance

Neighbourhood Region of the search space that is “near” to some particular point in that space. Defined in terms of a **move** operator or **distance metric**.

6

FEATURES OF LANDSCAPES

- Number, fitness, and distribution of local optima or peaks
- Distribution of fitness values (density of states)
- Fitness differences between neighbouring points (ruggedness)



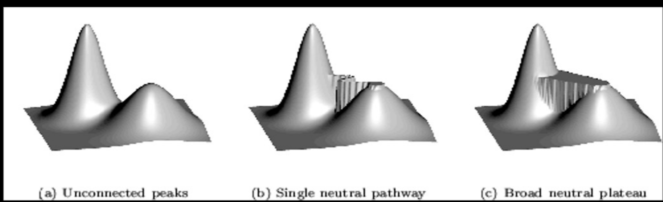
M. Fuji, Japan

Grandes Jorasses, French Alps

7

FEATURES OF LANDSCAPES

- Topology of the basins of attraction
- Deception: Info to guide the search, Neutrality: Presence and structure of terrains with equal fitness
- Evolvability/searchability
- Funnels: Global landscape structure



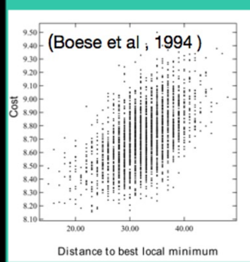
Smith T, et. al EC, 2002)

Auyantepui, Venezuela (Angel Falls, Highest Waterfall)

8

FITNESS DISTANCE CORRELATION (FDC)

- Explores the global structure of the landscape
- How closely related are fitness of solutions and distance to the nearest global optimum (or best-known solution x*)?
- A high correlation will indicate an easy search landscape ("*path*" to the optimum via solutions with increasing fitness)
- Procedure for FDC analysis
  - Generate a sample of random solutions (or local optima)
  - For each solution plot their difference in fitness vs. their distance (in genotypic space) from x*

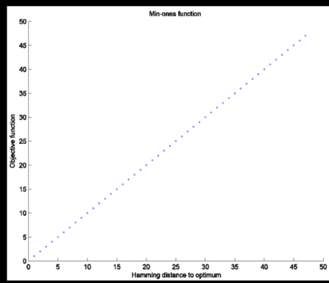


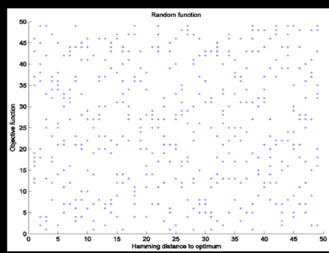
100 cities TSP instance, 2,500 2-opt local optima

9

FDC FOR TWO SIMPLE FUNCTIONS

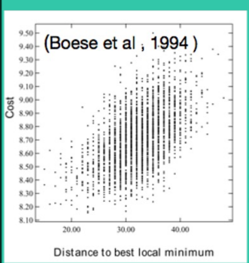
- Search space binary strings of length *n* (size 2^n)
- Hamming distance (# of different bits)
- Global optimum: x^*=0000000000 , $f(x^*) = 0$
- Min-ones function: minimise the number of ones in a binary string 
  - x_1=0010010000, $f(x_1)$=2, Hamming dist. = 2
  - x_2=1111000001, $f(x_2)$=5, Hamming dist. = 5
- Random function 
  - x_1=0010010000, $f(x_1)$=7, Hamming dist. = 2
  - x_2=1111000001, $f(x_2)$=3, Hamming dist. = 5



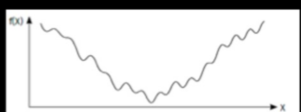


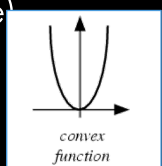
THE BIG-VALLEY HYPOTHESIS

- Several studies: NK landscapes (Kauffman, 1993), TSP (Boese et al, 1994), graph bipartitioning (Merz & Freisleben, 1998) flowshop scheduling (Reeves, 1999)
- Distribution of local optima is not uniform. Clustered in a big-valley (*globally convex* structure)

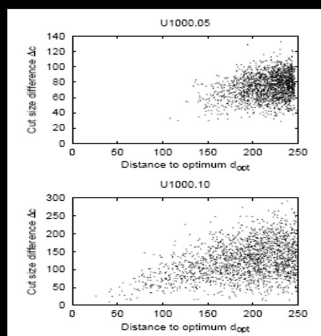


100 cities TSP instance, 2,500 2-opt local optima





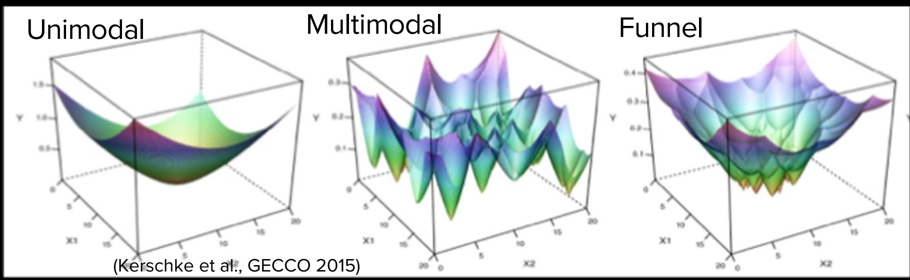
TSP: big-valley Local optima confined to a small region



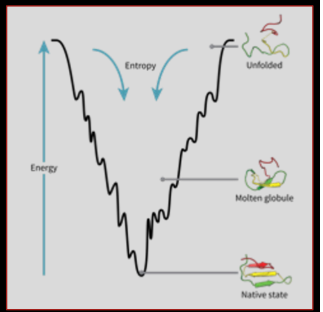
Graph bipartitioning big-valley correlation still apparent. Local minima are not as strongly confined to a small region of the solution space

WHAT IS A *FUNNEL*?

Continuous Optimisation

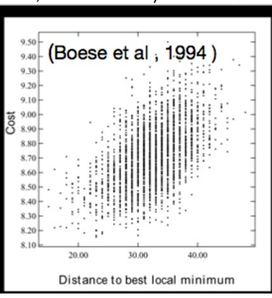


Protein Folding

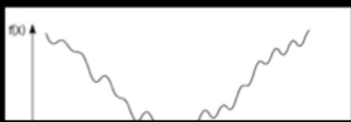


"a set of downhill pathways that converge on a single low-energy minimum." (Doye, 1999)

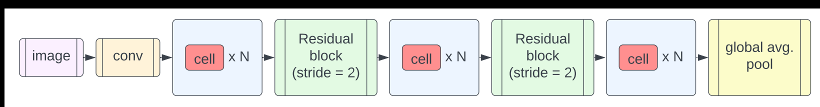
Combinatorial Optimisation



Related to the notion "big-valley"



12



LANDSCAPE ANALYSIS CASE STUDY

Neural Architecture Search (NAS)

G Ochoa, N Veerapen (2022) Neural Architecture Search: A Visual Analysis
Parallel Problem Solving from Nature – PPSN XVII. PPSN 2022.

13

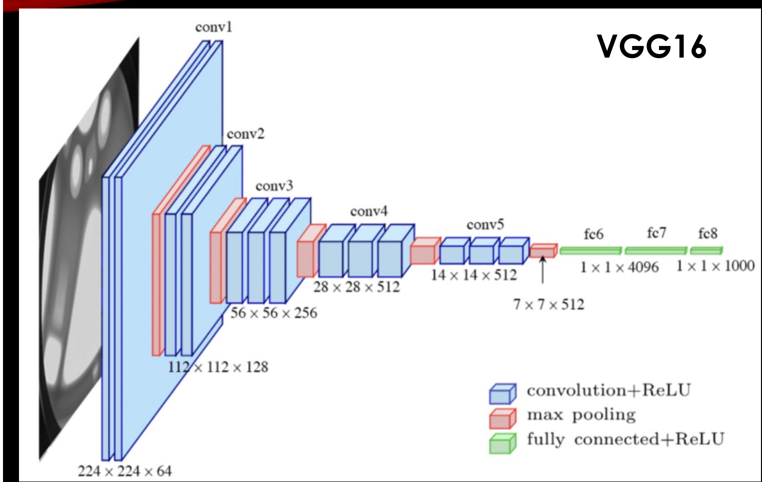
NEURAL ARCHITECTURE SEARCH (NAS)

- Use of search heuristics to optimize the topology of deep neural networks
- Has produced models that surpass the performance of huma-designed models
- Formulated as a discrete optimisation problem.

$$A^* = \underset{A \in \mathring{A}}{\operatorname{argmax}} \operatorname{Acc}(A)$$

14

DEEP CNN ARCHITECTURE

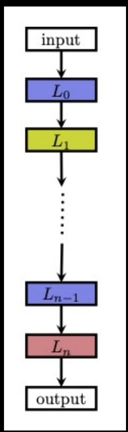


A classic convolutional NN architecture
- Total of 16 layers with weights
- Over 130 Million weights!
- 5 blocks: convolution layers followed by a max-pooling layer
- Resolution halved after each block

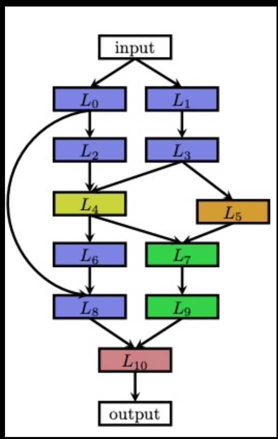
VGG: Visual Geometry Group
Karen Simonyan and Andrew Zisserman
University of Oxford, 2014
Winner of the Visual Recognition Challenge

15

ALTERNATIVE NAS SEARCH SPACES

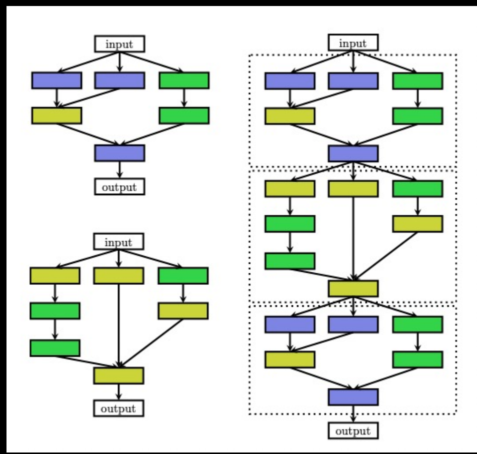


Chain Structure



Multi-branch Structure

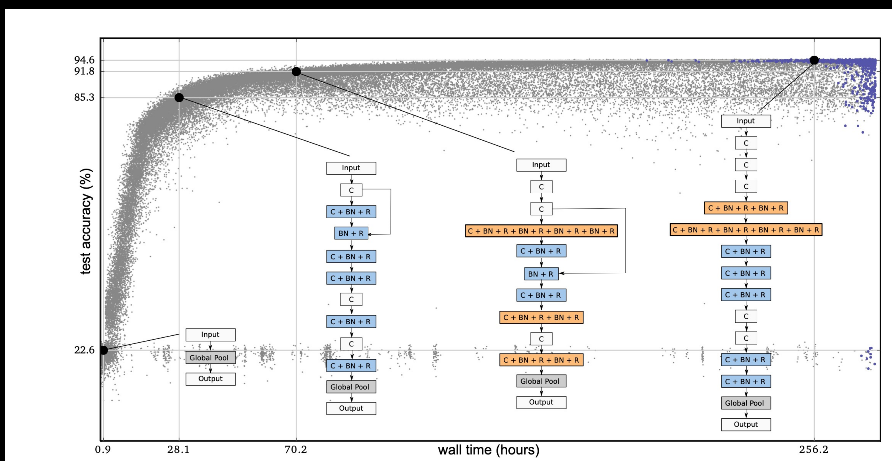
Two cells:
Normal &
Reduction



Cell (Block) Search Space

16

EXAMPLE OF SEARCH PROGRESS



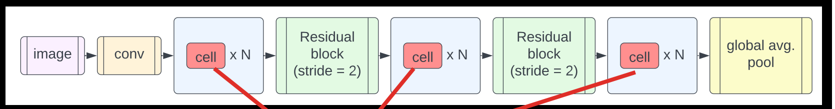
E. Real etal (2017) Large-scale evolution of image classifiers.
International Conference on Machine Learning - (ICML'17)

17

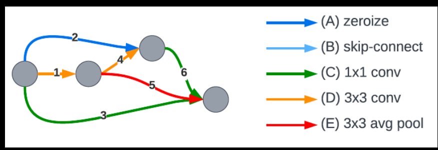
CELL BASED TABULAR BENCHMARK

Tabular benchmark (NATS-Bench - Image classification)

Cell-based search space



Fixed macro-skeleton



Cell (DAG)

Genotype

D	A	C	D	E	C

Discrete optimisation problem
max $a \in A$ $f(a)$
$f(a)$ validation accuracy

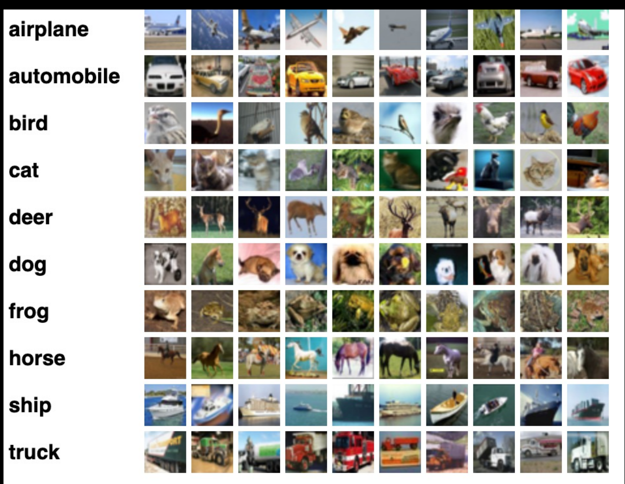
- DAG with 6 edges.
- Each edge can be one of 5 operations
- Size of the search space is $5^6 = 15,625$

Dong, X. et. al. (2021) NATS-Bench: Benchmarking NAS algorithms for architecture topology and size. *IEEE Trans. Pattern Analysis and Machine Intelligence*

18

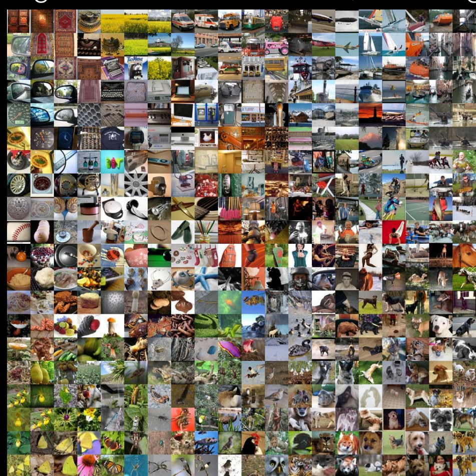
IMAGE CLASSIFICATION DATASETS

CIFAR10, CIFAR100 60,000 images, 10 &100 categories



https://www.cs.toronto.edu/~kriz/cifar.html

ImageNet ~ 14 Million images 1,000 categories
ImageNet-16-120 reduced res, 120 categories



https://cs.stanford.edu/people/karpathy/cnnembed/

19

NAS LANDSCAPES

Example data: cifar_10 data

sol
532322
445242
544555
521112
223223
323415
411544
125244
352214
511214
444111
445145
111145
152433
413411
252215
224214
523323
554423

valid-accuracy
81.98266666
90.76933332
81.91066666
84.37733332
85.12533334
87.812
88.80933334
89.15466667
89.24399999
84.46266668
9.711999999
87.94
9.712000001
88.34533333
88.59466666
56.54266666
90.26666667
86.70266667
87.784
84.80266667

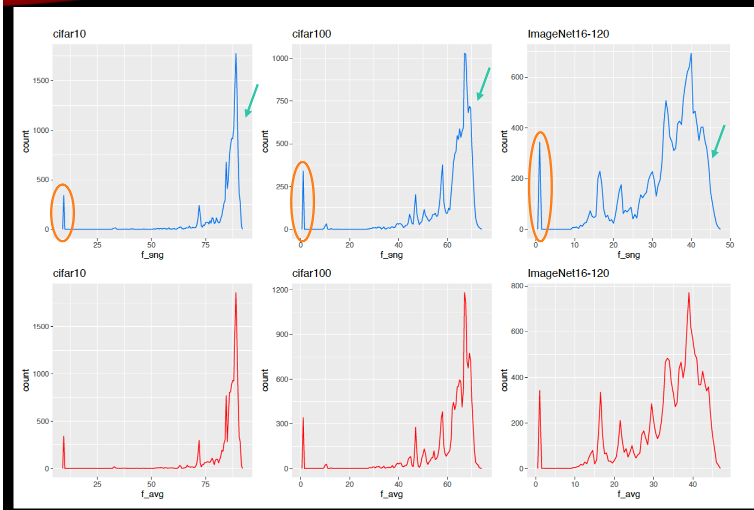
A	B	C	D	E	
index	sol	name	C	N	ar
0	532322	infer.tiny	16	5	|avg_pool_3
1	445242	infer.tiny	16	5	|nor_conv_3
2	544555	infer.tiny	16	5	|avg_pool_3
3	521112	infer.tiny	16	5	|avg_pool_3
4	223223	infer.tiny	16	5	|skip_connec
5	323415	infer.tiny	16	5	|nor_conv_1
6	411544	infer.tiny	16	5	|nor_conv_3
7	125244	infer.tiny	16	5	|none~0|+|s
8	352214	infer.tiny	16	5	|nor_conv_1
9	511214	infer.tiny	16	5	|avg_pool_3
10	444111	infer.tiny	16	5	|nor_conv_3
11	445145	infer.tiny	16	5	|nor_conv_3
12	111145	infer.tiny	16	5	|none~0|+|n
13	152433	infer.tiny	16	5	|none~0|+|a
14	413411	infer.tiny	16	5	|nor_conv_3
15	252215	infer.tiny	16	5	|skip_connec
16	224214	infer.tiny	16	5	|skip_connec
17	523323	infer.tiny	16	5	|avg_pool_3
18	554423	infer.tiny	16	5	|avg_pool_3
19	443551	infer.tiny	16	5	|nor_conv_3
20	155244	infer.tiny	16	5	|none~0|+|a

index	I	J	K	L	M	N	O	S	T
index	in-loss	train-accuracy	rain-per-time	train-all-time	valid-loss	valid-accuracy	valid-per-	t-per-time	test-all-time
0	1902812	88.9186667	7.221093	1444.2186	0.56585078	81.9826667	2.54897	1.0195917	203.918339
1	0176718	99.9853333	10.8482069	2169.64138	0.4389296	90.7693333	3.75419	.50167959	300.335918
2	0833942	89.3146667	11.3358884	2267.17768	0.5661588	81.9106667	3.97424	.58969879	317.939758
3	5283107	91.24	4.16524494	833.048987	0.47687674	84.3773333	2.43655	.97462394	194.924788
4	5593634	91.292	6.48429906	1296.85981	0.45933634	85.1253333	2.41316	.96526407	193.052814
5	0046234	99.9533333	9.96243	1992.486	0.57506105	87.812	3.24370	.29748196	259.496391
6	0267631	99.9653333	11.2853785	2257.0757	0.57231554	88.8093333	3.93284	1.5731383	314.62766
7	0695222	99.9213333	8.70576368	1741.15274	0.4590636	89.1546667	3.21537	.28615121	257.230241
8	0440578	99.9613333	8.20519018	1641.03804	0.47918321	89.244	2.87121	.14848718	229.697435
9	5870412	90.944	6.53986319	1307.97264	0.47469274	84.4626667	2.68279	.07311813	214.623626
10	0246029	10.288	10.2379573	2047.59146	2.30295691	9.712	3.65178	.46071284	292.142567
11	0996415	99.7813333	12.4620568	2492.41136	0.55911125	87.94	4.15125	.66050069	332.100137
12	0246032	10.288	6.65870827	1331.74166	2.30295848	9.712	2.7200	.08800188	217.600376
13	0758837	99.864	10.1294189	2025.88378	0.53853208	88.3453333	3.20340	.28136018	256.272036
14	0728176	99.888	9.83898797	1967.79759	0.51229702	88.5946667	3.37143	.34857213	269.714426
15	7659096	57.8653333	4.62867395	925.73479	1.22572374	56.5426667	2.39076	.95630461	191.260921
16	0235439	99.9773333	7.61449607	1522.89921	0.46507577	90.2666667	2.91691	1.1667647	233.35294
17	0106718	99.8506667	9.23993437	1847.98687	0.59593535	86.7026667	2.94650	.17860394	235.720789
18	0439566	99.9546667	11.6001165	2320.02329	0.58827107	87.784	3.75118	.50047463	300.094926
19	2584353	99.3773333	11.4089125	2281.7825	0.69198427	84.8026667	3.78325	.51330337	302.660674
20	0188796	99.58	10.0112177	2002.24354	0.57879267	87.2853333	3.48725	.39490385	278.980769

…

21

DENSITY OF STATES



Density of states (DOS) number of solutions with a certain fitness value

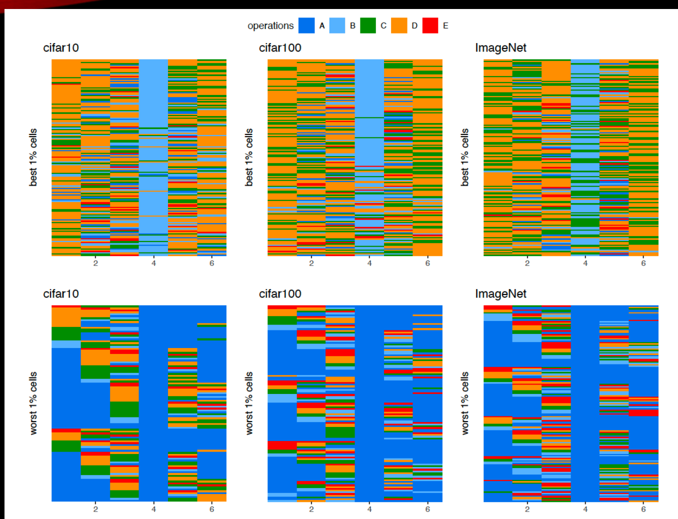
Gives an idea of the distribution of fitness values, amount of neutrality.

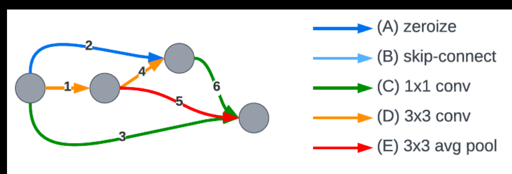
Fitness. fun: *f_sng* (top), *f_avg* (bottom)

x-axis - whole range of validation accuracy values grouped in bins of width 0.5 - frequency polygons.

22

NAS LANDSCAPES GENOTYPE MAPS

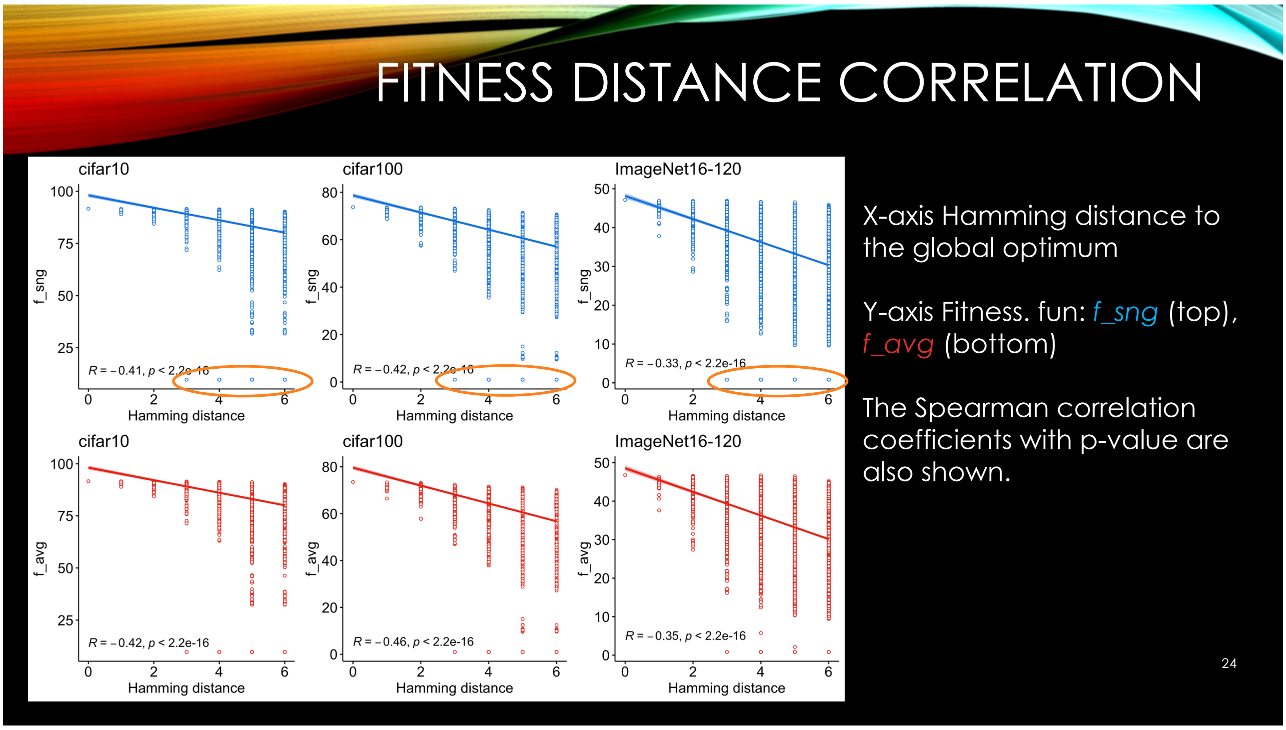


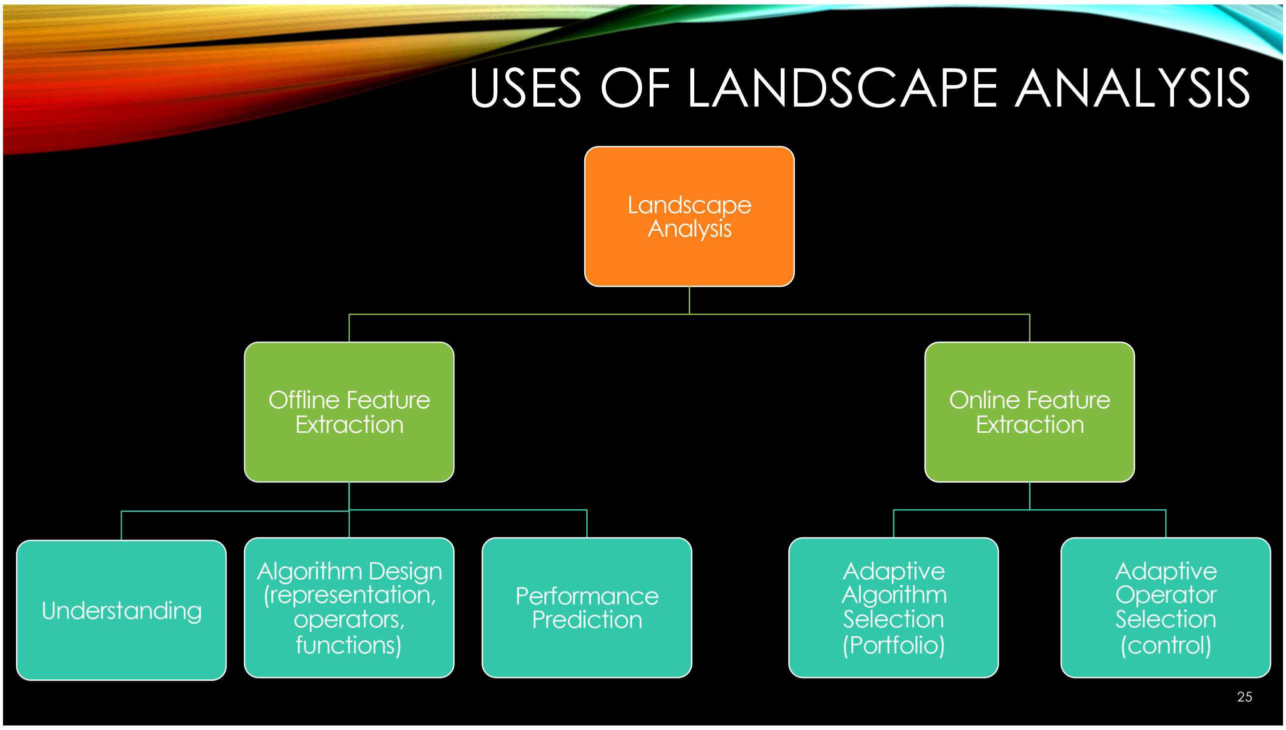


Genotype maps of the best 1% (top), and worst 1% (bottom) performing cells, sorted according to *f_avg*.

Each line visualises a cell. Positions are coloured according to the respective operation.

23


FITNESS DISTANCE CORRELATION
cifar10
cifar100
ImageNet16-120
f_sng
f_avg
Hamming distance
R = −0.41, p < 2.2e-16
R = −0.42, p < 2.2e-16
R = −0.33, p < 2.2e-16
R = −0.42, p < 2.2e-16
R = −0.46, p < 2.2e-16
R = −0.35, p < 2.2e-16
X-axis Hamming distance to the global optimum
Y-axis Fitness. fun: f_sng (top), f_avg (bottom)
The Spearman correlation coefficients with p-value are also shown.
24


USES OF LANDSCAPE ANALYSIS
Landscape Analysis
Offline Feature Extraction
Online Feature Extraction
Understanding
Algorithm Design (representation, operators, functions)
Performance Prediction
Adaptive Algorithm Selection (Portfolio)
Adaptive Operator Selection (control)
25

BEYOND ***FITNESS*** LANDSCAPES

Idea of *fitness* landscapes have been applied in non-evolutionary contexts, so many are dropping the fitness term.

New kinds of landscapes

- Multiobjective fitness landscapes
- Constraint violation landscapes
- Dynamic and coupled fitness landscapes
- Error / loss landscapes in the context of neural networks

Search landscape analysis

Exploratory landscape analysis

Error/loss landscape analysis

Landscape analysis

26

CONCLUSION

- There are many different techniques for conducting landscape analysis.
- Landscape analysis has evolved into a useful tool in different contexts

understanding and analysing complex problems and benchmark suites

understanding and explaining algorithm behaviour

predicting algorithm performance

algorithm configuration

automated algorithm selection

27

REFERENCES

Early/Classic

- P. F. Stadler (2002) Fitness Landscapes. In: Biological Evolution and Statistical Physics, pp. 187–207. Springer, Heidelberg
- E. Weinberger (1990) Correlated and uncorrelated fitness landscapes and how to tell the difference. *Biol. Cybern*. 63(5), 325–336
- S. Wright (1932) *The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in evolution*, Proc. Sixth International Congress on Genetics, pp. 356–366.

Surveys

- K. M. Malan (2021). A Survey of Advances in Landscape Analysis for Optimisation. *Algorithms*, 14(2).
- K. M. Malan, A. P. Engelbrecht (2013)A survey of techniques for characterising fitness landscapes and some possible ways forward. *Inf. Sci*. 2013, 241, 148–163.
- E. Pitzer, M. Affenzelle (2012) A Comprehensive Survey on Fitness Landscape Analysis. In Recent Advances in Intelligent Engineering Systems; Springer, 2012; pp. 161–191.
- F Zou, D Chen, H Liu, S Cao, X Ji, Y Zhang (2022) A survey of fitness landscape analysis for optimization, *Neurocomputing*, Volume 503, pp. 129-139,

REFERENCES

Book (Edited)

- Hendrik Richter, Andries Engelbrecht (2014) Recent Advances in the Theory and Application of Fitness Landscapes, Emergence, Complexity and Computation (ECC, volume 6), Springer

Book Chapters

- H. H. Hoos and T. Stutzle (2005) Stochastic Local Search: Foundations & Applications. Morgan Kaufmann Publishers Inc. Chapter 5 Search Space Structure and SLS Performance
- Reeves, C.R. (2005). Fitness Landscapes. In: Burke, E.K., Kendall, G. (eds) Search Methodologies. Springer, Boston, MA. https://doi.org/10.1007/0-387-28356-0_19