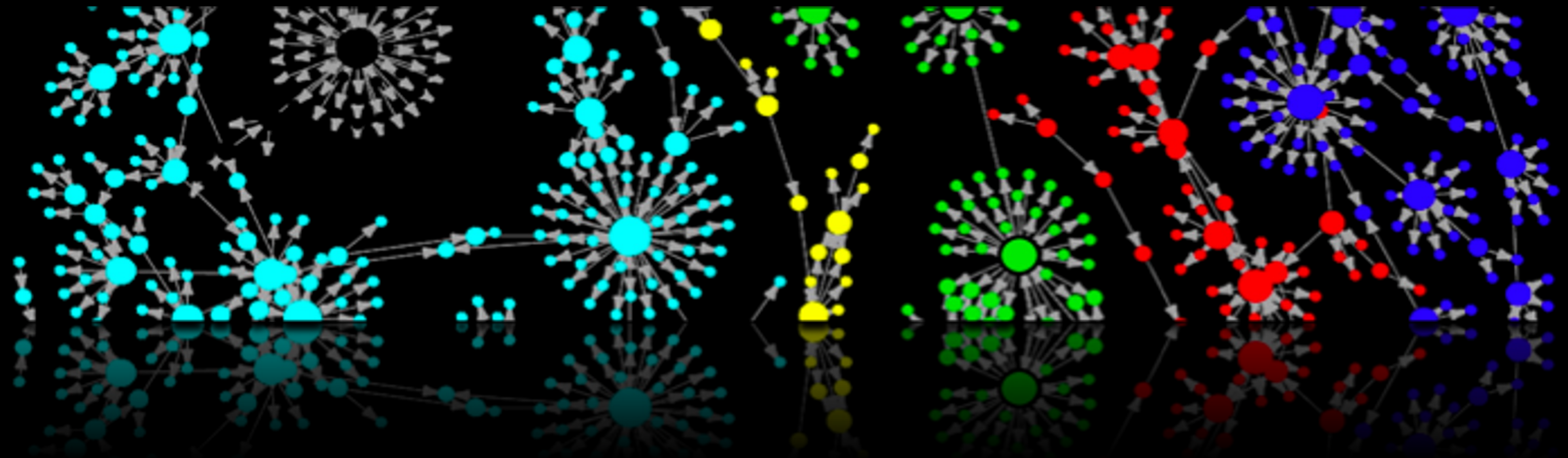
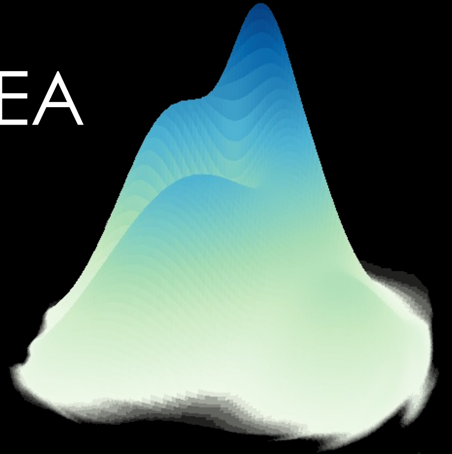


JoLEA



NEUROEVOLUTION LANDSCAPES AND TRAJECTORIES

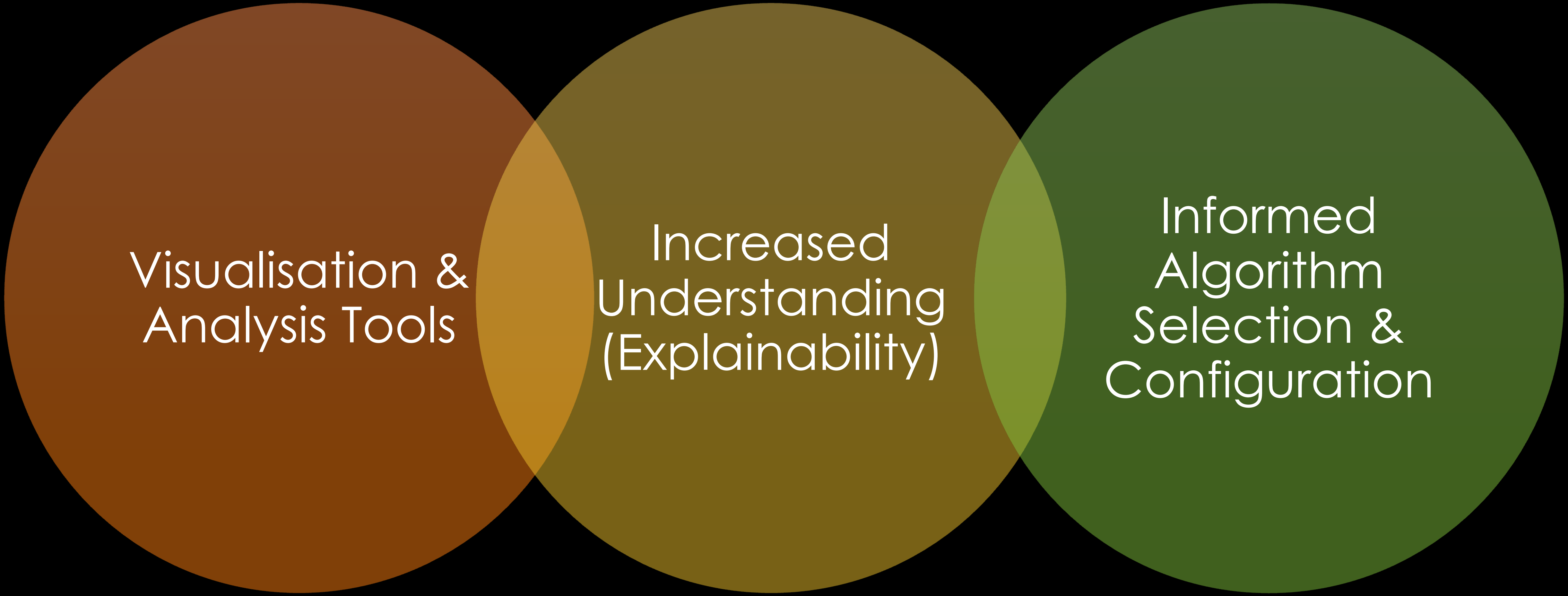
Collaborators
Jason Adair, Katherine Malan, Nuno Rodrigues, Stefano Sarti, Sara Silva, Leonardo Vanneschi, Nadarajen Veerapen

Gabriela Ochoa

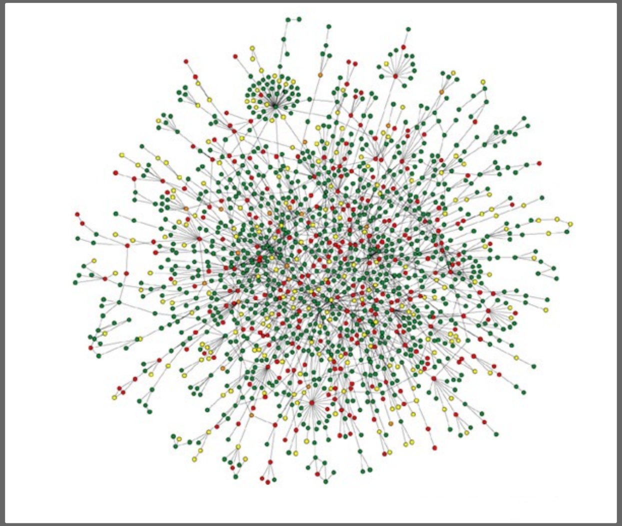
UNIVERSITY of STIRLING

UNIVERSITY of STIRLING

MOTIVATION

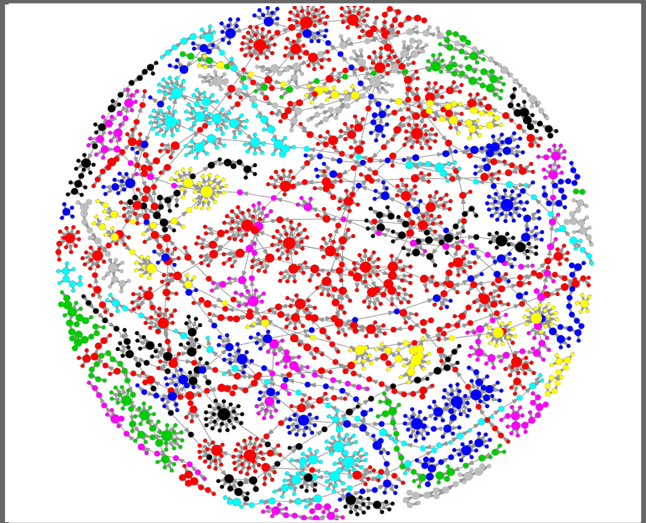

Visualisation & Analysis Tools
Increased Understanding (Explainability)
Informed Algorithm Selection & Configuration

OUTLINE



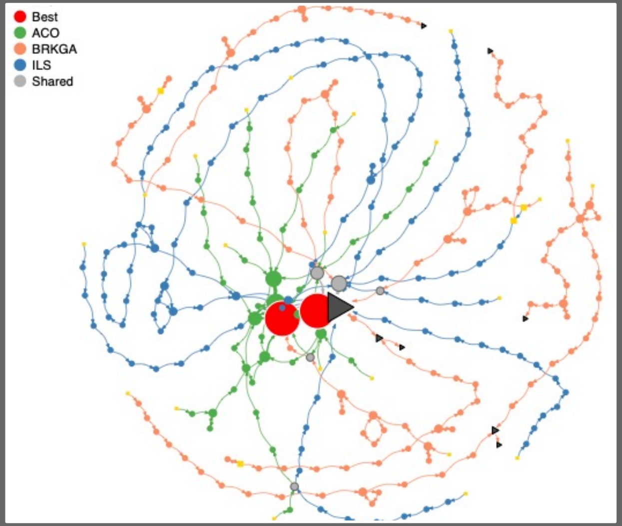
G. Theory ~1700, ~2000

Complex Networks



2008, NAS 2022

Local Optima Networks (LONs)

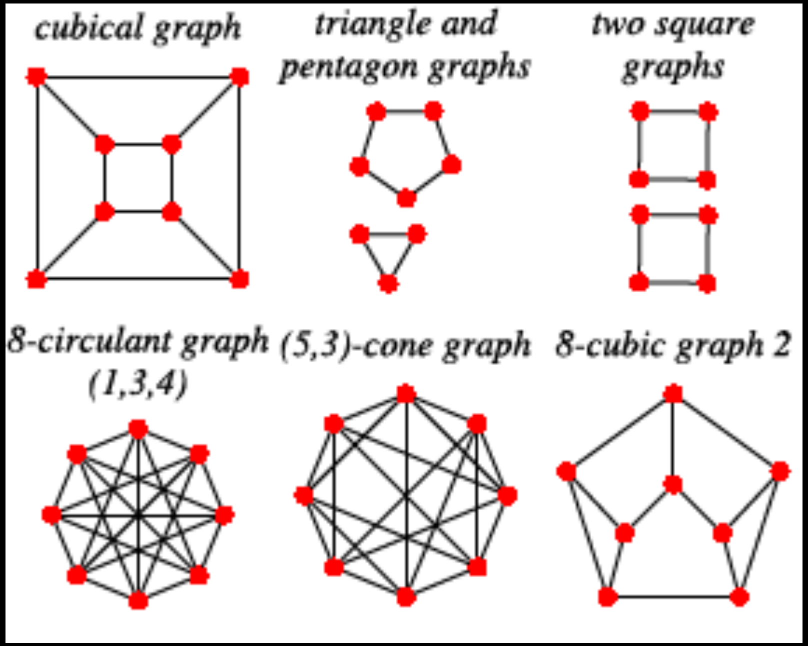


2020, NE 2021, 2022

Search Trajectory Networks (STNs)

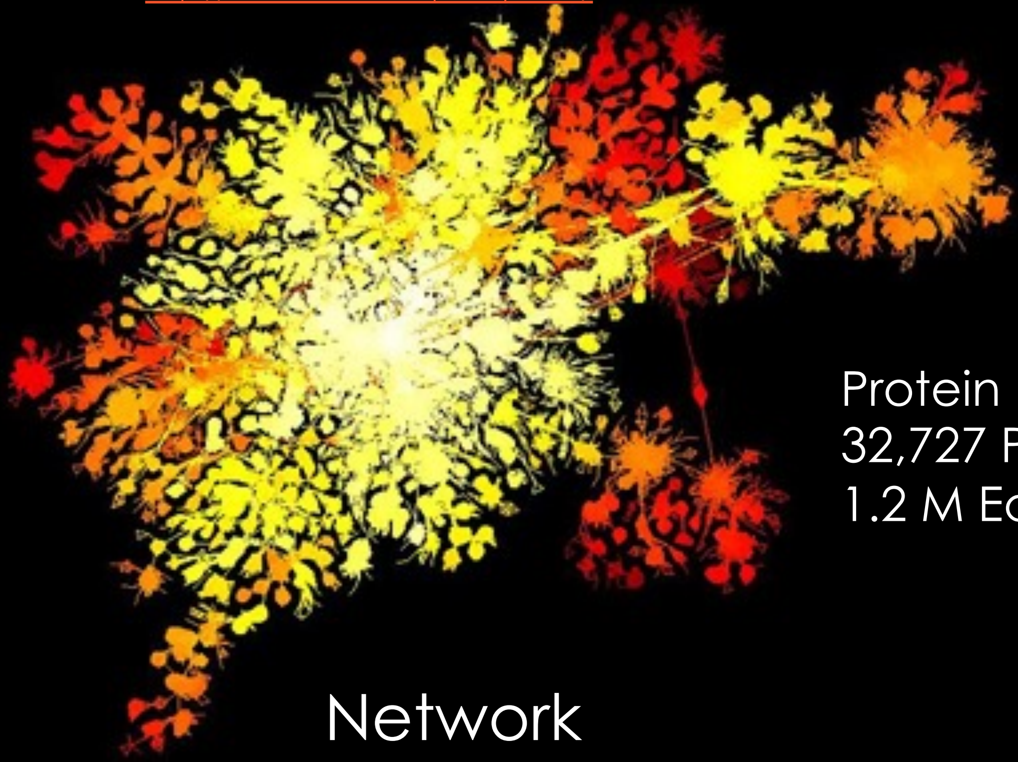
WHAT IS A (COMPLEX) NETWORK

https://mathworld.wolfram.com/WeaklyRegularGraph.html



Graph
Mathematical Object

http://www.visualcomplexity.com/



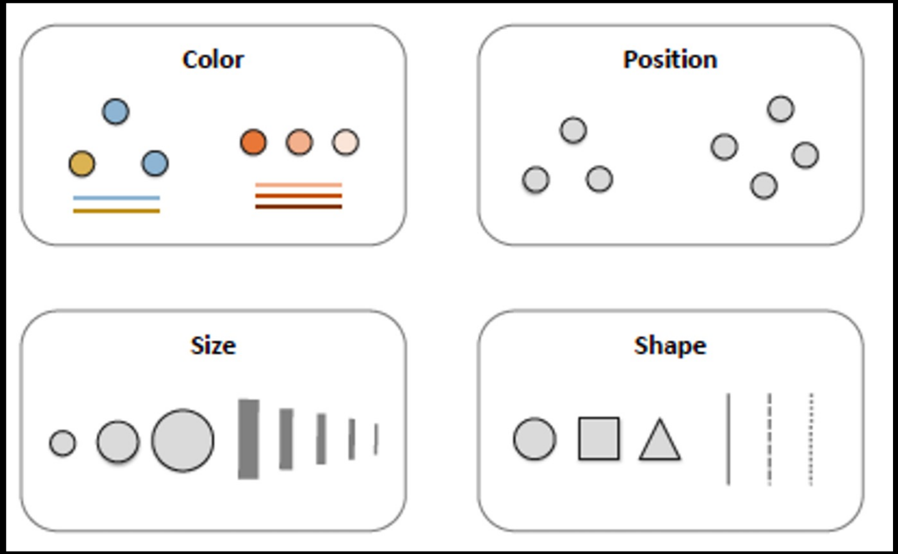
Protein Homology
32,727 Proteins
1.2 M Edges

Network
Data Driven Instantiation

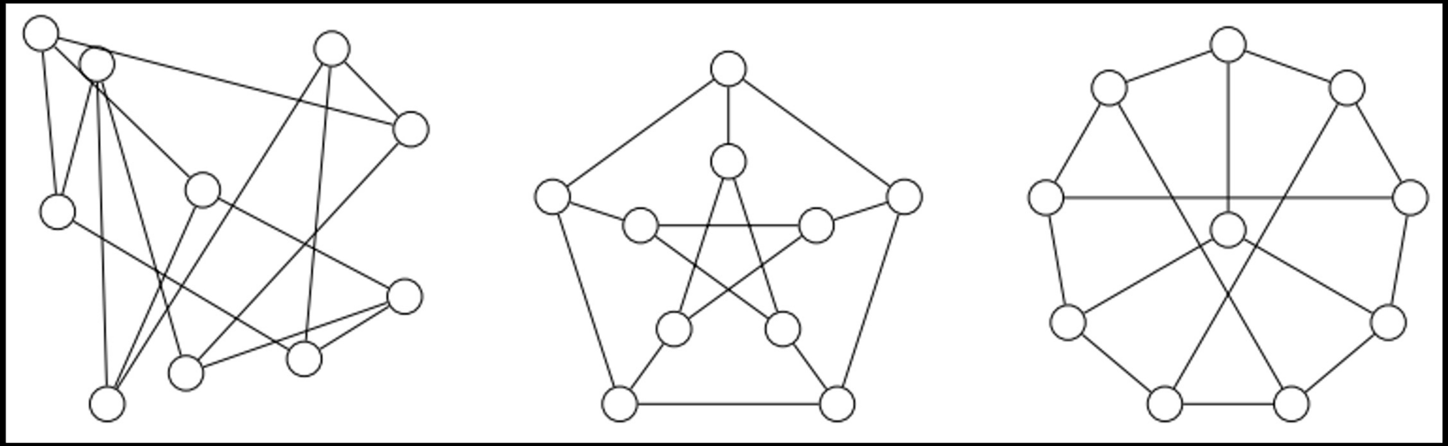
Behind each complex system, there is an intricate network that encodes the interactions between the system's components. Albert-László Barabási, Network Science

NETWORK VISUALISATION

Art of choosing an appropriate representation that is *aesthetically pleasing* and highlights important *structural properties*



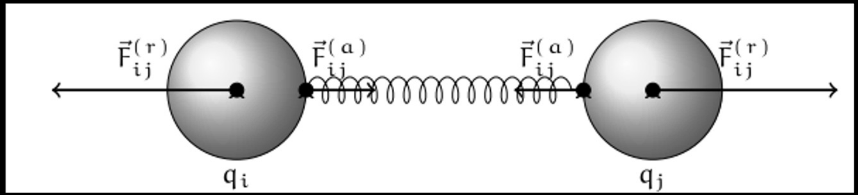
https://kateto.net/network-visualization



Petersen graph

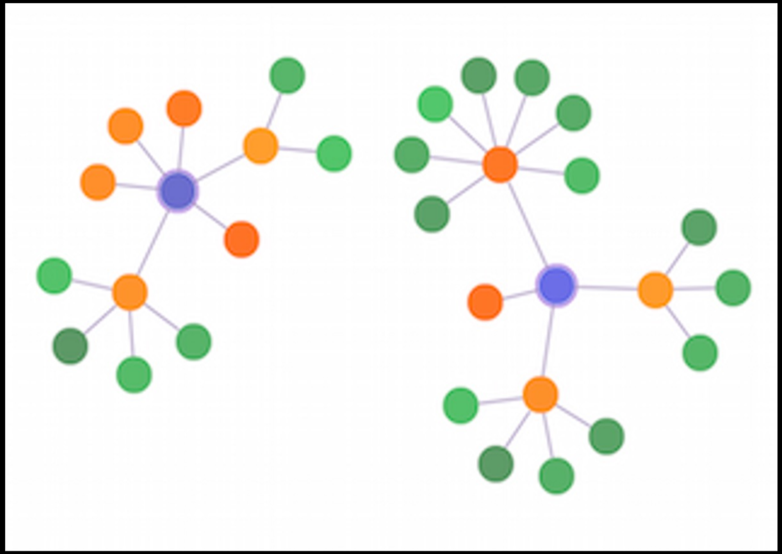
Node-edge diagram graph layouts

FORCE-DIRECTED GRAPH LAYOUT ALGORITHMS

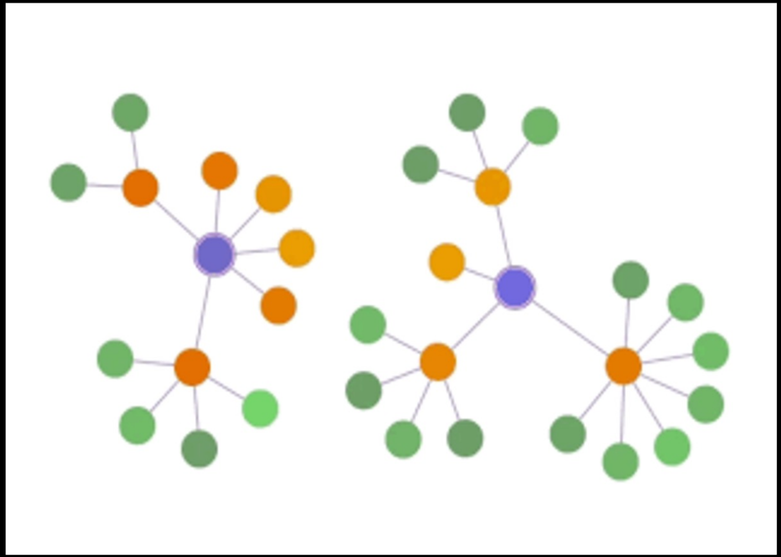


Aesthetic criteria

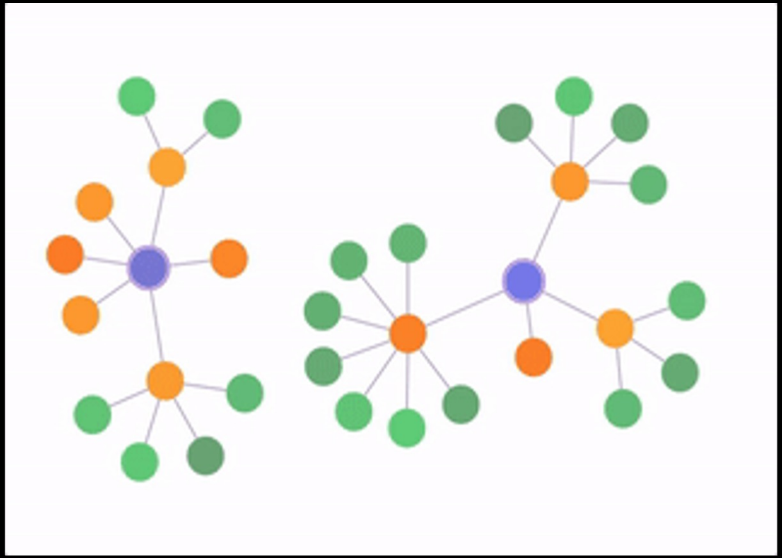
- Vertices are evenly distributed
- The number of edge crossings is minimised
- The lengths of edges are approximately uniform
- Inherent symmetries in the graph are respected



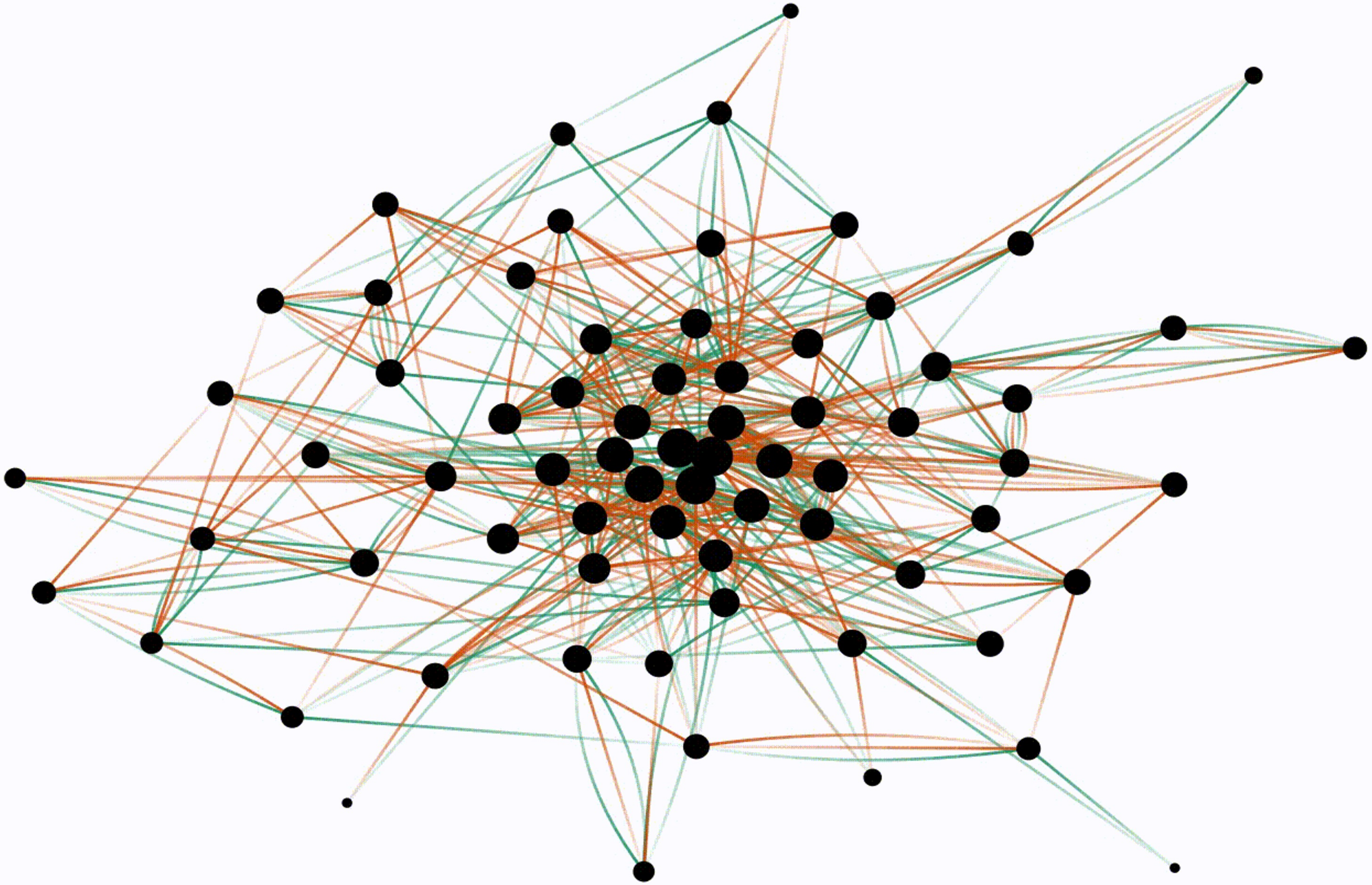
Fruchterman-Reingold



Kamada–Kawai (*organic*)



Reingold-Tilford (trees)



Graph Layout

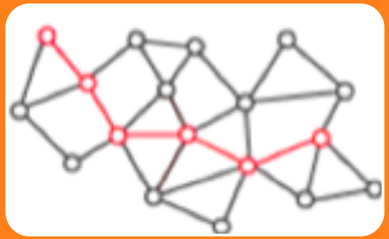
Many Algorithms

Some are simple: circular, star, grid, random.

Others optimise the position of nodes based on different characteristics of the graph.

Source https://www.data-imaginist.com/2017/ggraph-introduction-layouts/

NETWORK METRICS

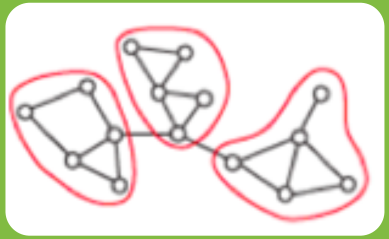


Distance

- Diameter
- Average shortest path

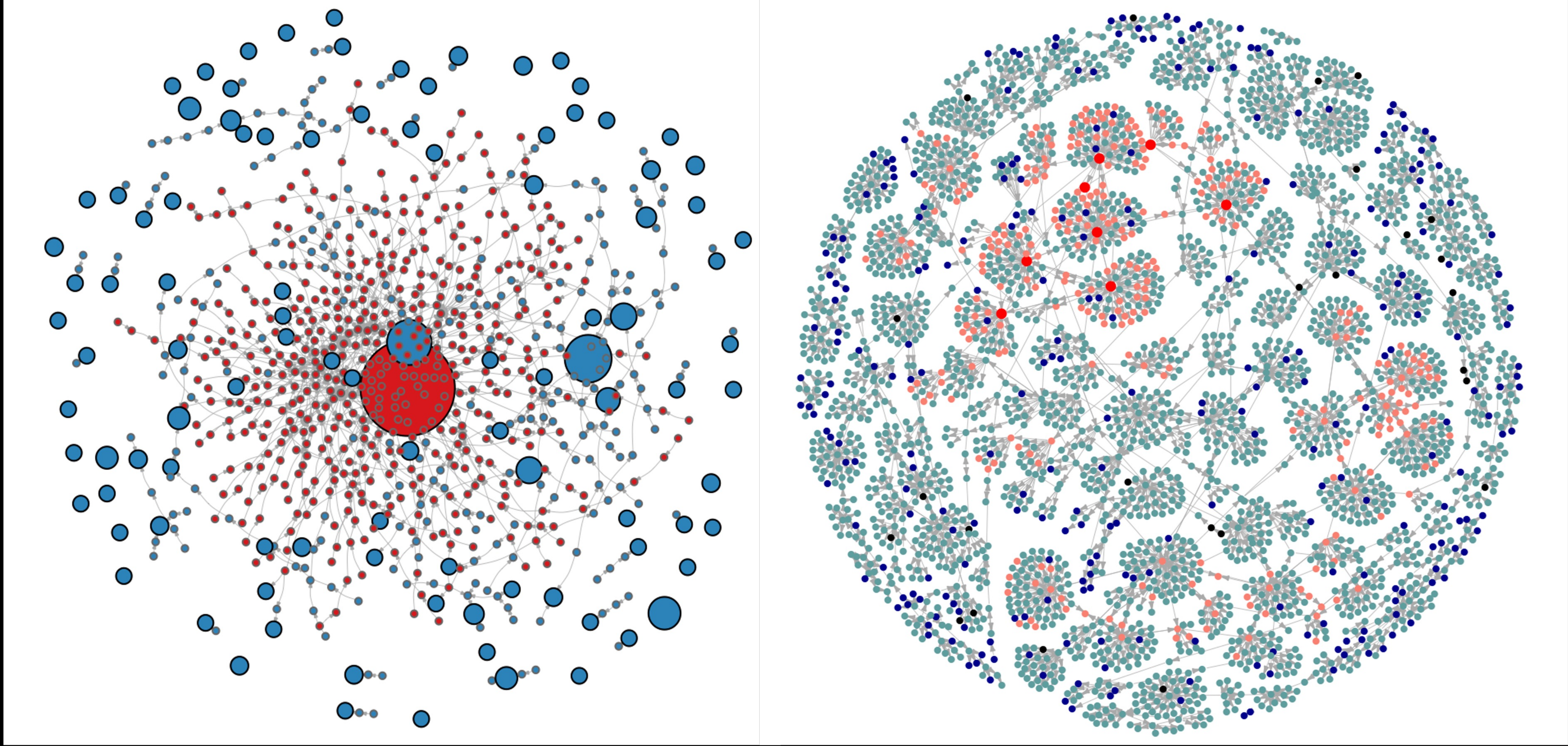
Topology

- Degree distribution
- Mean degree, Assortativity, disparity, centrality



Cohesion

- Clustering coefficient
- Community structure



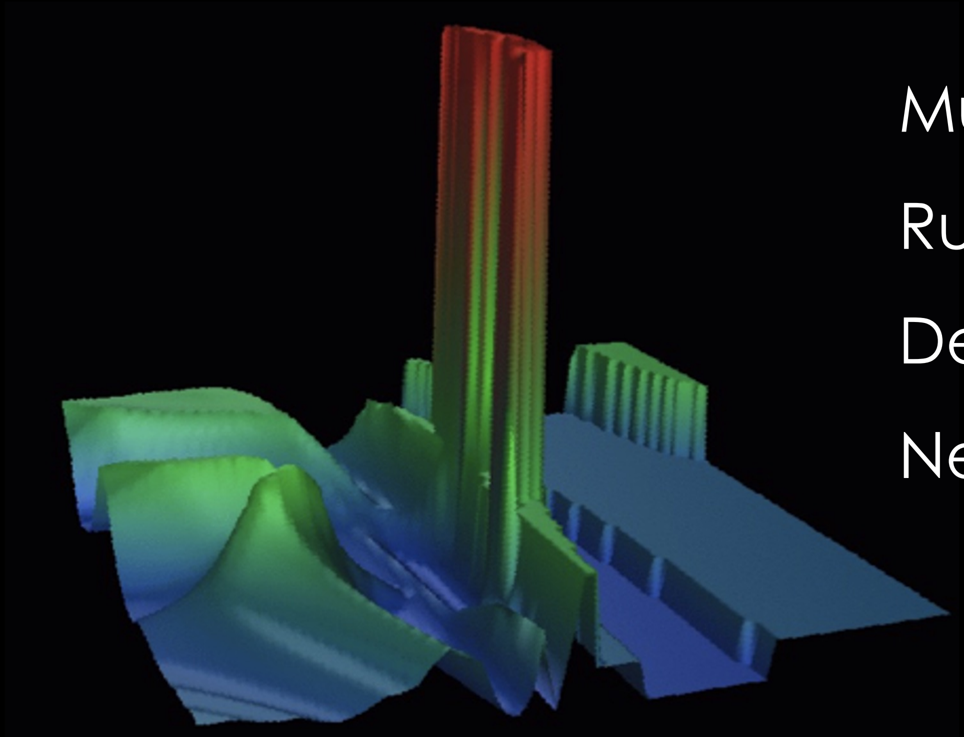
LOCAL OPTIMA NETWORKS (LONS)

FITNESS LANDSCAPES

(S, N, f)

- Search Space
- Neighbourhood Structure
- Fitness Function

Biology, Computational Chemistry
Evolutionary Computation & Metaheuristics
Machine Learning



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

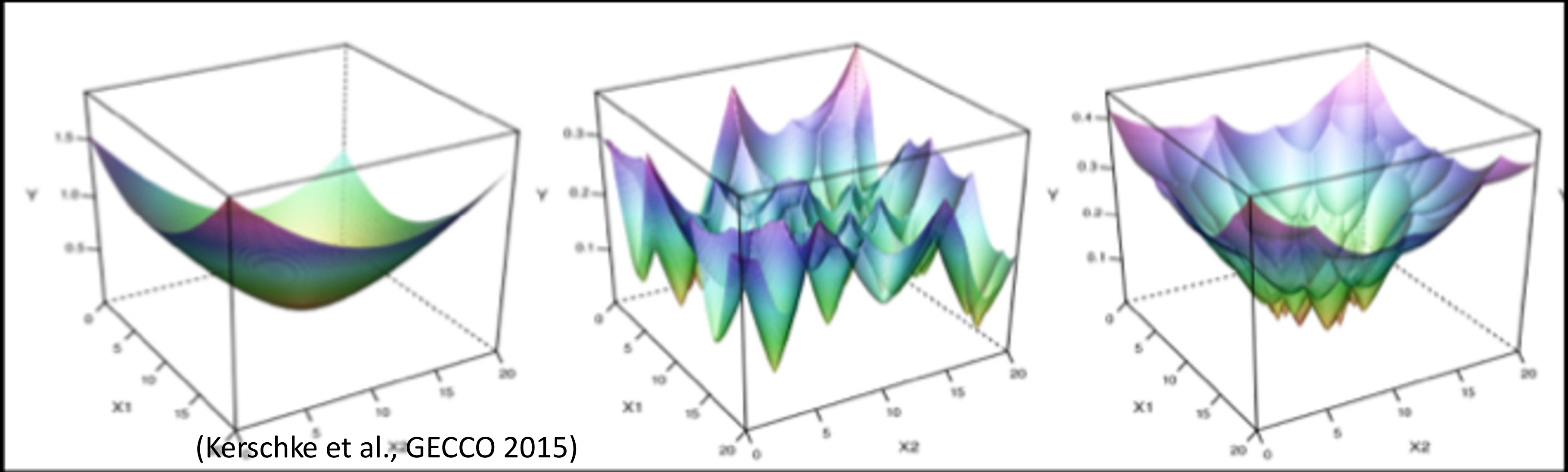
Multimodality

Ruggedness

Deceptivenes

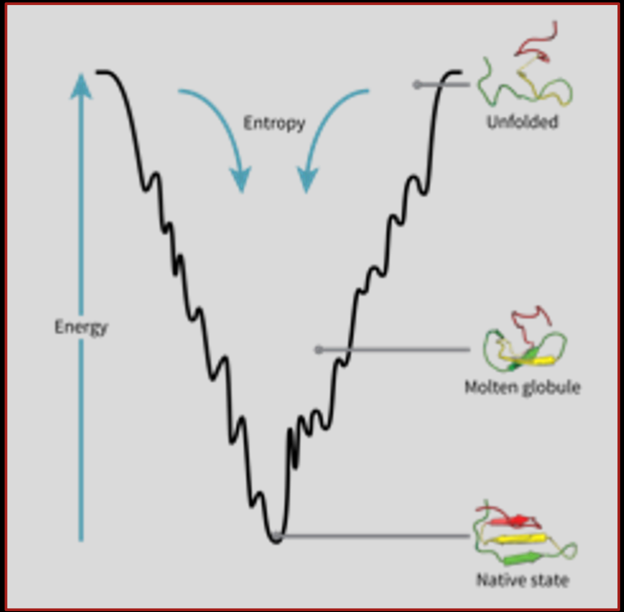
Neutrality

WHAT IS A *FUNNEL*?

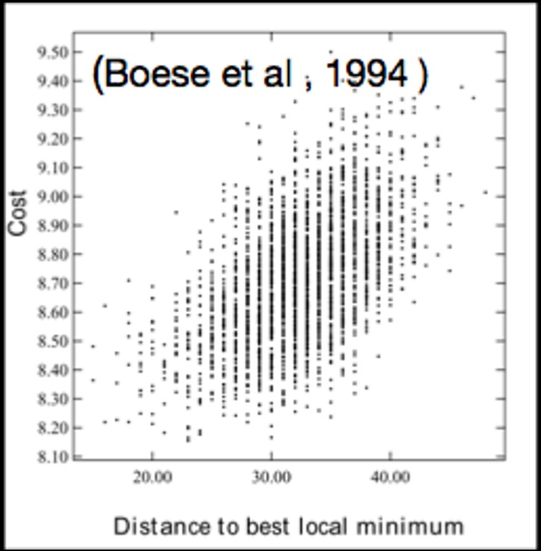


(Kerschke et al., GECCO 2015)

Protein Folding

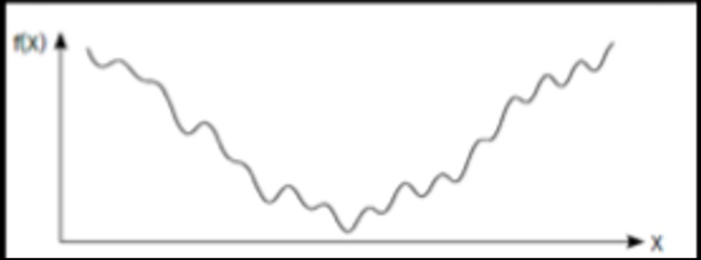


T. Splettstoesser
www.scistyle.com



(Boese et al , 1994)

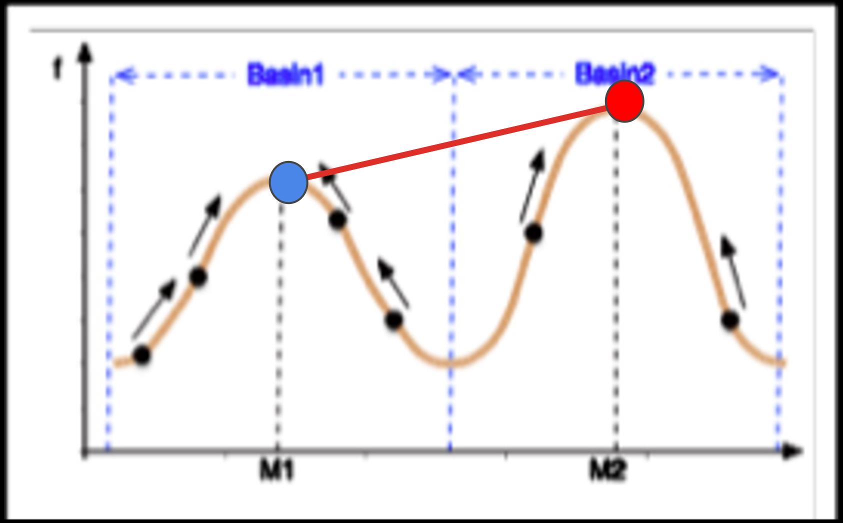
Related to the notion "big-valley"



LOCAL OPTIMA NETWORKS (LONS)

Nodes - local optima according to a hill-climbing heuristic

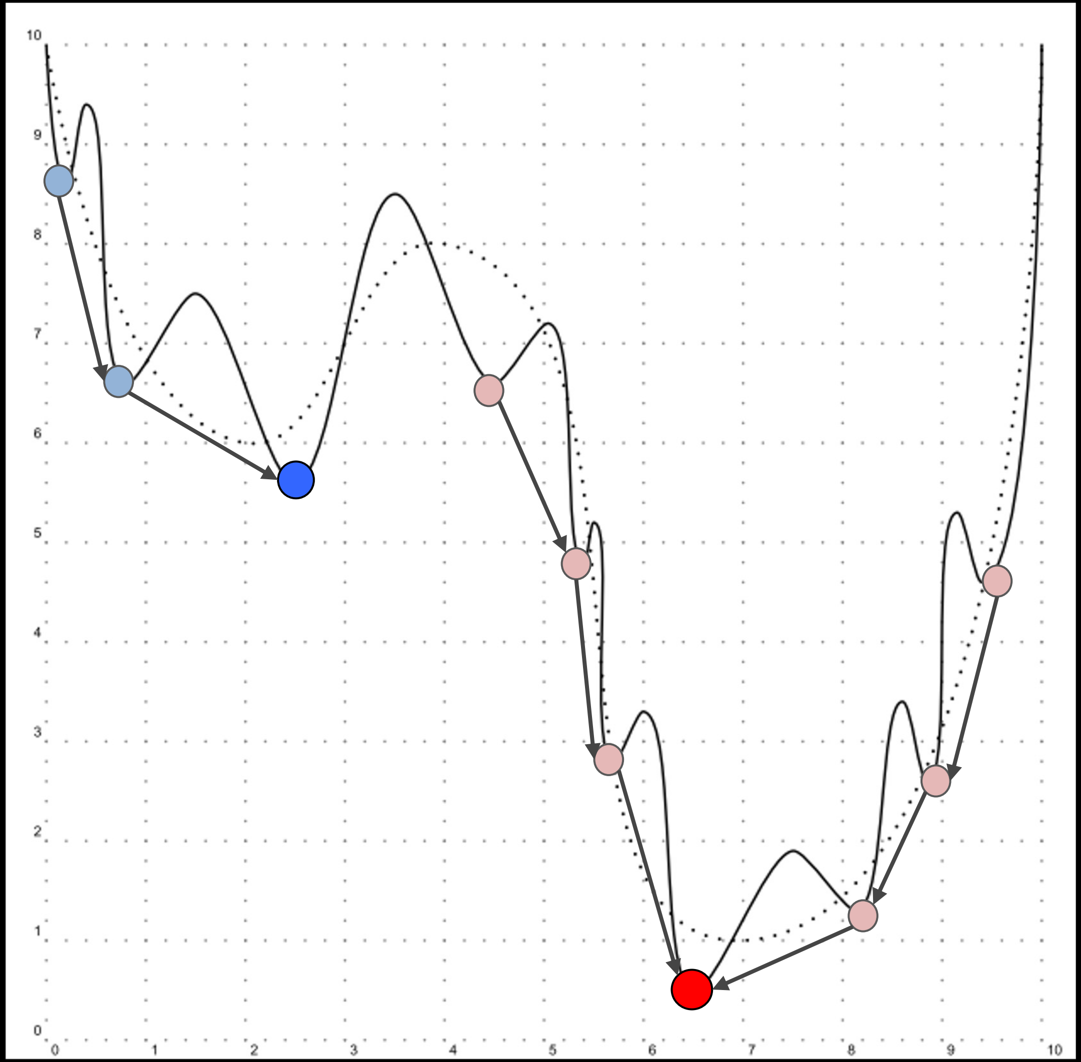
Edges - transitions between optima (basin, escape, crossover)



P. K. Doye. The network topology of a potential energy landscape: a static scale-free network. *Physical Review Letter*, 2002.

G. Ochoa, M. Tomassini, S. Verel, and C. Darabos. A study of NK landscapes' basins and local optima networks. GECCO 2008

Characterising Funnels with LONs



- Global minimum
- Sub-optimal sink
- Local minimum in optimal funnel
- Local minimum in sub-optimal funnel
- Monotonic edge

MLONs

Monotonic Sequence: a descending sequence of adjacent minima

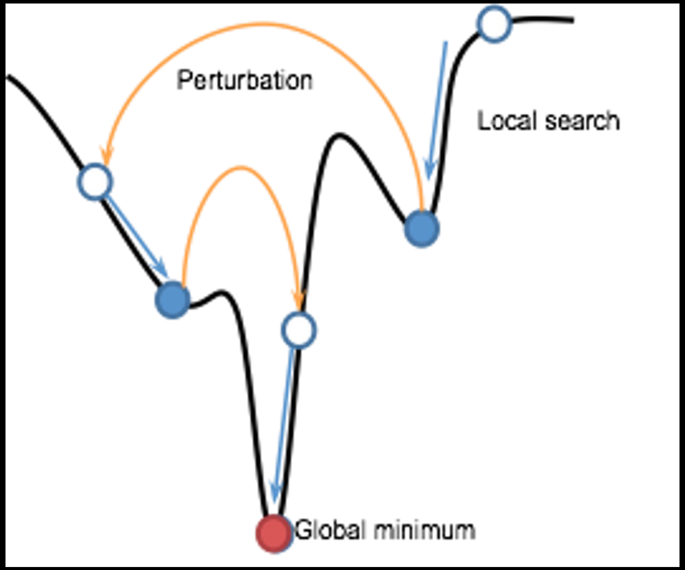
Funnel: the aggregation of all monotonic sequences ending at the same point (or sink).

TRAVELLING SALESMAN (TSP)

Sampling and constructing LONs with escape edges

Algorithm 1 ILS sampling

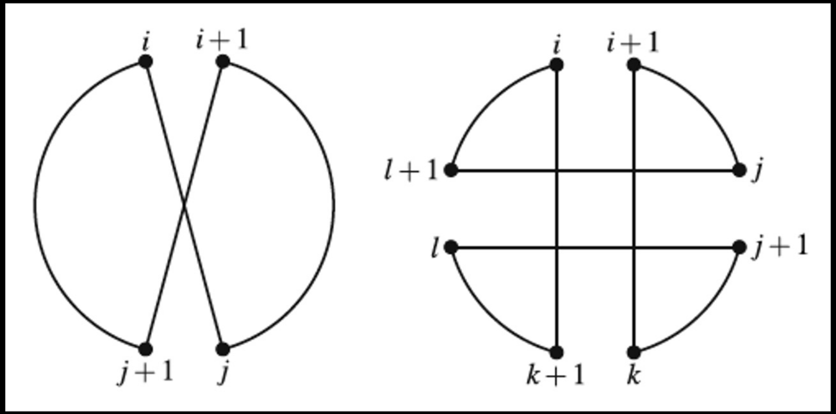
$L \leftarrow \{\}$; $E \leftarrow \{\}$ (initialise LON data)
repeat
  $s_0 \leftarrow \mathrm{RandomInitialSolution}$
  $s^* \leftarrow \mathrm{HillClimber}(s_0)$
  $L \leftarrow L \cup \{s^*\}$ (record LON data)
  repeat
    $s' \leftarrow \mathrm{RandomMutation}(s^*)$
    $s'^* \leftarrow \mathrm{HillClimber}(s')$
    if $f(s'^*) \leq f(s^*)$ **then**
      $L \leftarrow L \cup \{s'^*\}$ (record LON data)
      $E \leftarrow E \cup \{(s^*, s'^*)\}$ (record LON data)
      $s^* \leftarrow s'^*$
    end if
  until *ILS termination condition is met*
until *sampling termination condition is met*



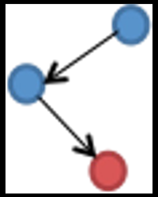
Chained Lin-Kernighan

(Martin, Otto, Felten, 1992)

- Form of Iterated Local Search
- Diversification & Intensification

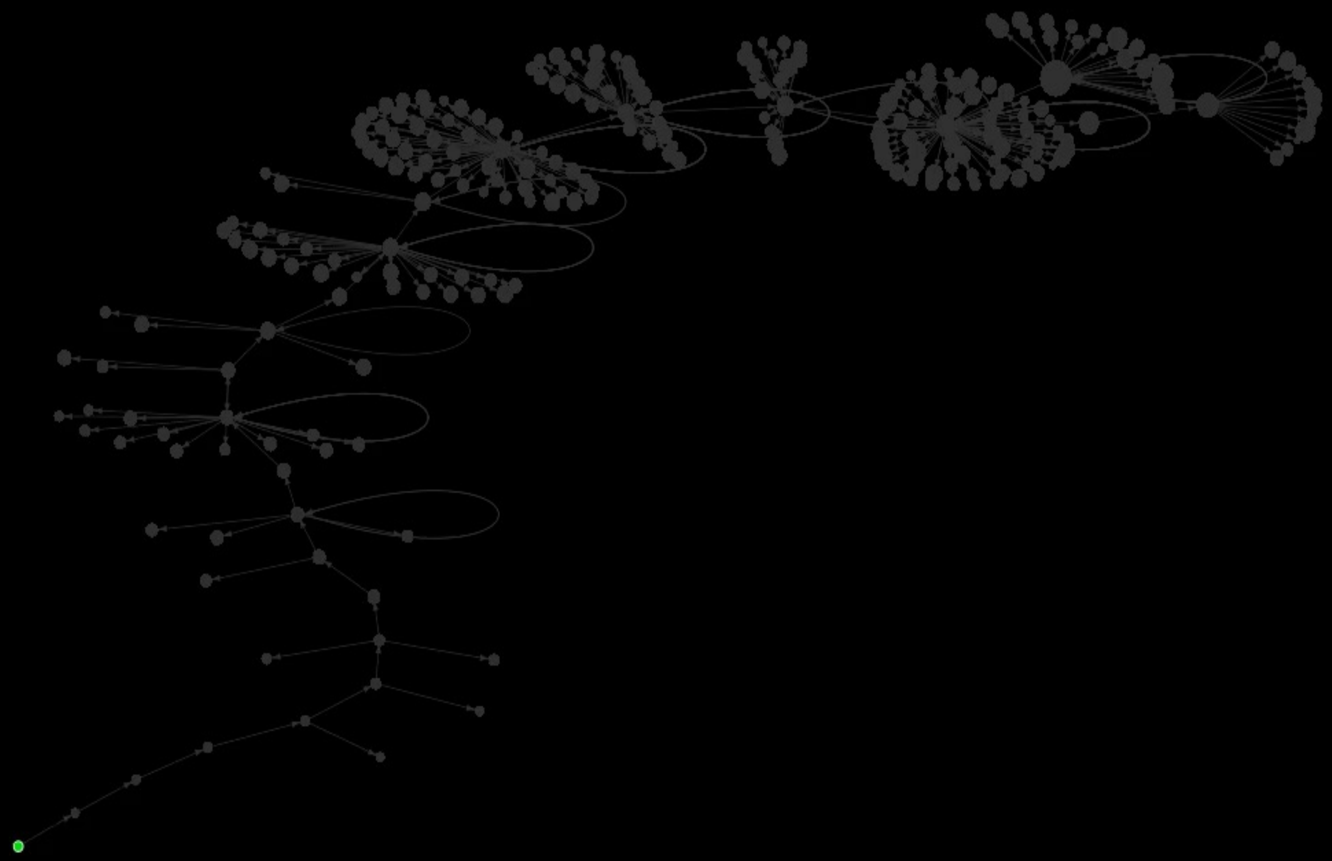


2-exchange double-bridge

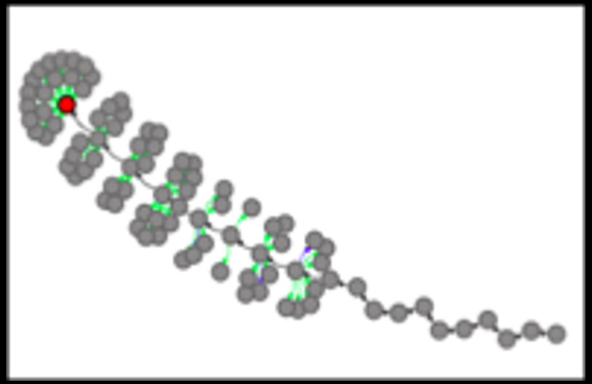


Nodes: LK local optima

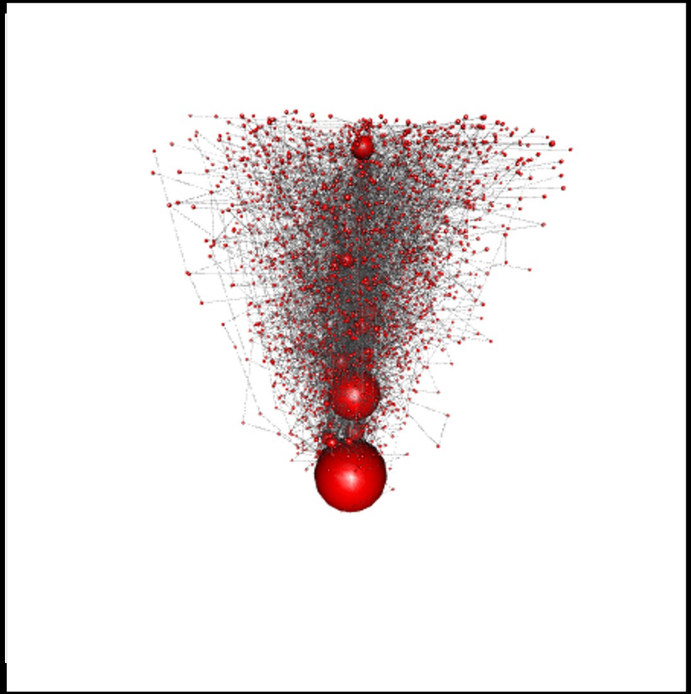
Edges: double-bridge escapes



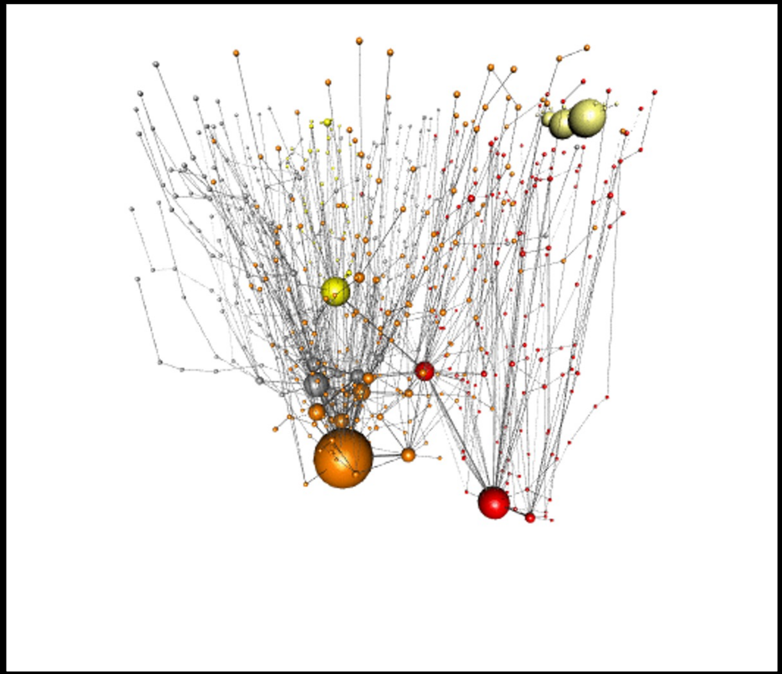
A Chained-LK run on instance lin318
318 points, drilling application



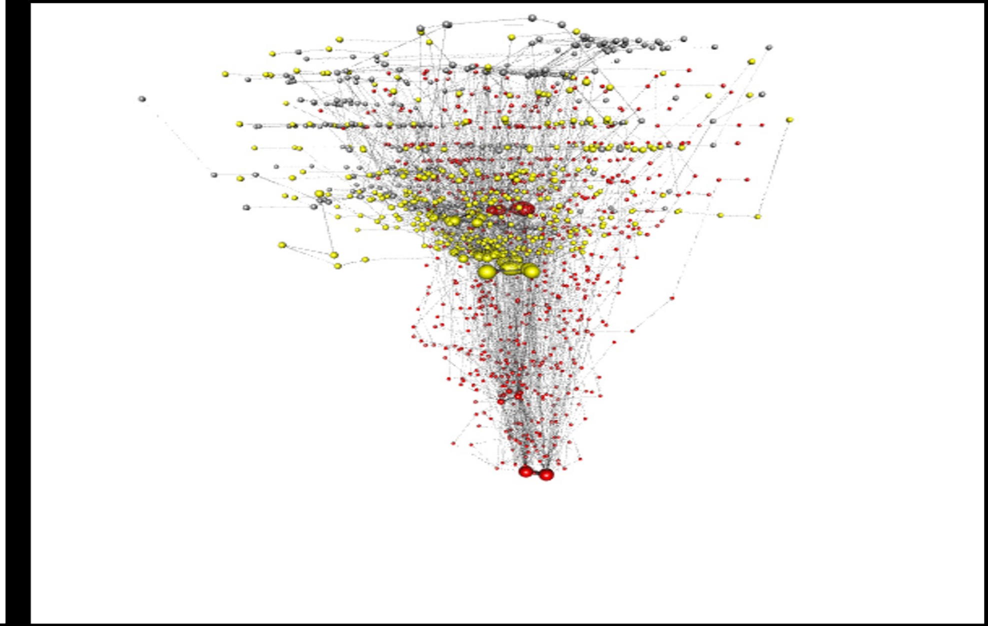
TSP INSTANCES



c755 Clustered Cities
Funnels: 1, Success: 100%

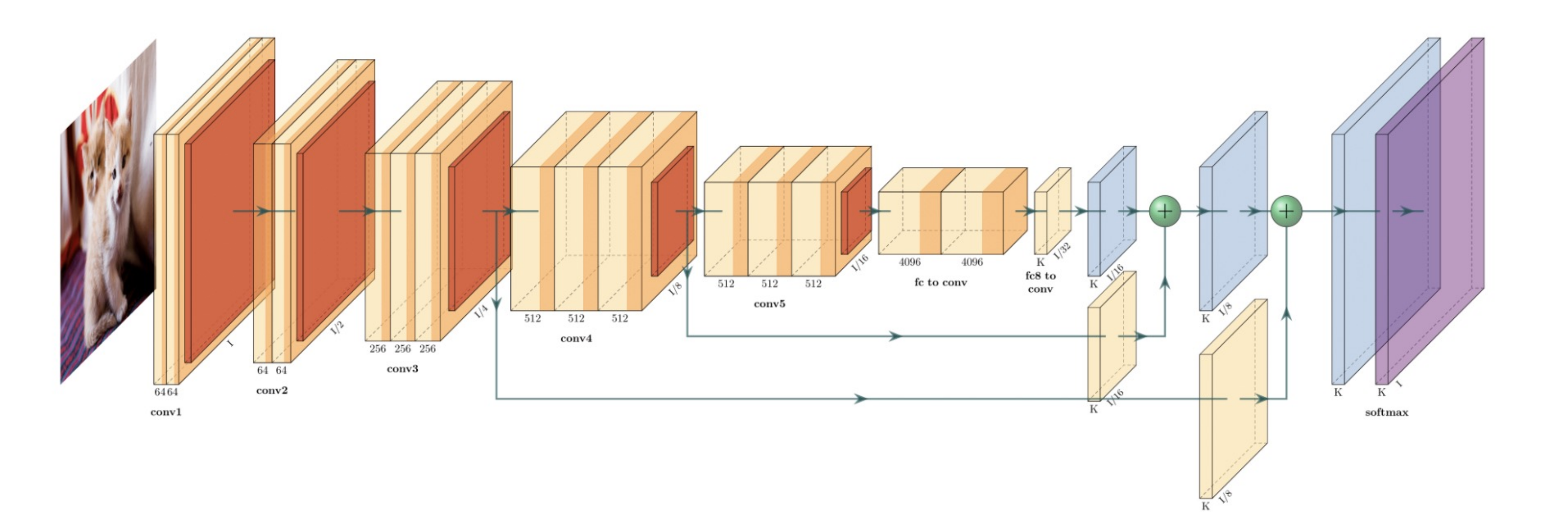


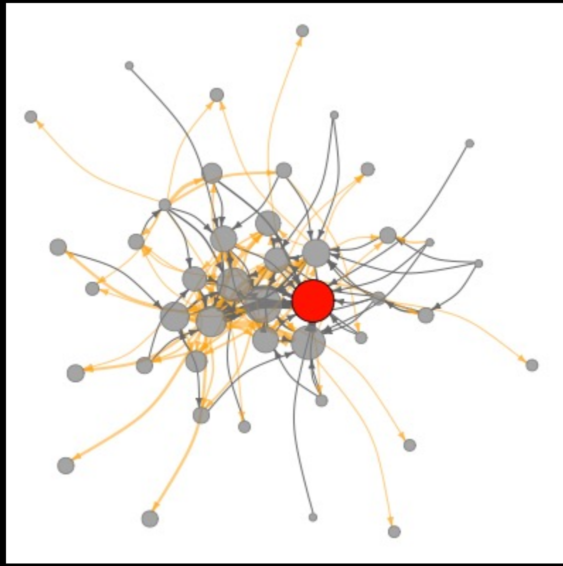
E755 Uniform Cities
Funnels: 4, Success: 13%



att532 (cities in the US)
Funnels: 2, Success: 44%

Mapping the global structure of TSP fitness landscapes
G Ochoa, N Veerapen (2018) *Journal of Heuristics*

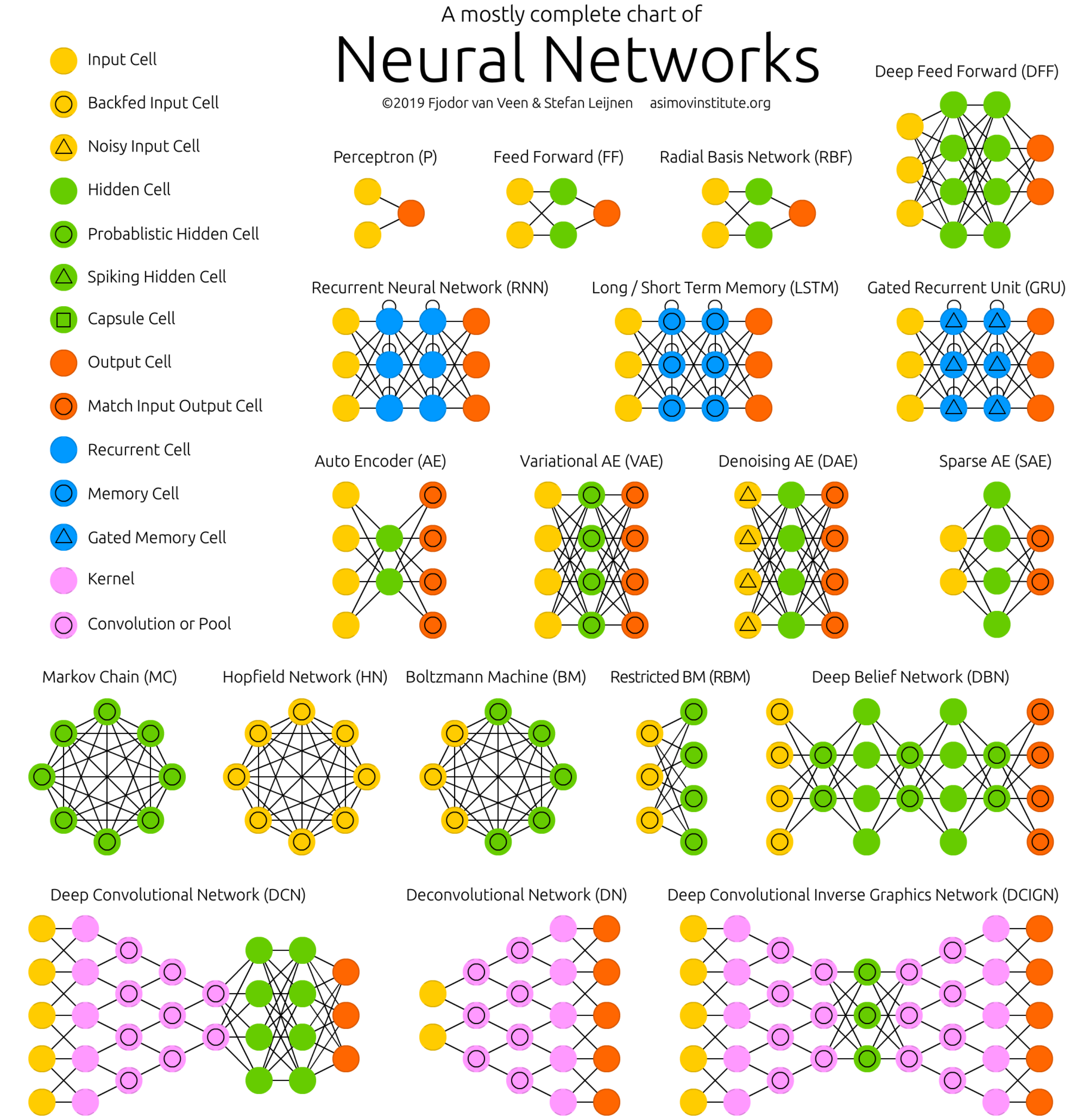


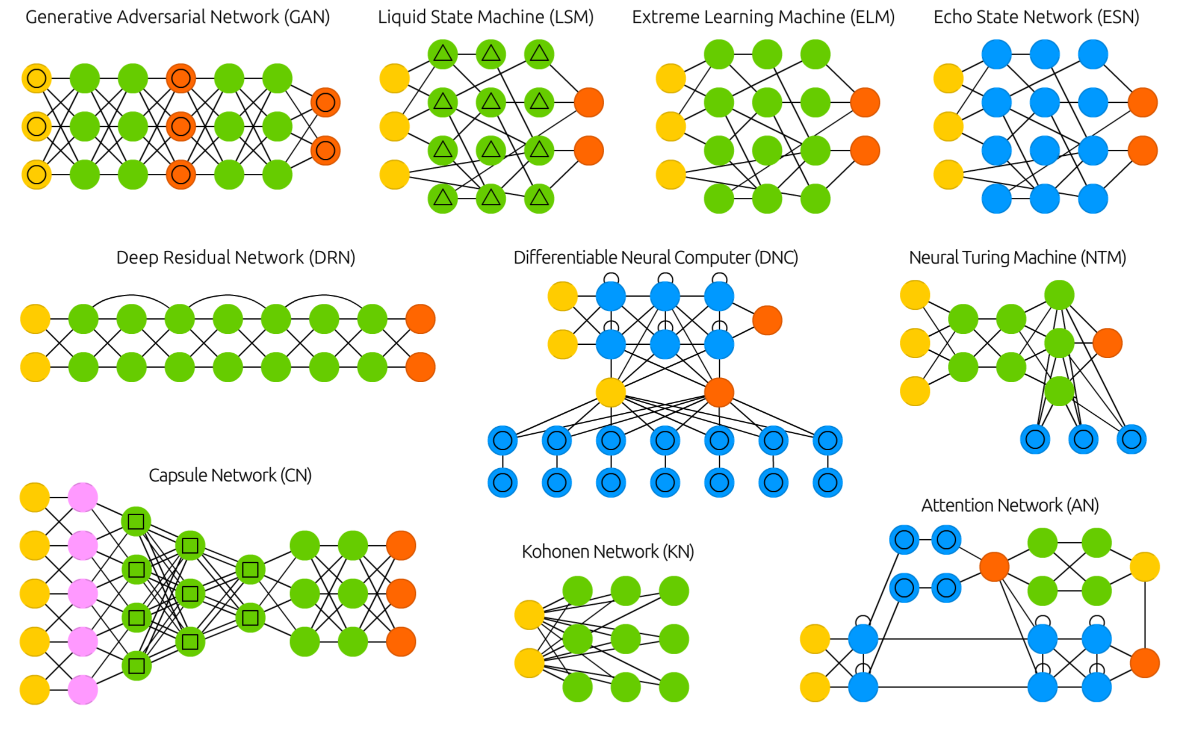


LONS FOR NEURAL ARCHITECTURE SEARCH

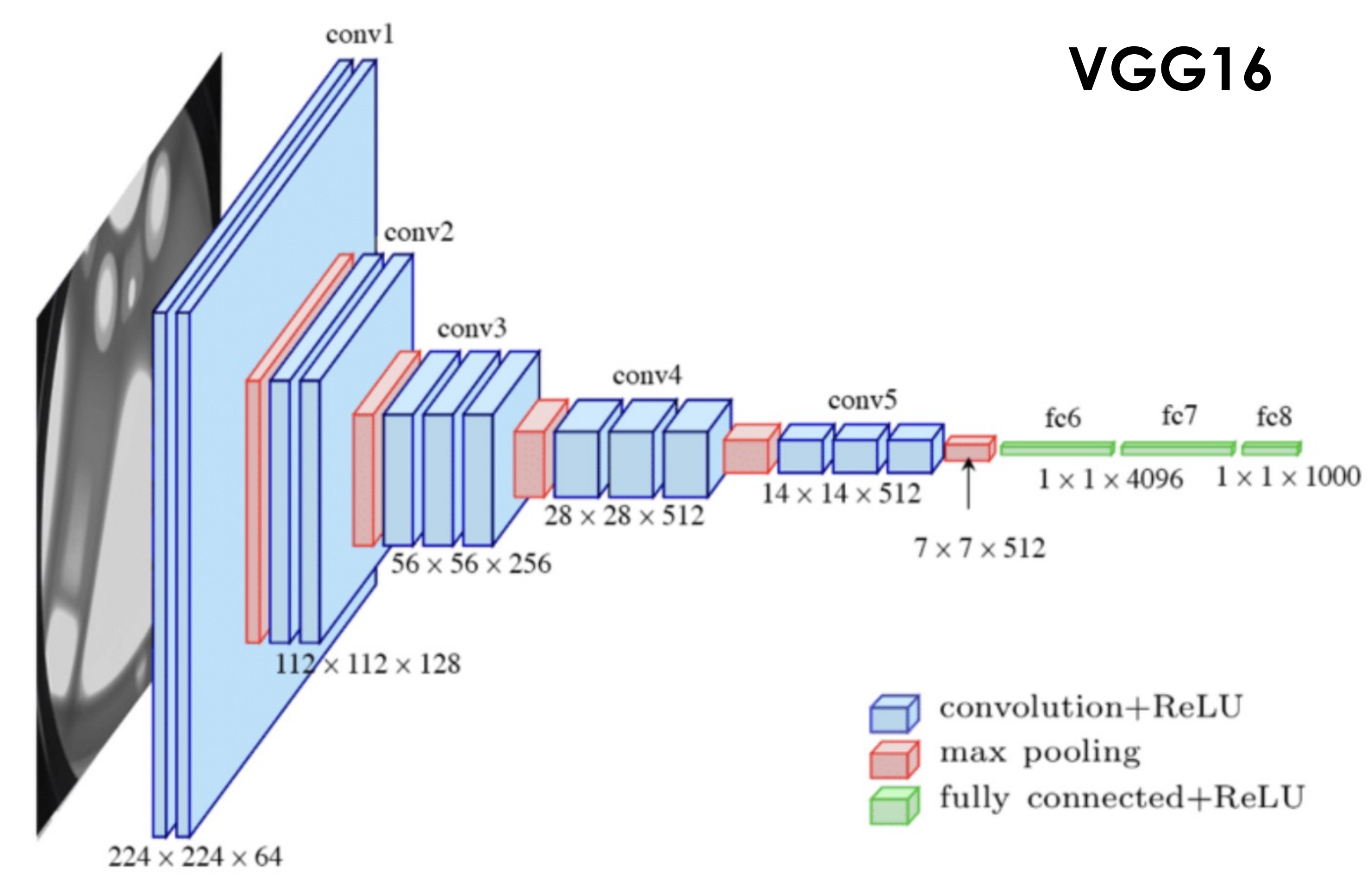
THE NEURAL NETWORK ZOO

"With new neural network architectures popping up every now and then, it's hard to keep track of them all." Fjodor Van Veen
https://www.asimovinstitute.org/neural-network-zoo/





DEEP CNN ARCHITECTURE



A classic convolutional NN architecture

- Total of 16 layers with weights
- Over 130 Million weights!
- 5 blocks: convolution layers follow ed by a max-pooling layer
- Resolution halved after each

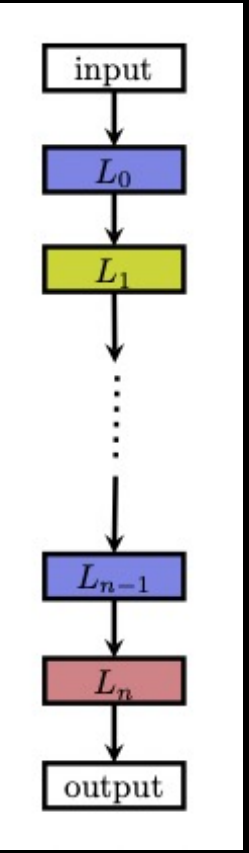
VGG: Visual Geometry Group
Karen Simonyan and Andrew Zisserman
University of Oxford, 2014
Winner of the Visual Recognition Challenge

NEURAL ARCHITECTURE SEARCH (NAS)

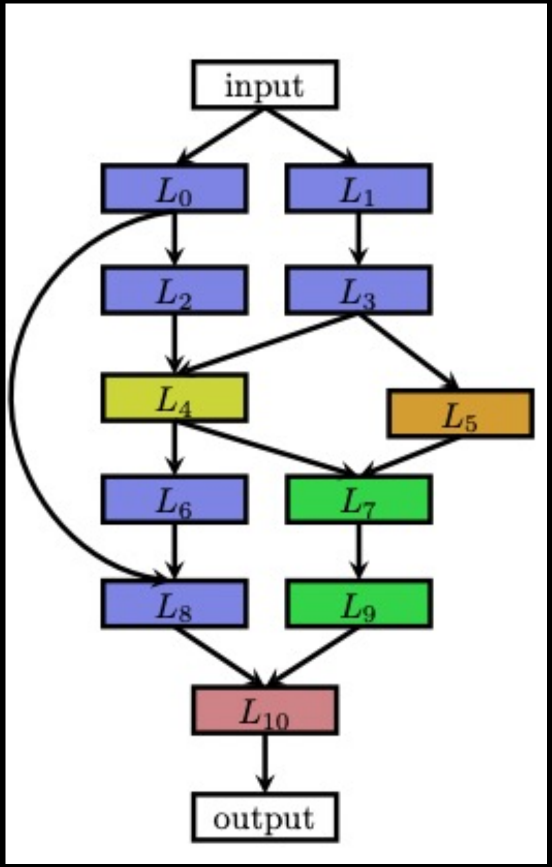
- Use of search heuristics to optimize the topology of deep neural networks
- Has produced models that surpass the performance of huma-designed models
- Formulated as a discrete optimisation problem.

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

ALTERNATIVE NAS SEARCH SPACES

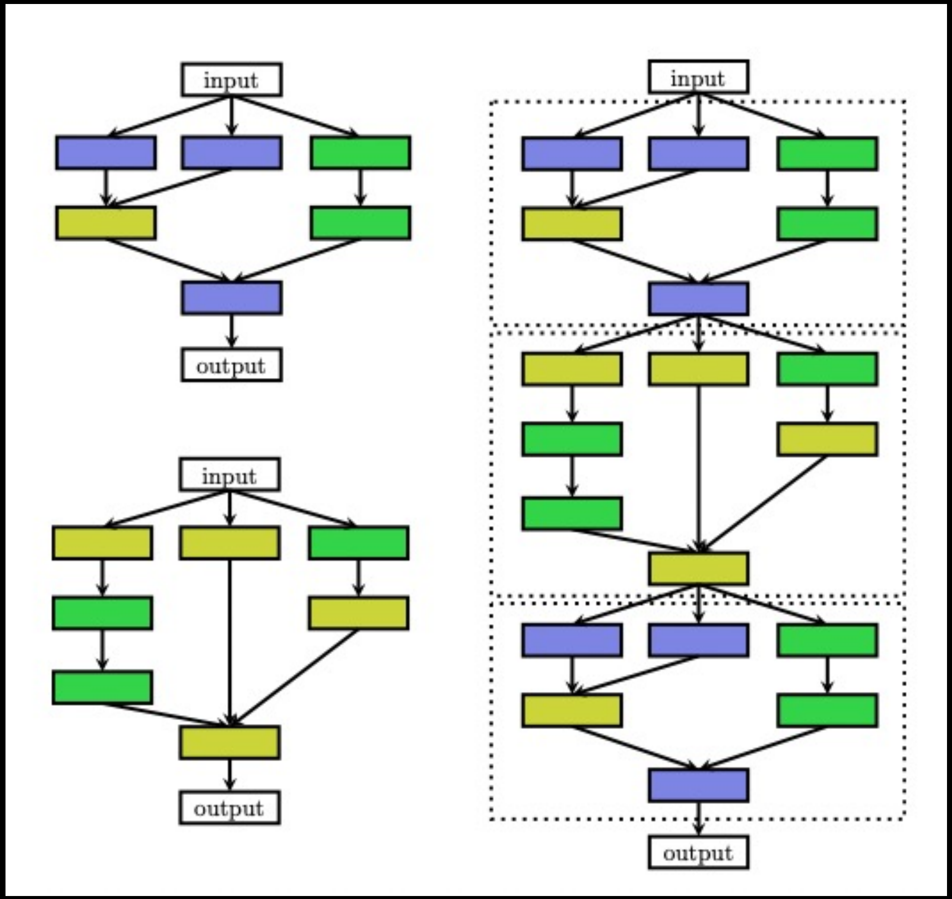


Chain Structure



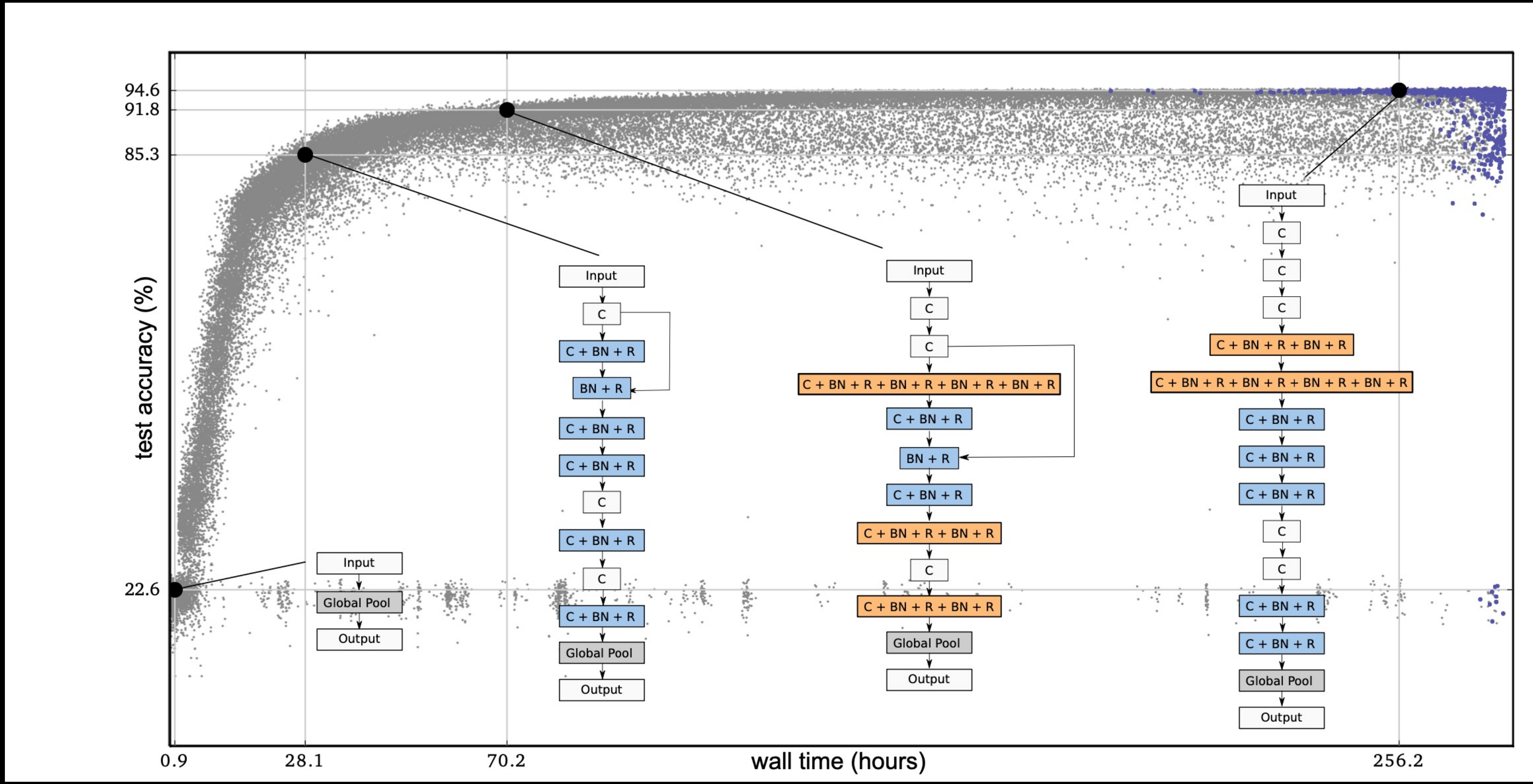
Multi-branch Structure

Two cells: Normal & Reduction



Cell (Block) Search Space

EXAMPLE OF SEARCH PROGRESS



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

CELL BASED TABULAR BENCHMARK

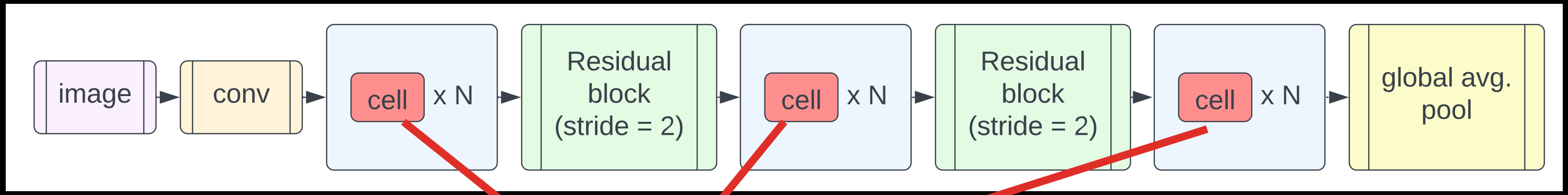
Dong, X. et. al. (2021) NATS-Bench: Benchmarking NAS algorithms for architecture topology and size. *IEEE Trans. Pattern Analysis and Machine Intelligence*

Tabular benchmark (NATS-Bench - Image classification)

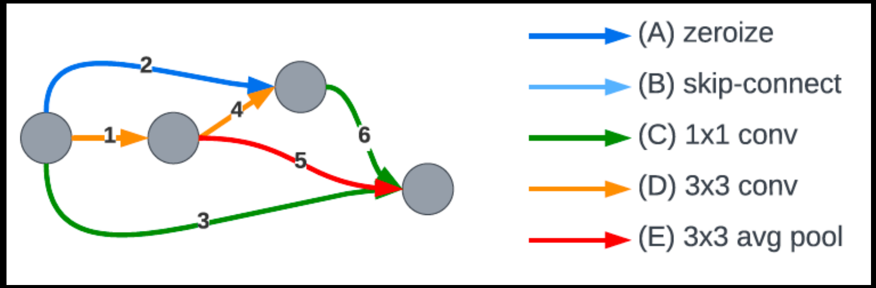
Cell-based search space

Fitness Landscape Analysis (DOS, FDC, LONS)

Impact of noise in performance estimation

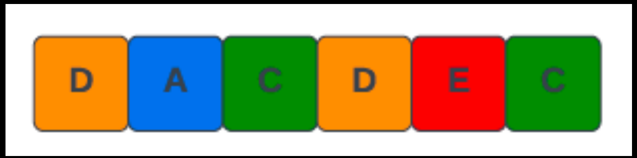


Fixed macro-skeleton



Cell (DAG)

Genotype



Discrete optimisation problem

max $a \in A$ $f(a)$

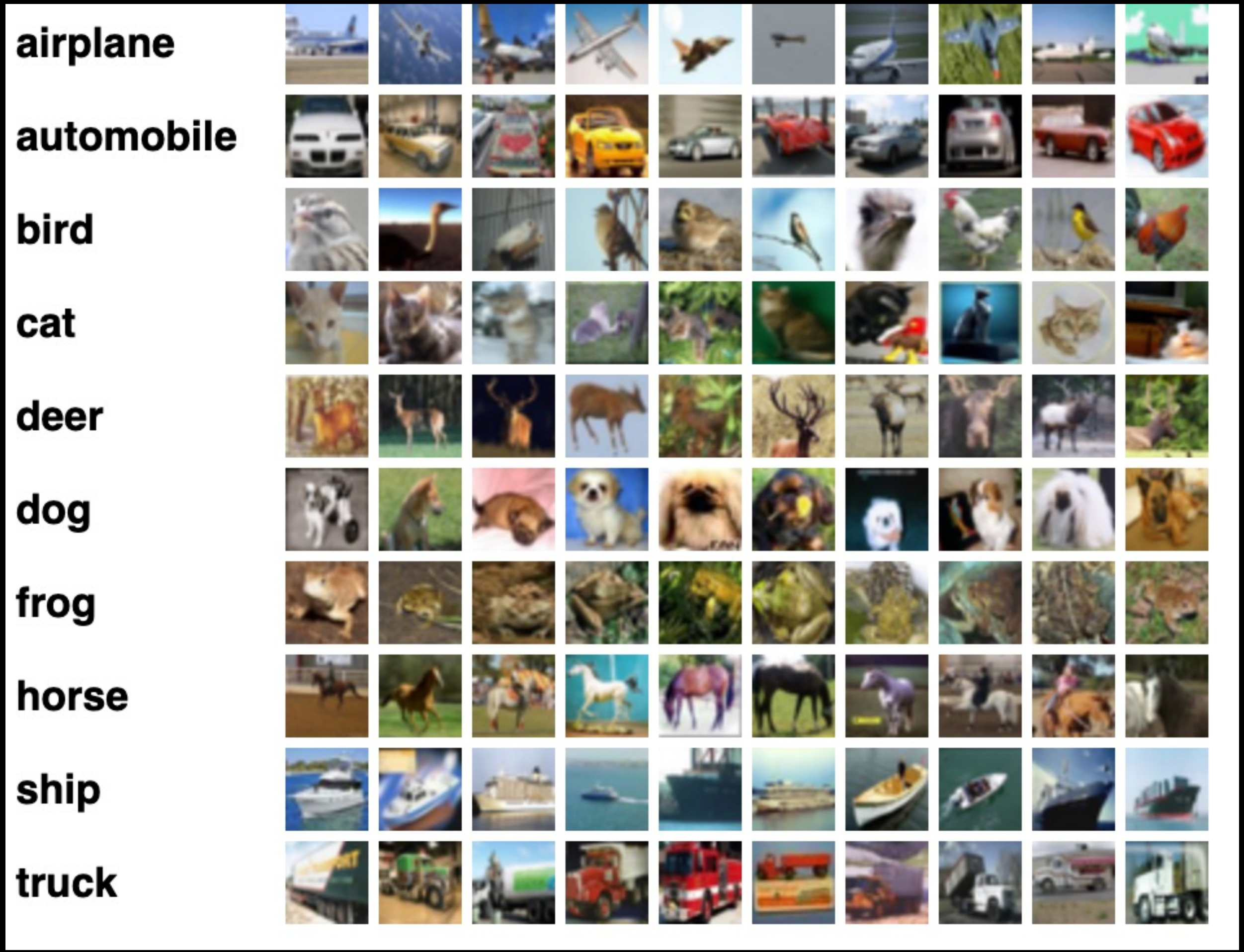
$f(a)$ validation accuracy

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

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

IMAGE CLASSIFICATION DATASETS

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

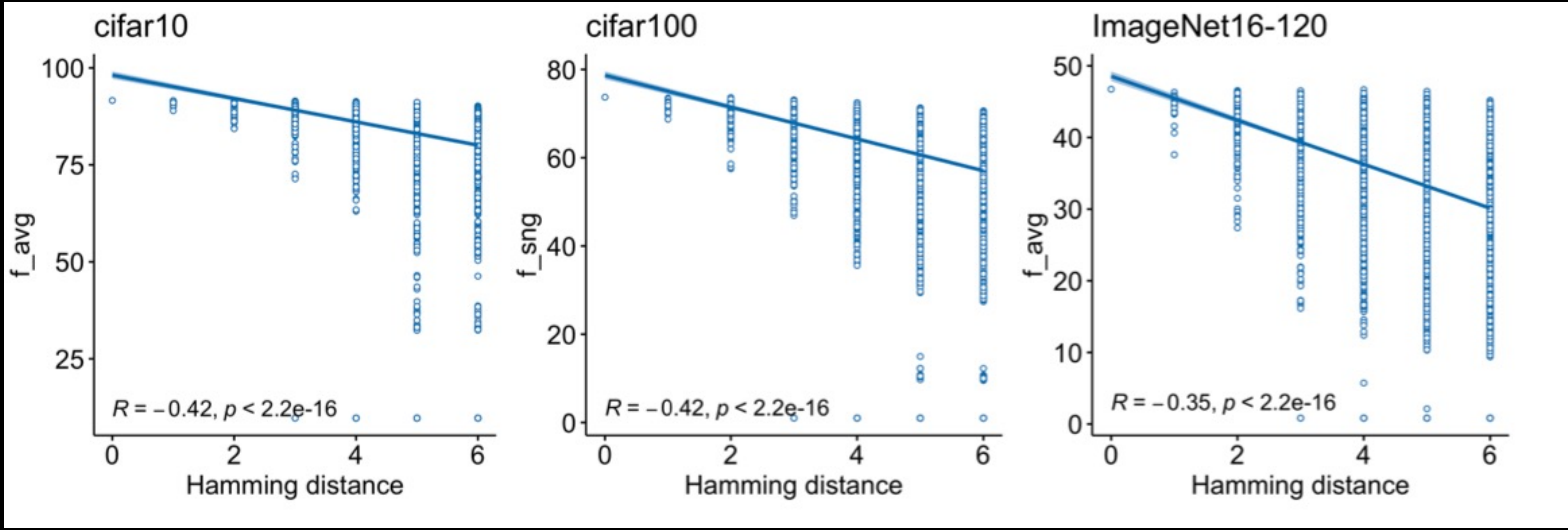


https://www.cs.toronto.edu/~kriz/cifar.html

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

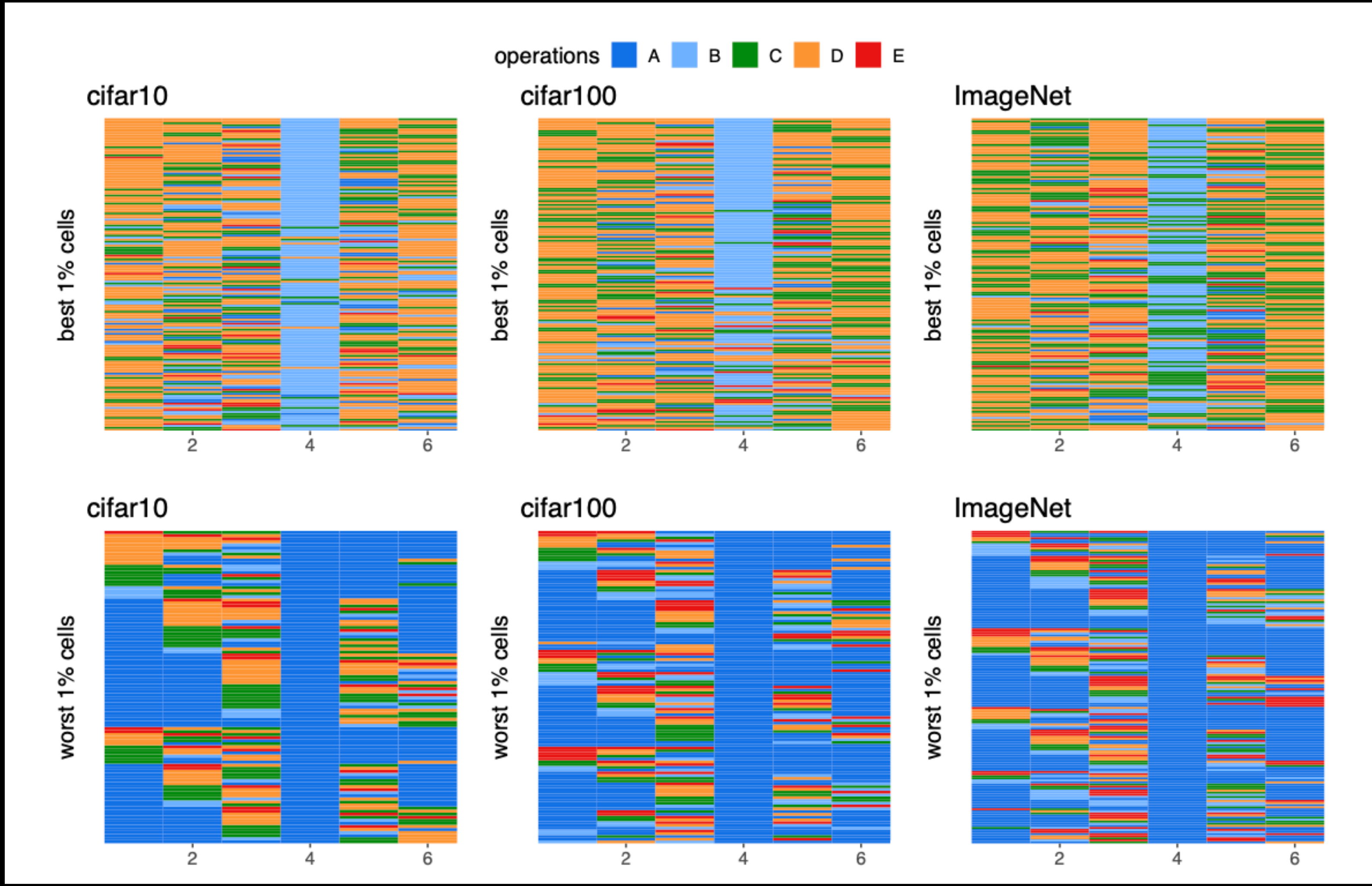
https://cs.stanford.edu/people/karpathy/cnnembed/

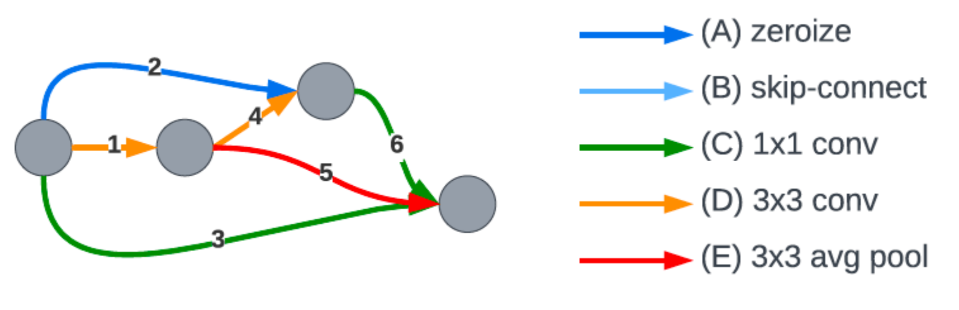
FITNESS DISTANCE CORRELATION



Horizontal axes -Hamming distance to the global optimum,
The Spearman correlation coefficients with p-value are also shown.

GENOTYPE MAPS

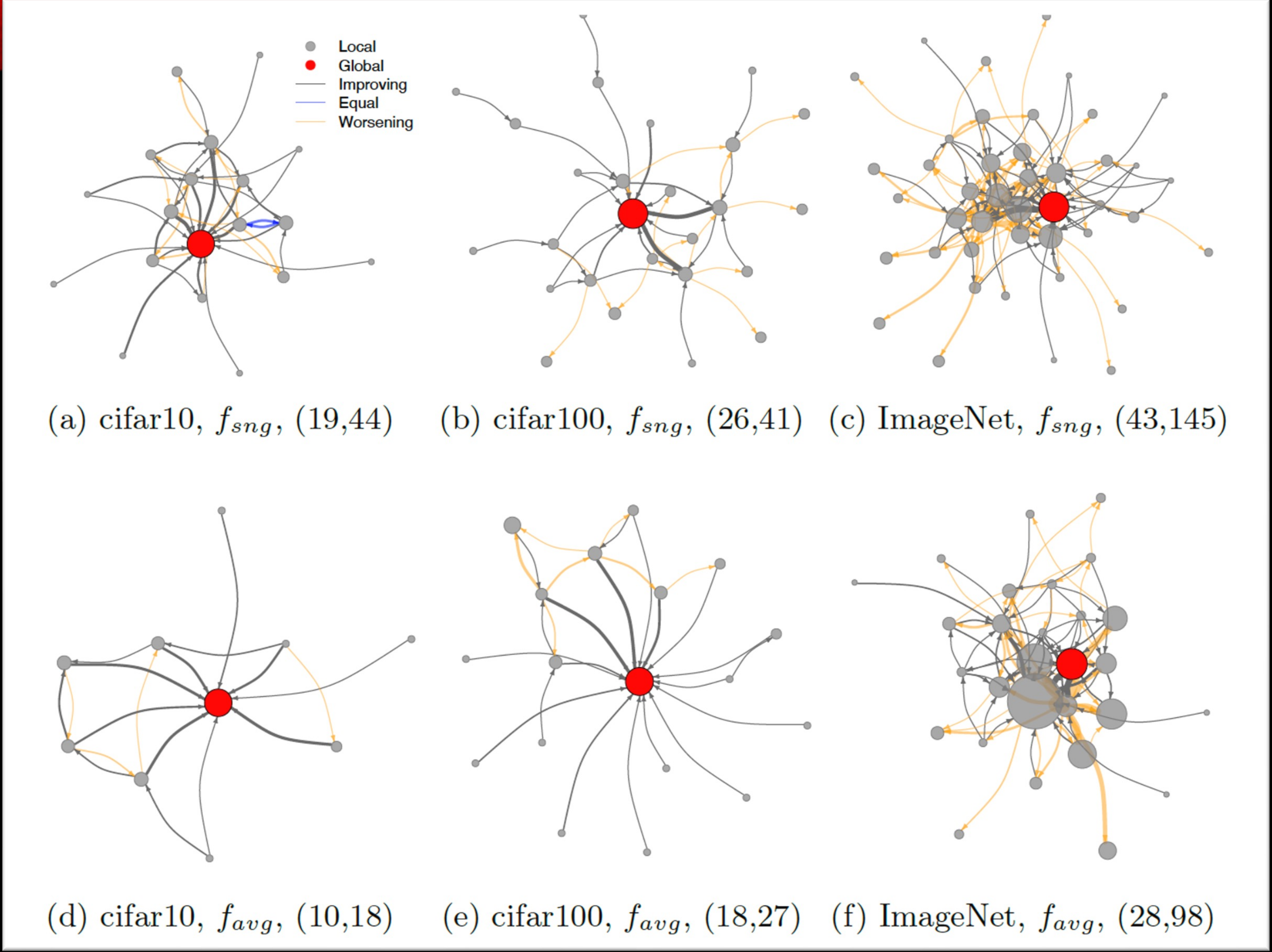




Genotype maps of the best 1% (top), and worst 1% (bottom) performing cells, sorted according to *f_avg*.

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

LOCAL OPTIMA NETWORKS



(a) cifar10, f_{sng}, (19,44) (b) cifar100, f_{sng}, (26,41) (c) ImageNet, f_{sng}, (43,145)

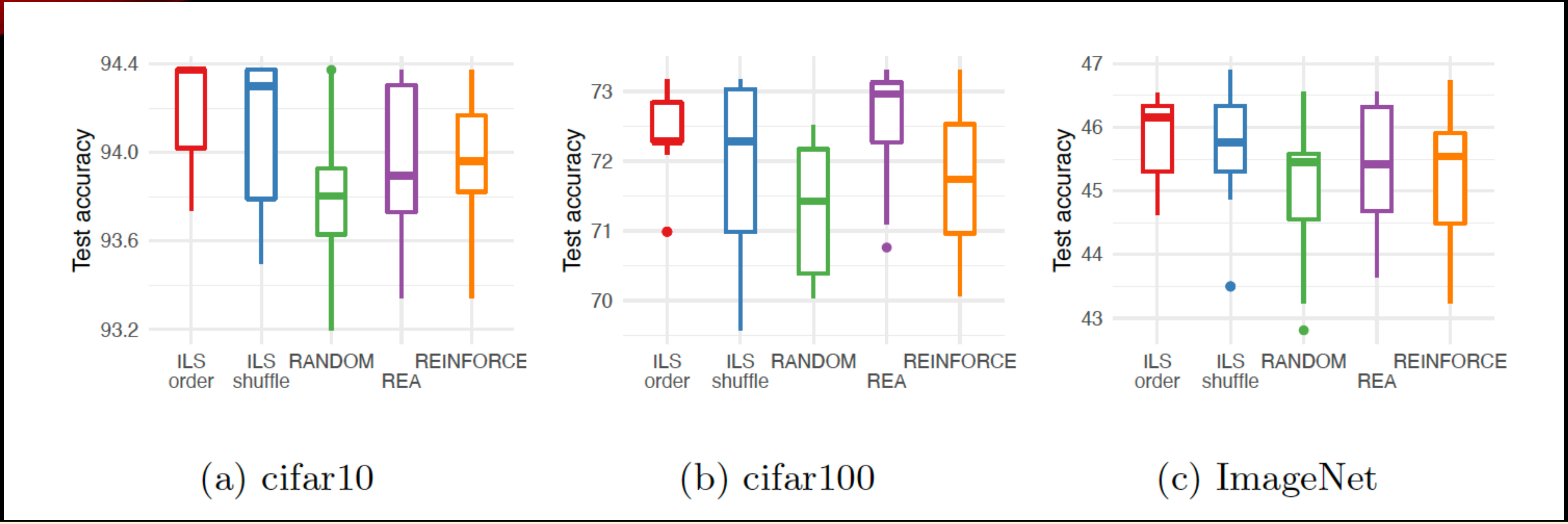
(d) cifar10, f_{avg}, (10,18) (e) cifar100, f_{avg}, (18,27) (f) ImageNet, f_{avg}, (28,98)

Local optima networks (**LONs**) for all datasets and the two fitness functions.

The number of nodes *n* and edges e are indicated as (*n, e*).

Size of nodes: proportional to incoming degree

Distribution of average test accuracy at the end of the run
Progression of average test accuracy



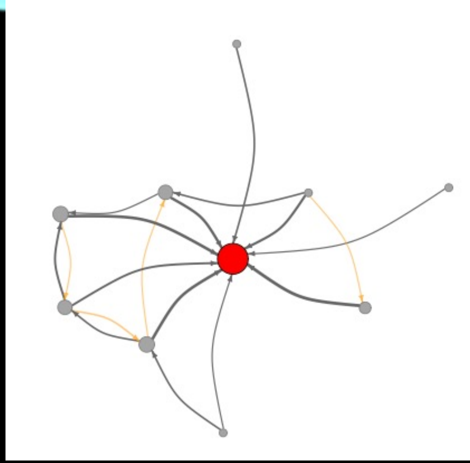
30 Runs
For each strategy

5 Search Strategies

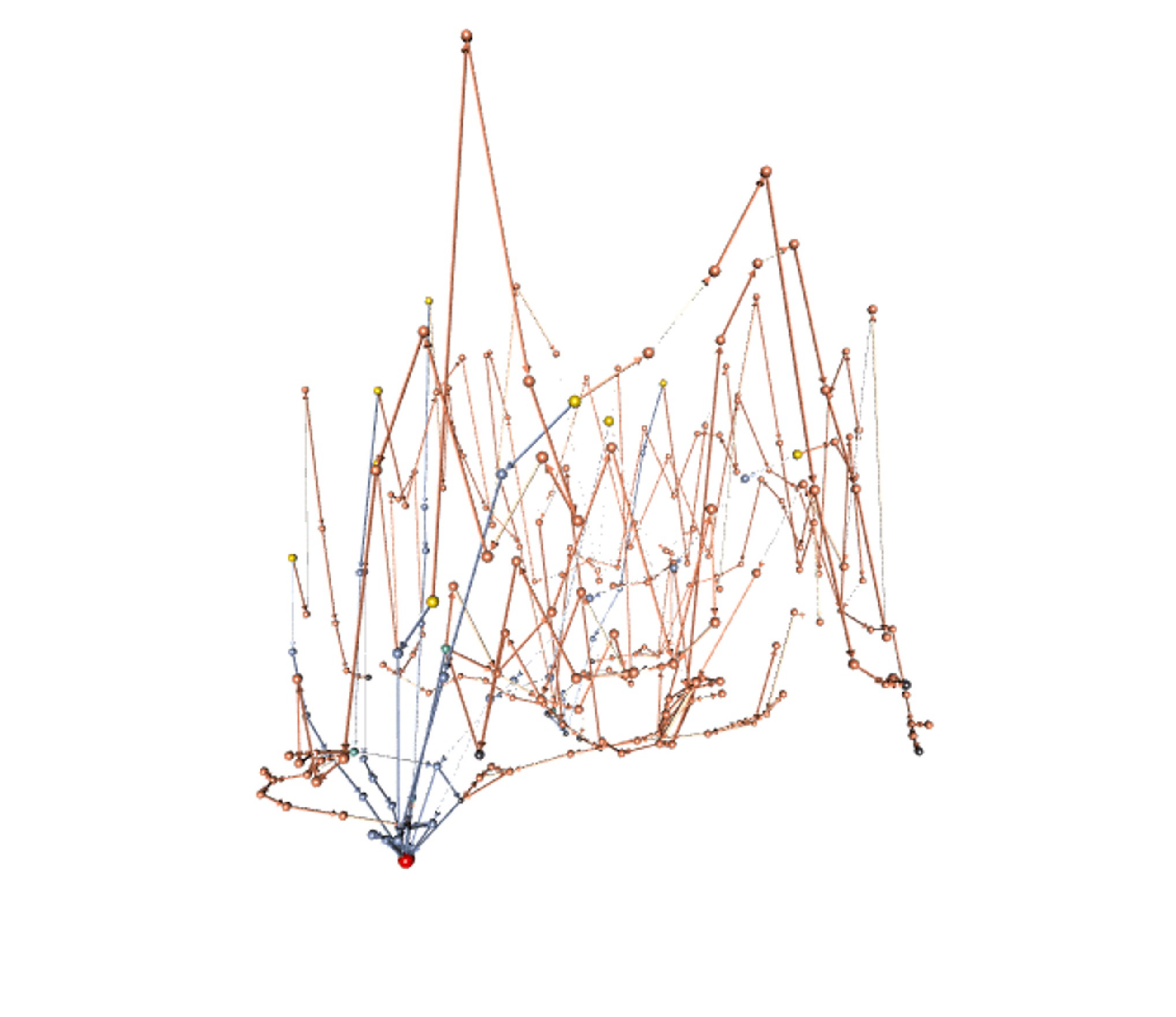
- Two versions of Iterated Local Search (ILS) (order, shuffle) – order to explore neighbours
- Random Search, Regularized Evolutionary Algorithm, Reinforcement Learning

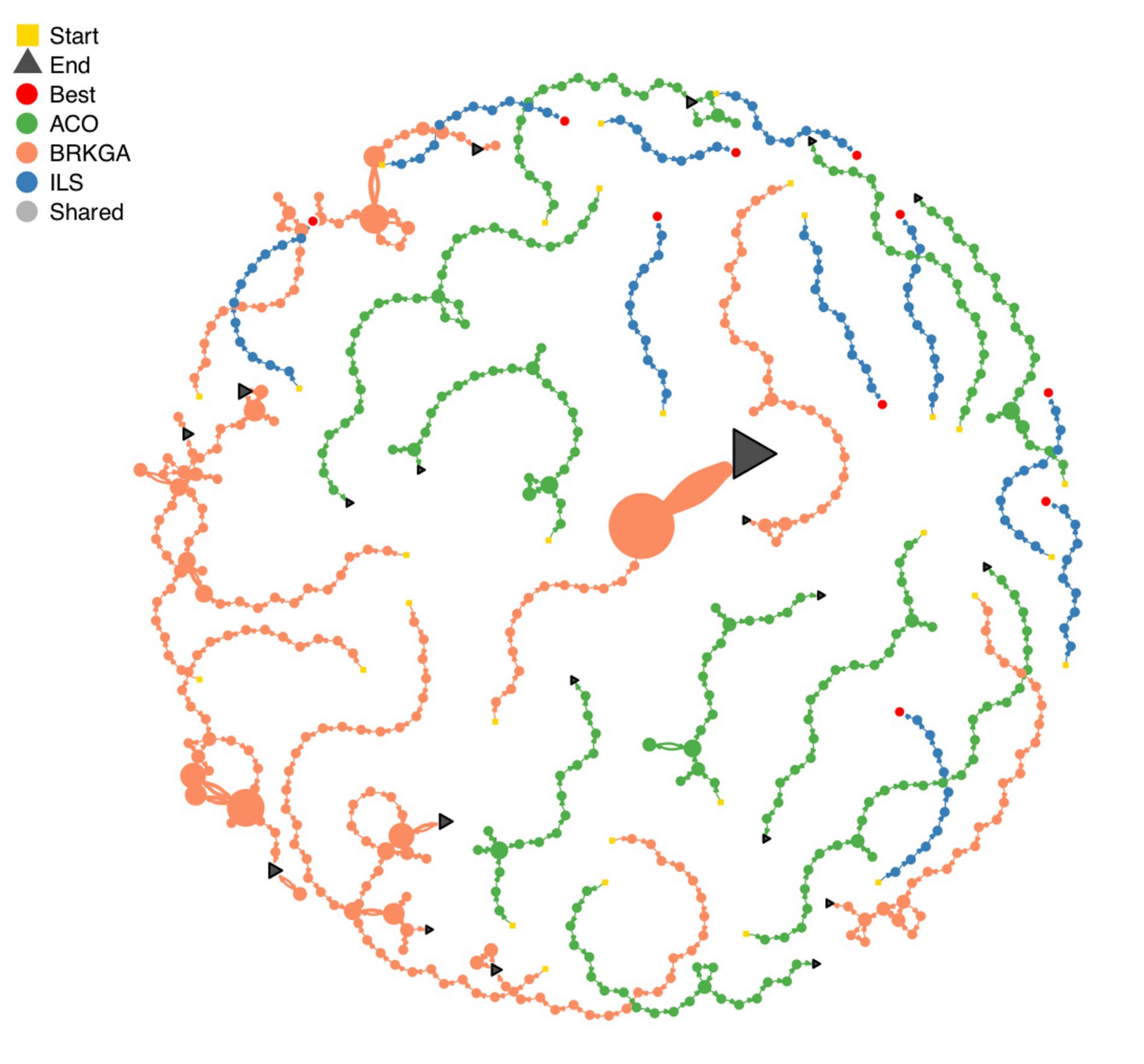
SEARCH STRATEGIES

INSIGHTS FROM LONs IN NAS



- NAS landscapes are rugged (multi-modal)
- Low number of local optima, from which it is not difficult to escape
- Reducing noise reduces the number of local optima
- ILS can outperform more elaborate evolutionary and reinforcement learning methods!
- No crossover in state-of-the-art EA method!

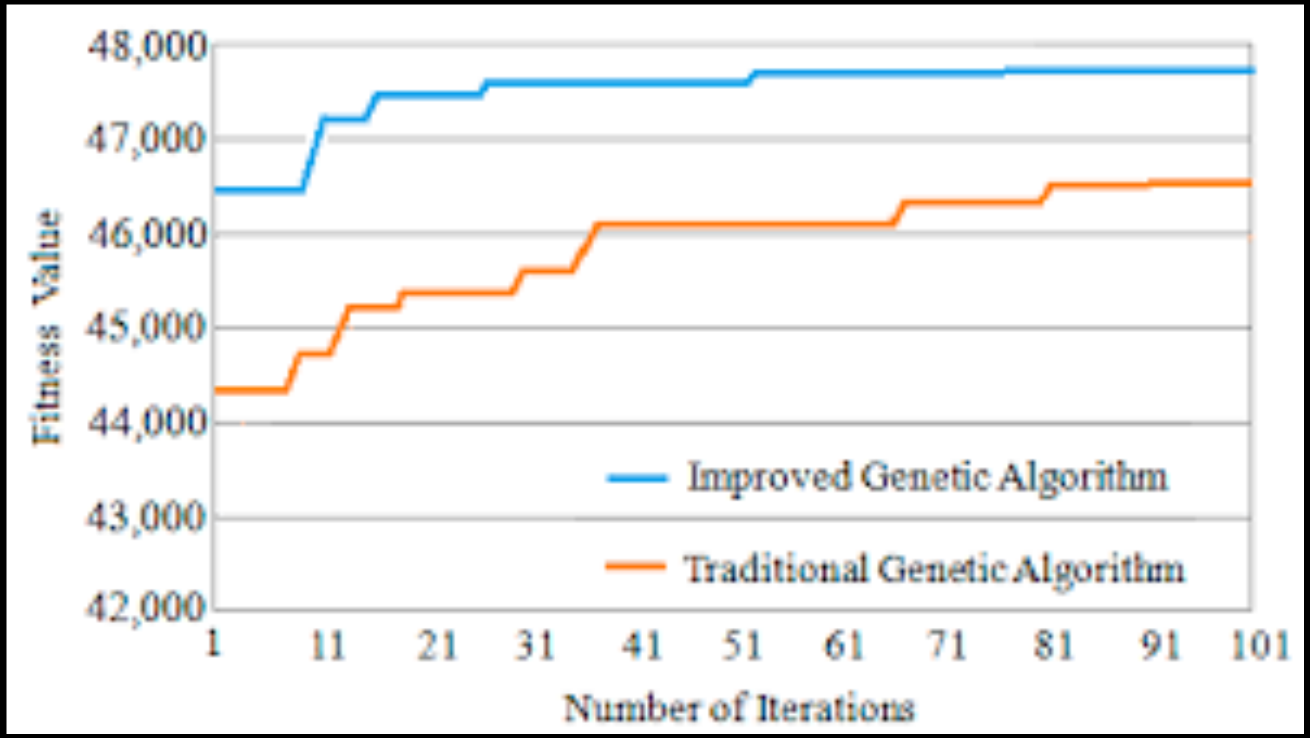




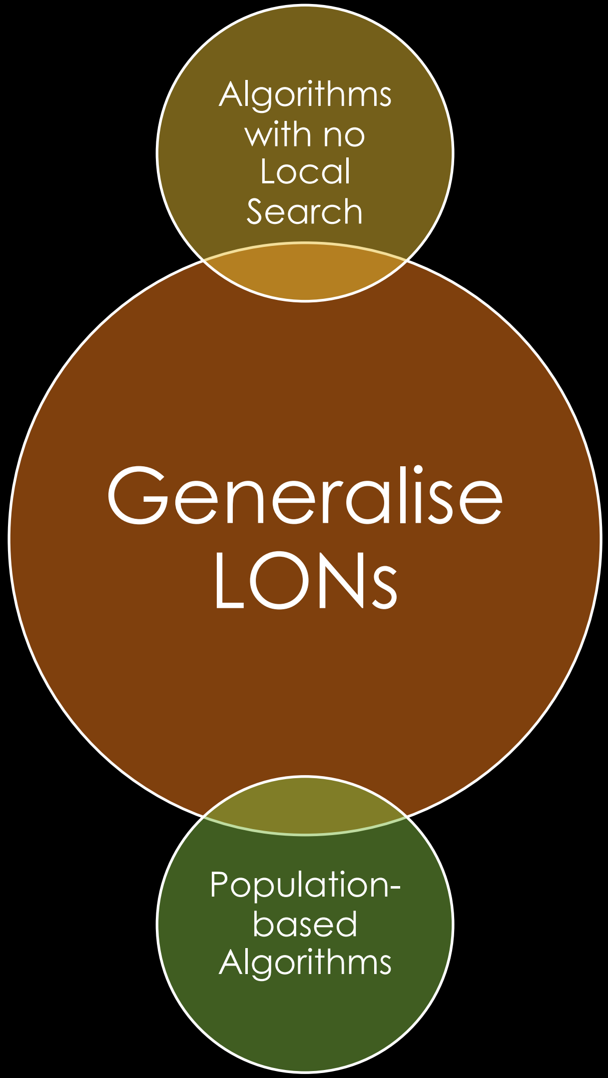
SEARCH TRAJECTORY NETWORKS

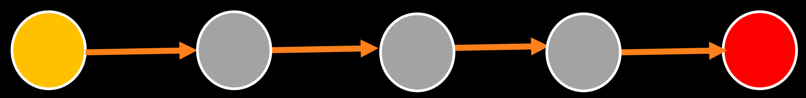
MOTIVATION FOR STNS

Few tools for studying/contrasting behaviour



Typical convergence plot – Only shows dynamics on the objective space



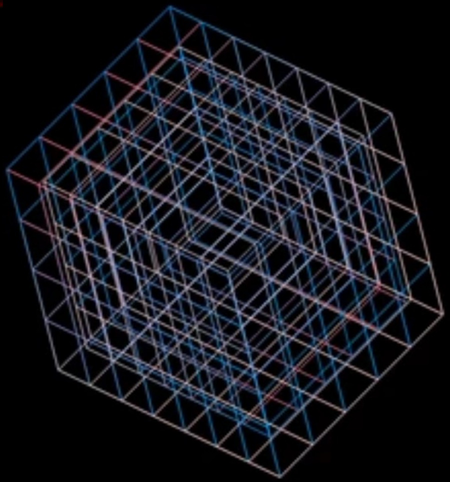


STN DEFINITIONS

Representative Solution	• Solution that represents status of the search process
Location	• Subset of solutions, partition of the search space
Nodes	• Locations of representative solutions
Edges	• Directed, connect two consecutive locations
STN	• Directed graph STN = (N, E)

MAPPING SOLUTIONS TO LOCATIONS

Solution space is divided into hypercubes



Example D = 3

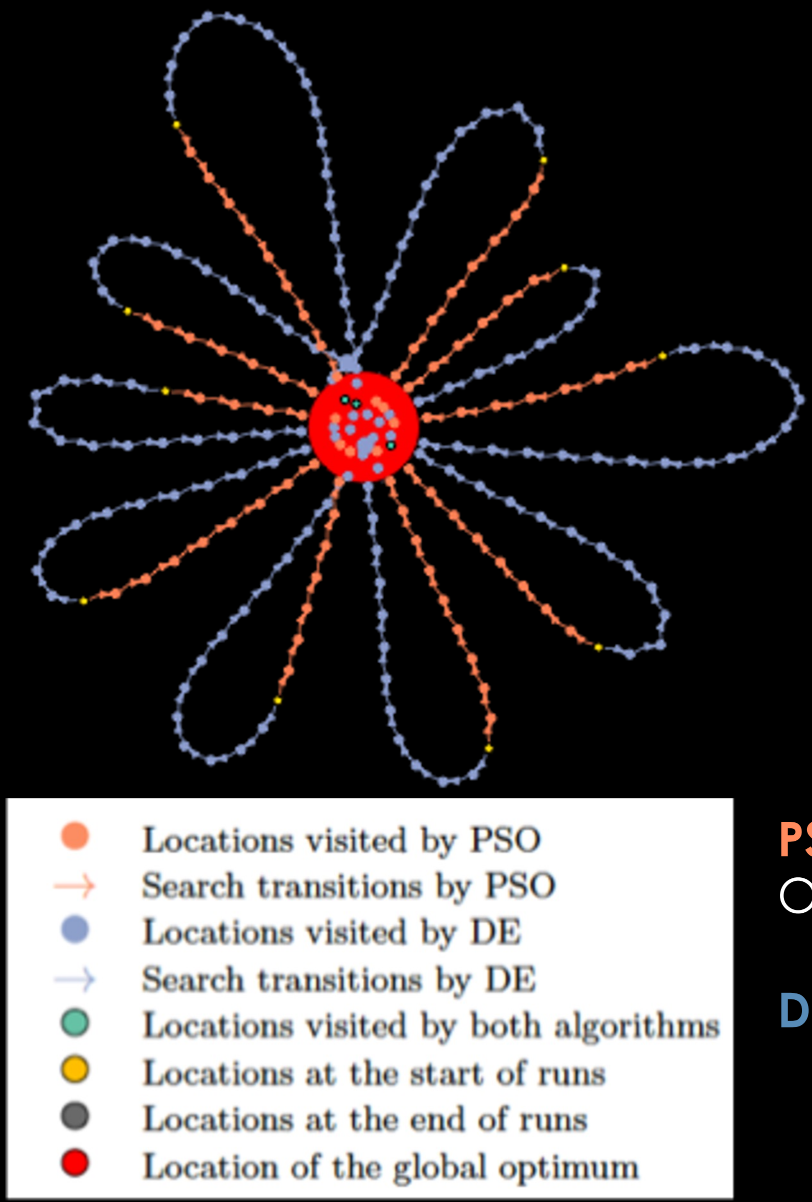
Solution
- [12.2029055, 1.570796323, 1.284991569]

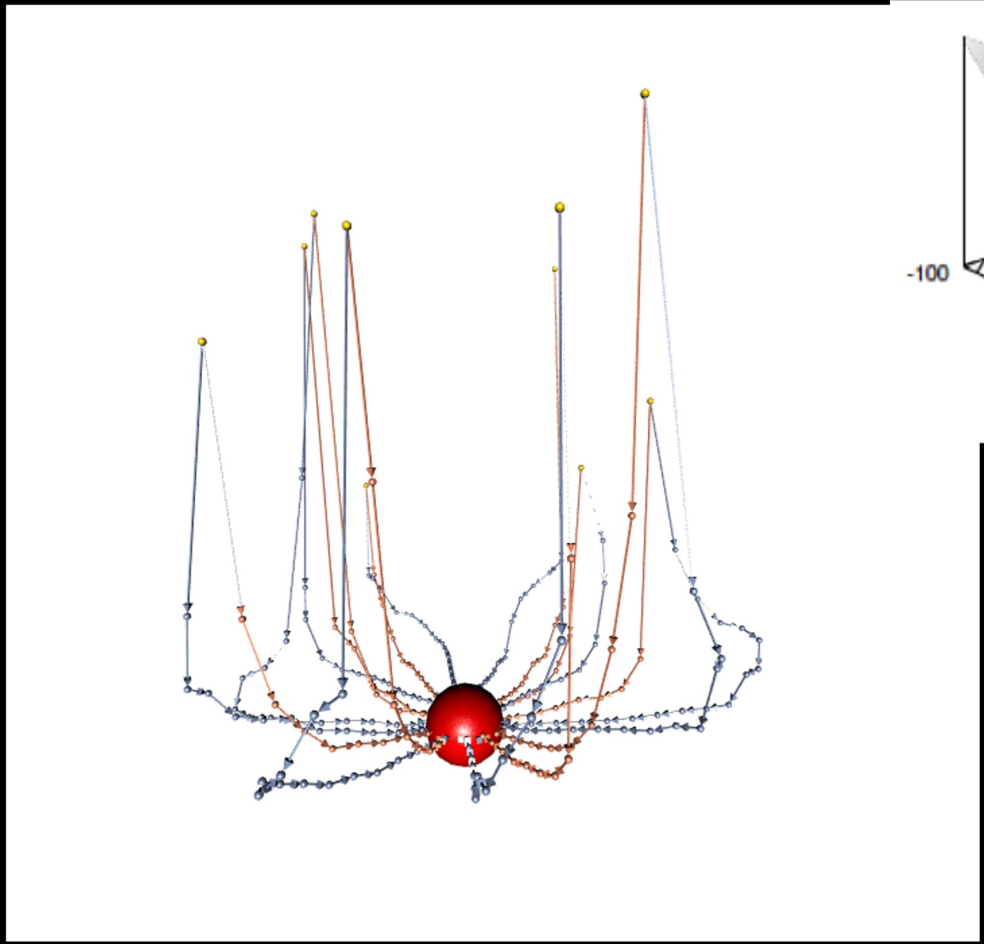
Rounding
- Precision (0.1)
- [12.2, 1.6, 1.3]

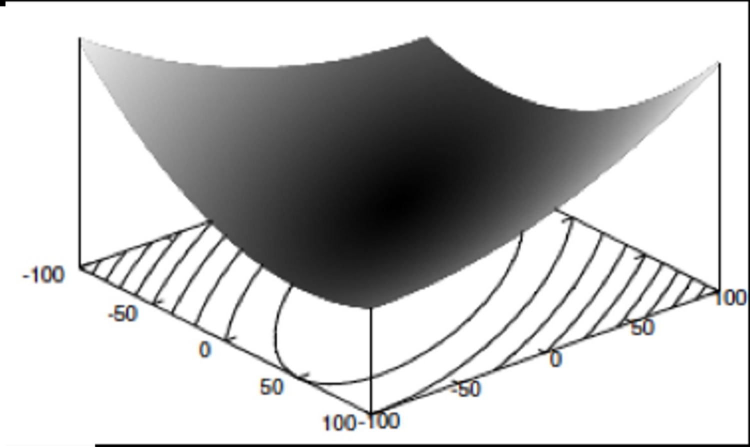
Integer
- [122, 16, 13]

Concatenate
- ID of node
- 122.16.13

QUADRIC, D = 10







D = 2

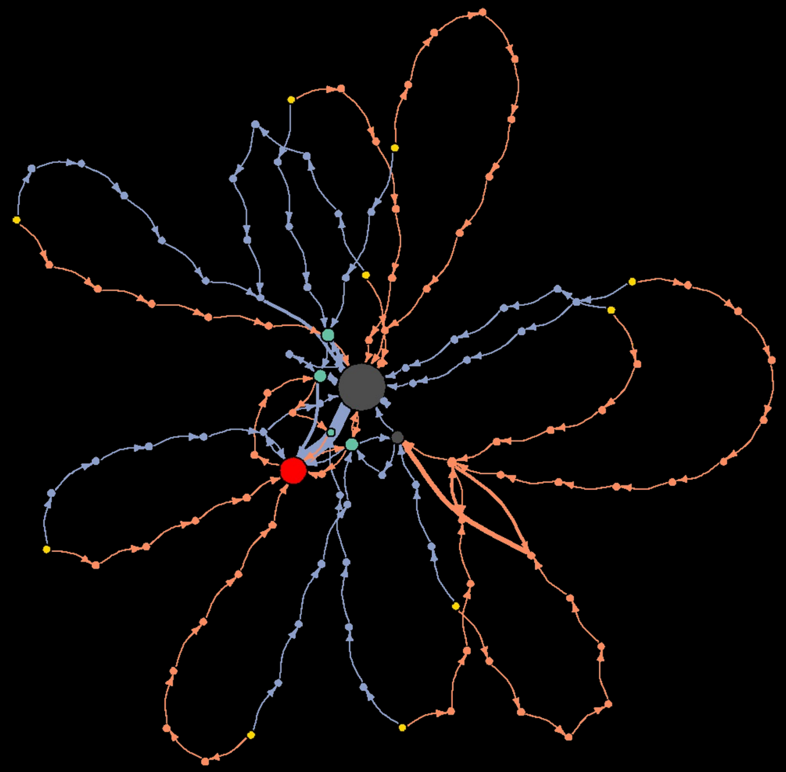
Unimodal function

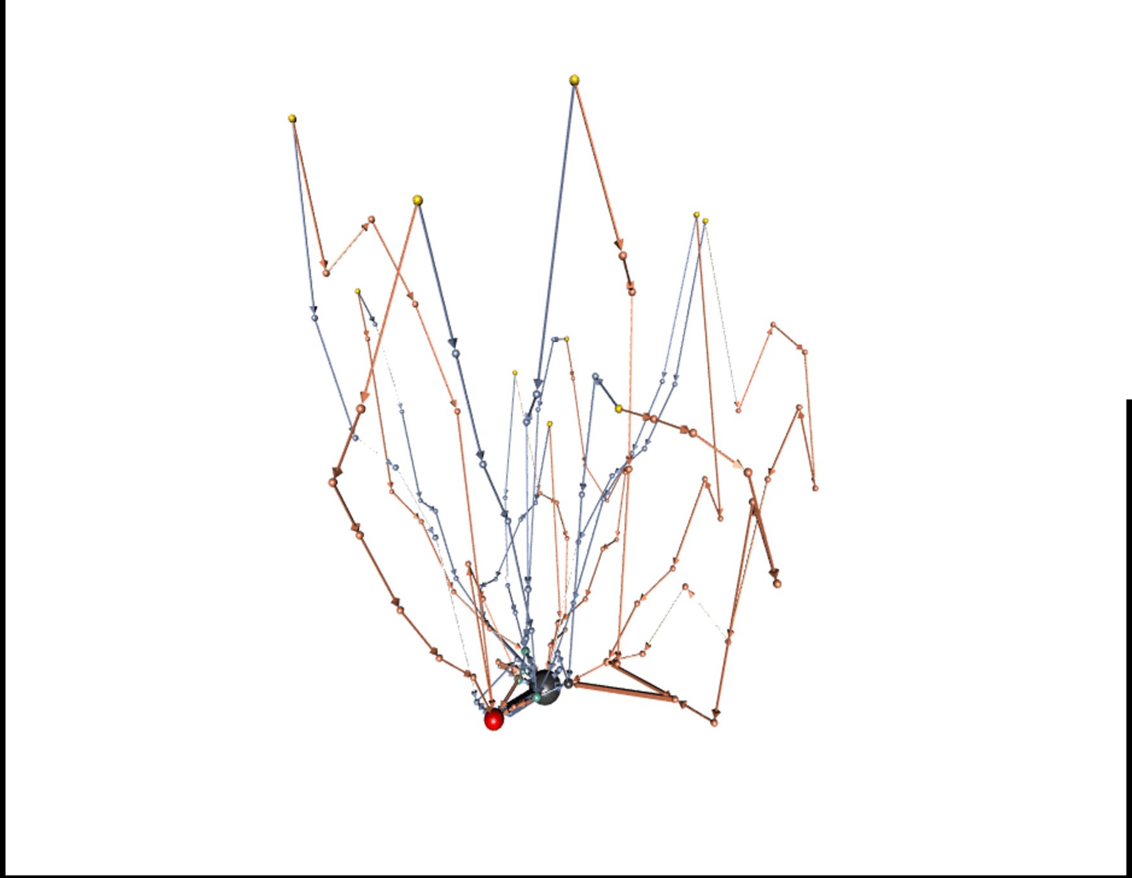
PSO – Particle Swarm Optimisation

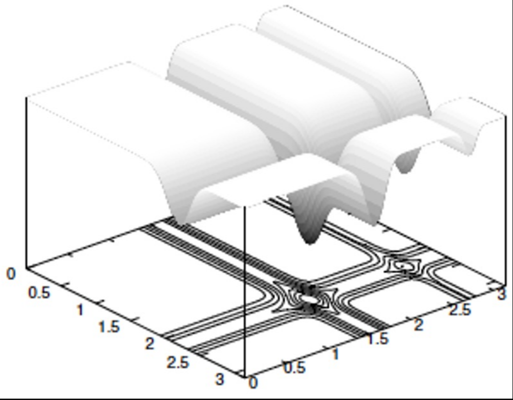
DE – Differential Evolution

	PSO	DE
Nodes	93	154
Path length	9.2	15.1
Success	100%	100%

Michalewicz, D = 5







Multimodal, large plateaus close to the global optimum

- Locations visited by PSO
- Search transitions by PSO
- Locations visited by DE
- Search transitions by DE
- Locations visited by both algorithms
- Locations at the start of runs
- Locations at the end of runs
- Location of the global optimum

PSO – Particle Swarm Optimisation

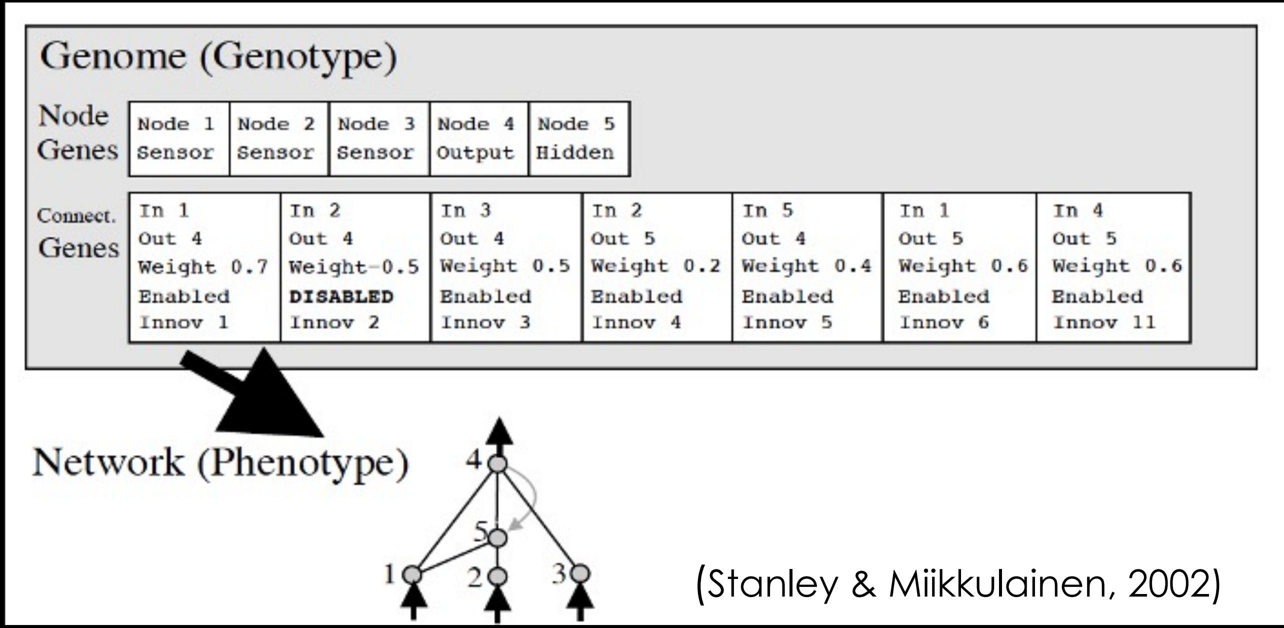
DE – Differential Evolution

	PSO	DE
Nodes	76	61
Path length	7.5	7.0
Success	40%	80%

NEUROEVOLUTION

- Evolve weights, fixed topology
- Evolve topology, learn weights with gradient-based methods
- Evolve both weights and topology
- Evolve components, hyperparameters, learning rules

NEAT – Neuroevolution of Augmented Topologies



(Stanley & Miikkulainen, 2002)

NEAT STNs

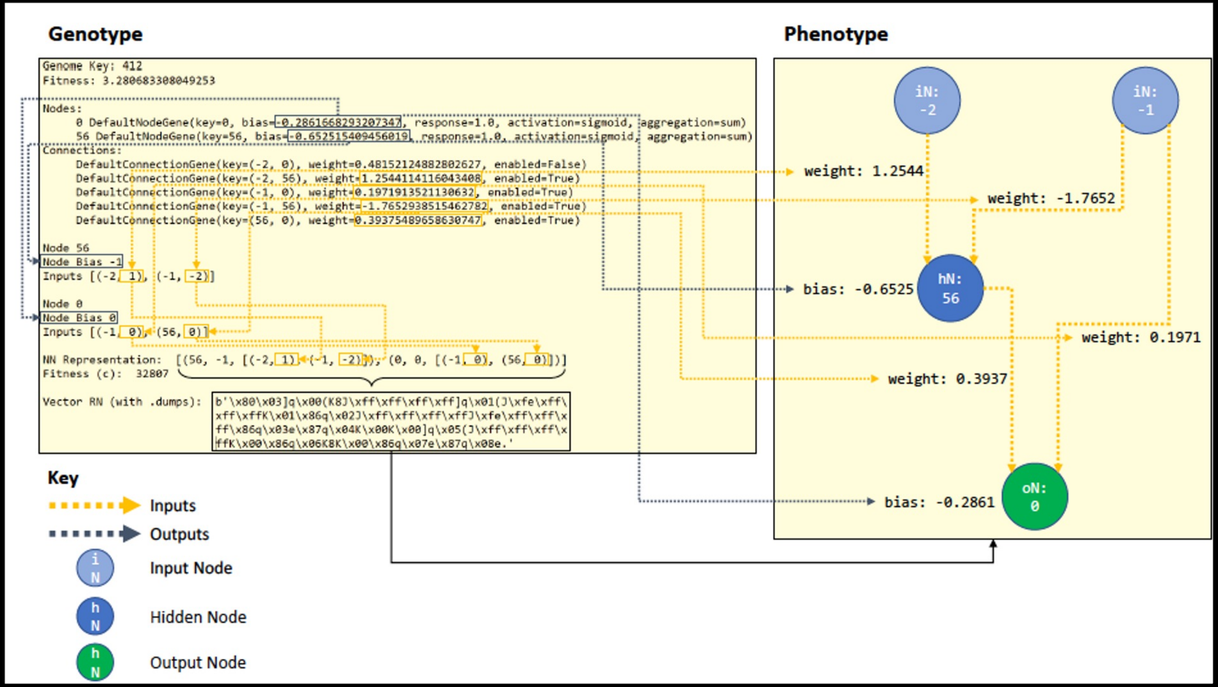
Neuroevolution of Augmented Topologies

Benchmarks

- Maze Navigation
- Deceptive problems

Algorithm Variants

- Novelty Search
- Crossover
- No Crossover



Mapping NEAT genotypes to locations

Object Serialisation (pickle.dumps)

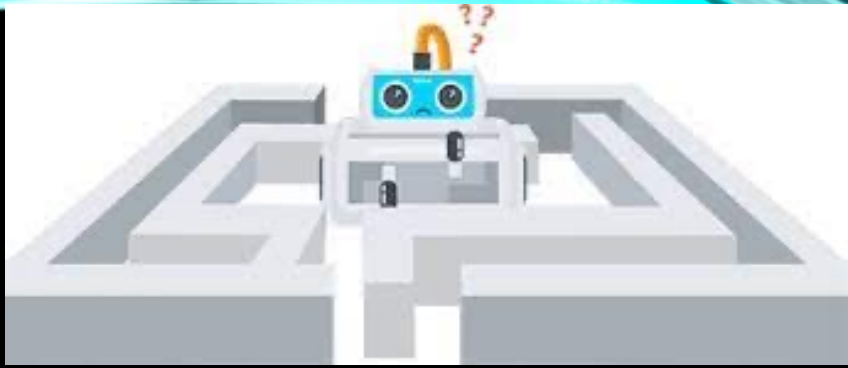
S Sarti, JAdair, G Ochoa (2022) Recombination and Novelty in Neuroevolution: A Visual Analysis. *SN COMPUT. SCI.* 3, 185

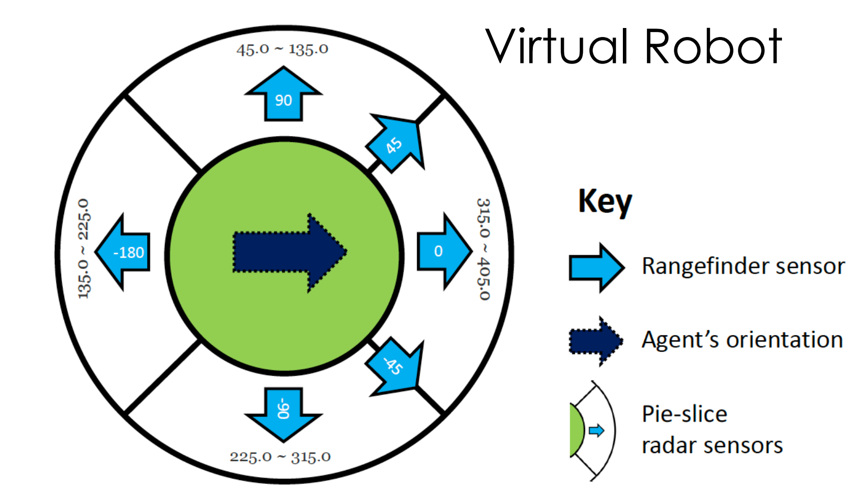
NOVELTY SEARCH

(Lehman Stanley, ECJ 2011)

- Many different genomes produce the same behaviour
- Genetic diversity is not sufficient in Neuroevolution
- Idea: directly reward a diversity of behaviours

MAZE NAVIGATION

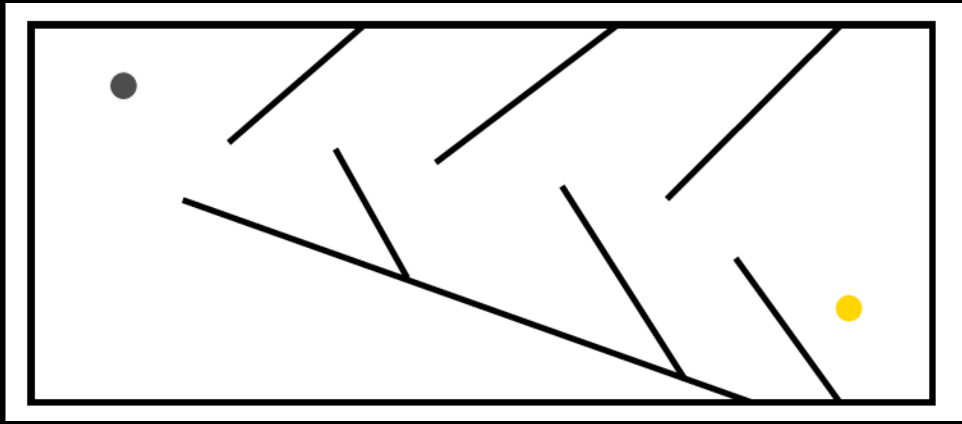




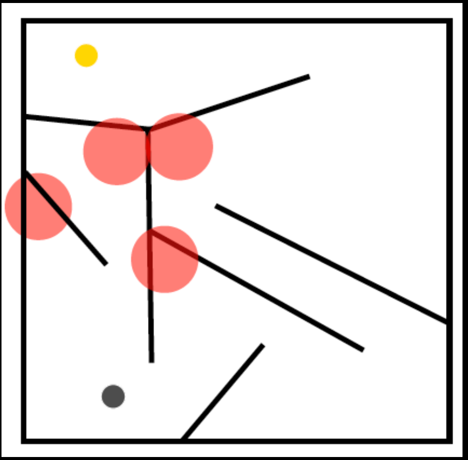
Fitness-based search

Novelty search

Crossover vs. no Crossover



Medium Maze



Goal

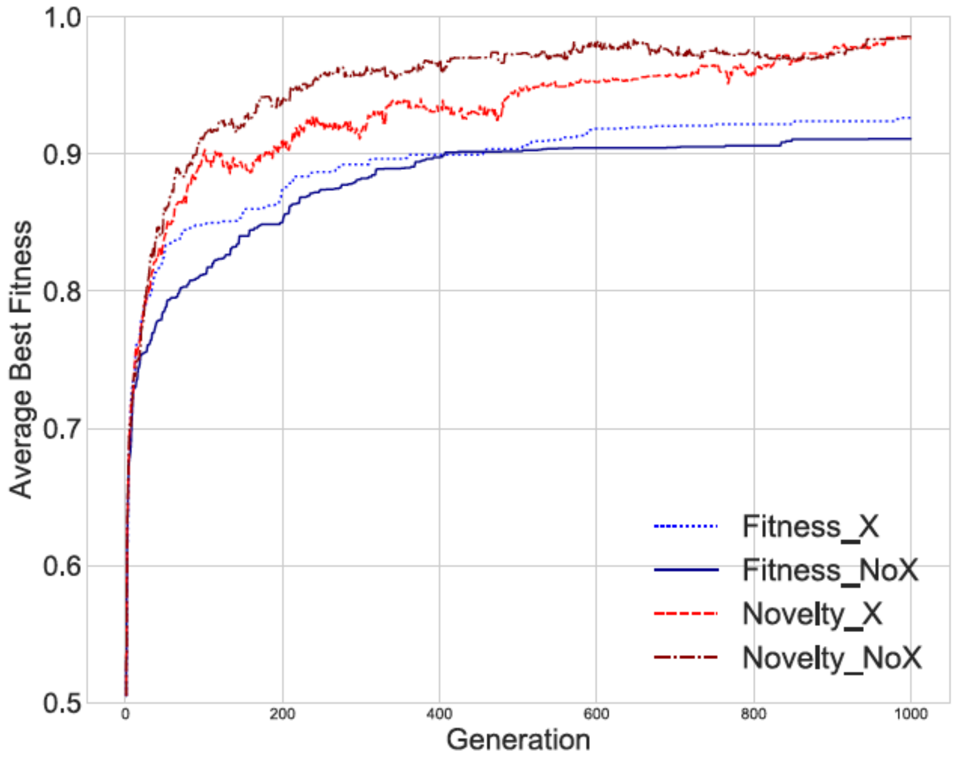
Start

Deceptive traps

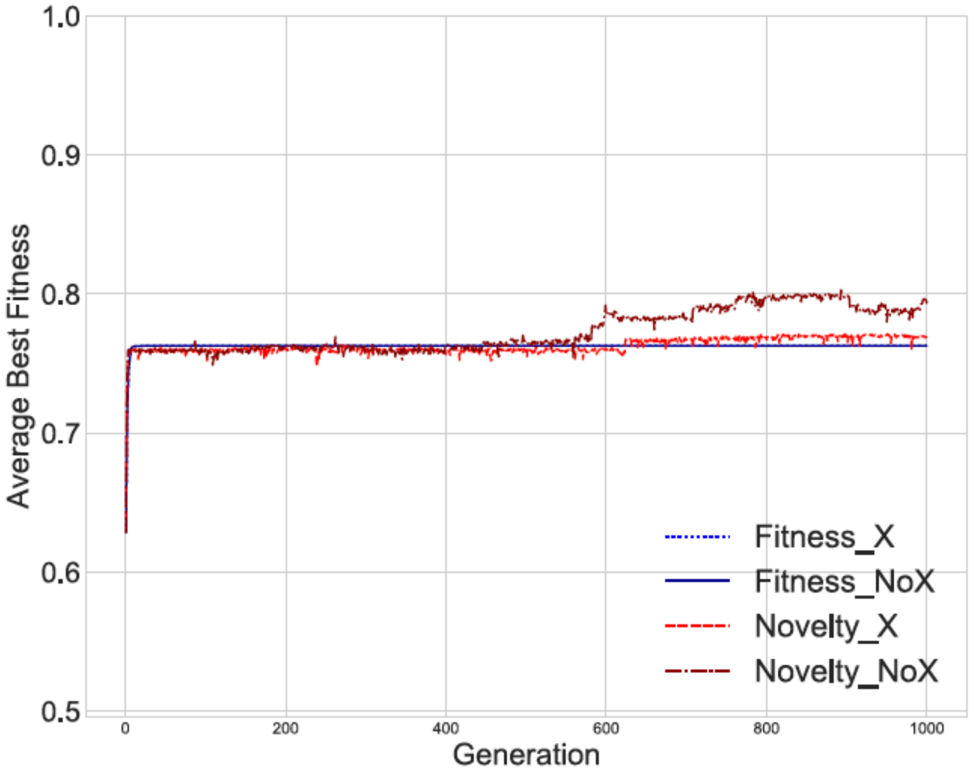
Hard Maze
Deceptive problem

NEAT PERFORMANCE CURVES

Average best fitness (30 runs) generations for the four NEAT variants



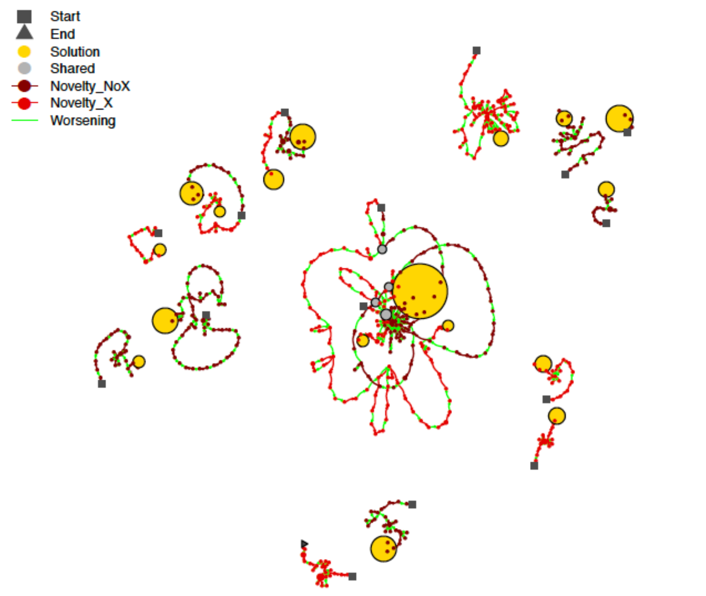
(a) Medium Maze Domain



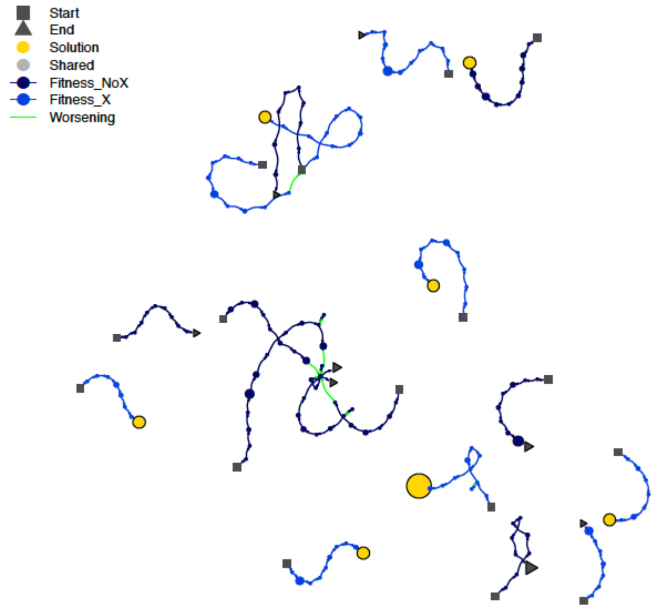
(b) Hard Maze Domain

NEAT STNs

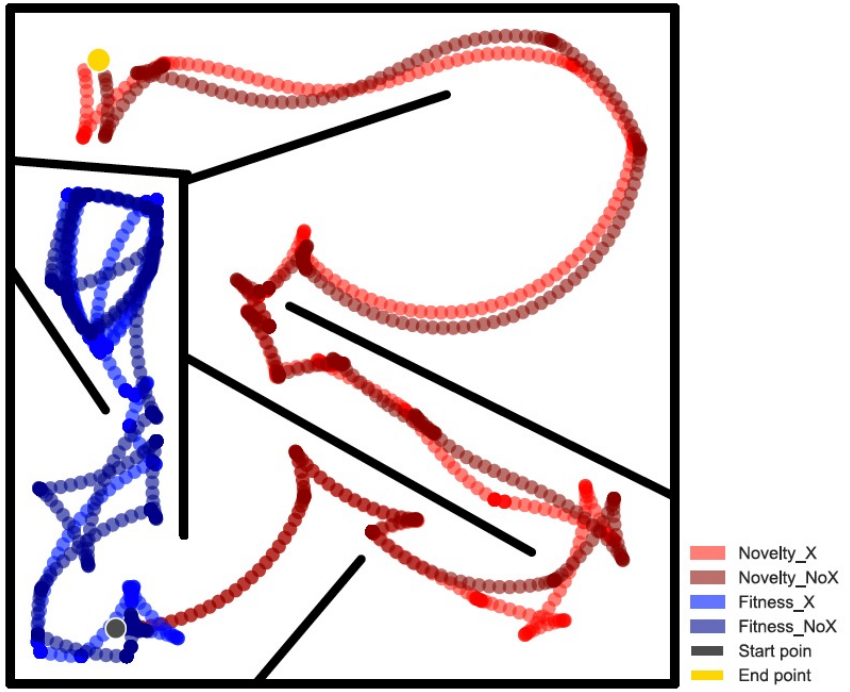
Merged models with and without Xover
- 9 runs for each strategy
- Size: number of compressed nodes

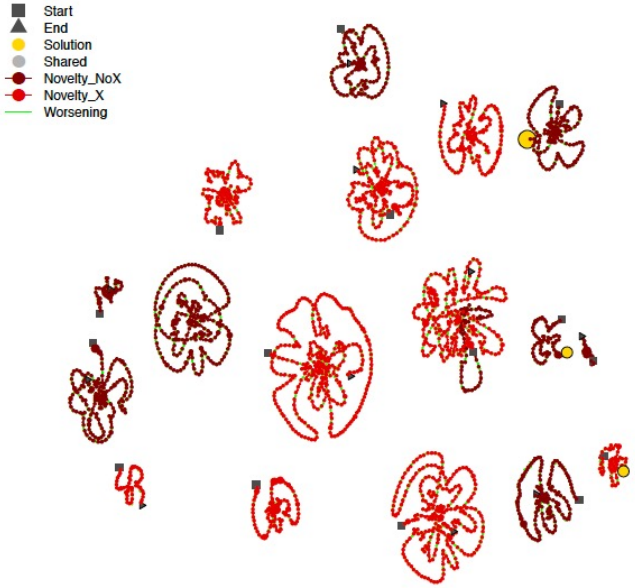


(a) Novelty, medium map

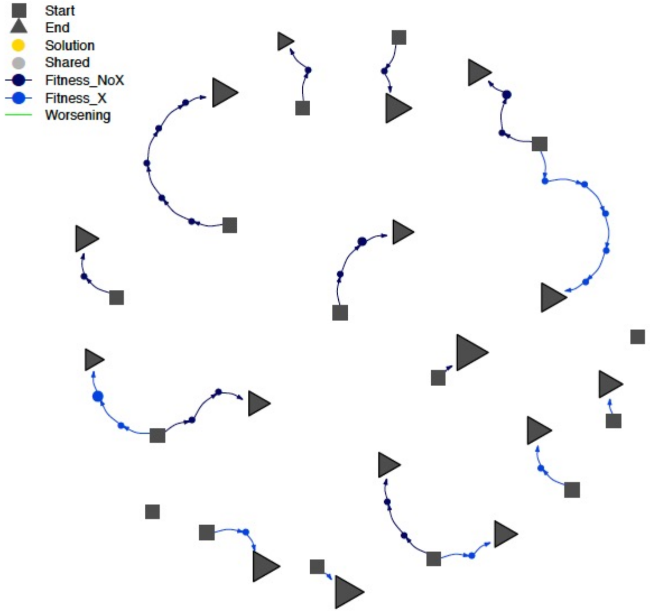


(b) Fitness, medium map



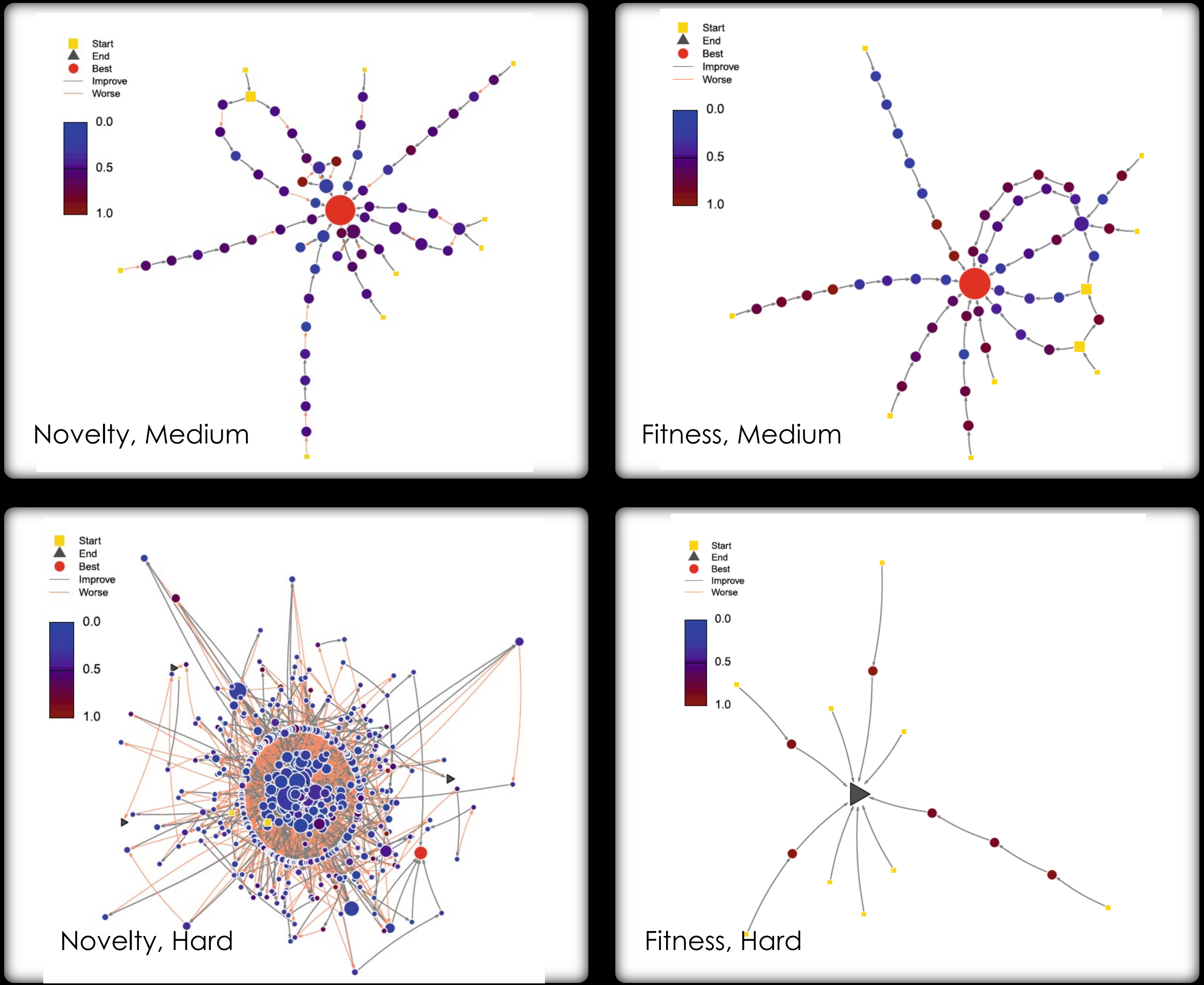


(c) Novelty, hard map



(d) Fitness, hard map

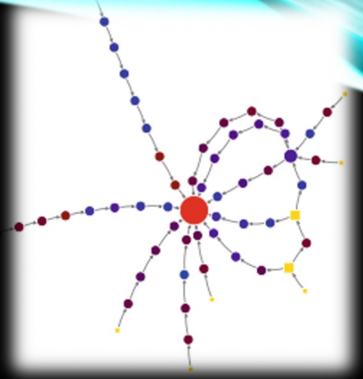
STNs IN BEHAVIOUR SPACE



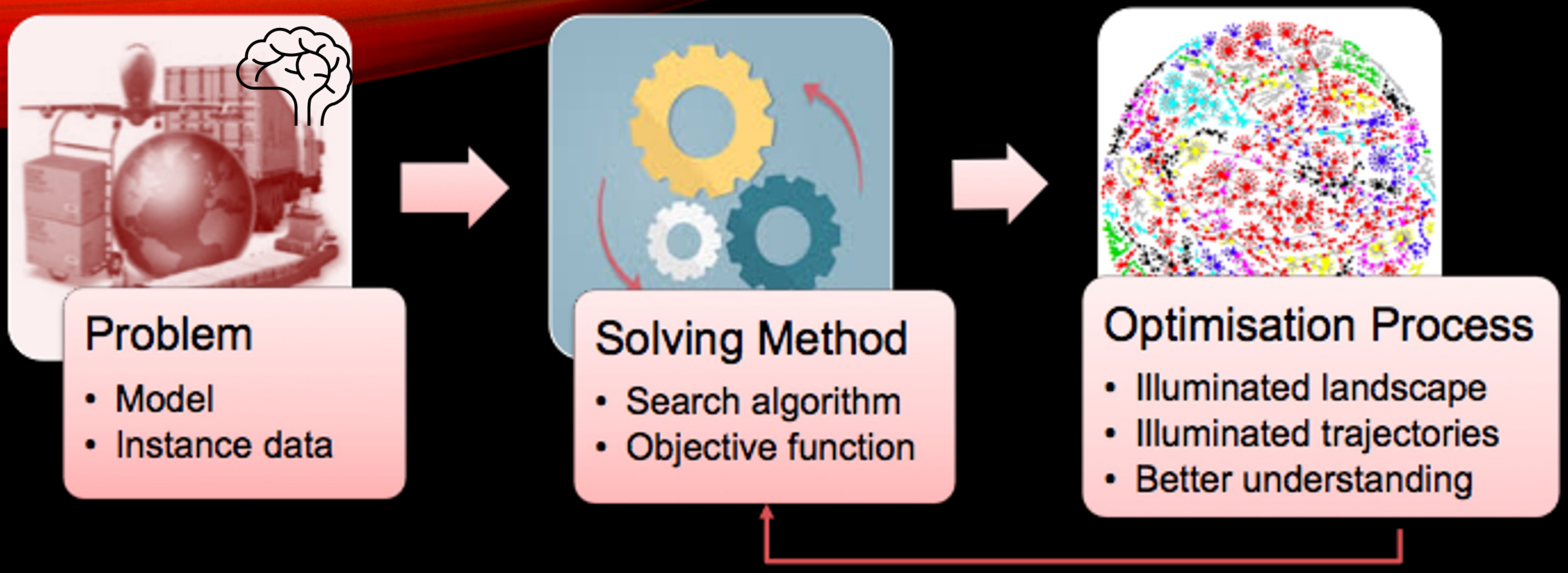
Behaviour in the maze domain is simply the (*x*, *y*) coordinate of the agent at the end of simulation

Sarti, S., Adair, J., Ochoa, G. (2022) Neuroevolution Trajectory Networks of the Behaviour Space. *Applications of Evolutionary Computation*. EvoApps 2022.

INSIGHTS FROM STNs IN NEAT



- So far applied only to early versions of NEAT with direct encoding
- Genotype STNs do not scale well to complex problems
- Phenotype/Behaviour space STNs can be useful
- Confirms that Novelty search is useful in deceptive and neutral domains
- Crossover does not seem to be useful!



Conclusion

An accessible and visual approach to understanding problems & methods

LON Resources

- lonmaps.com
- https://github.com/gabro8a/LONs-Numerical

STN Resources

- https://github.com/gabro8a/STNs
- Web Application http://45.32.184.82

Future work

Using knowledge to select/configure algorithms

Build up software tools!

REFERENCES

LONs for NAS

- N M Rodrigues, K M Malan, G Ochoa, L Vanneschi, S Silva (2022) Fitness landscape analysis of convolutional neural network architectures for image classification, *Information Sciences*, Vol. 609, pp. 711-726.
- *G Ochoa*, N Veerapen (2022). Neural Architecture Search: A Visual Analysis. *Parallel Problem Solving from Nature* – PPSN XVII. PPSN 2022. Lecture Notes in Computer Science, vol 13398. Springer, Cham

STNs for NEAT

- S Sarti, JAdair, G Ochoa (2022) Recombination and Novelty in Neuroevolution: A Visual Analysis. *SN COMPUT. SCI.* 3, 185
- Sarti, S., Adair, J., Ochoa, G. (2022) Neuroevolution Trajectory Networks of the Behaviour Space. *Applications of Evolutionary Computation*. EvoApps 2022. Lecture Notes in Computer Science, vol 13224. Springer.
- S Sarti, G Ochoa (2021) A NEAT Visualisation of Neuroevolution Trajectories. *Applications of Evolutionary Computation* - EvoApps 2021. Lecture Notes in Computer Science, vol 12694. Springer, Cham.