NEUROEVOLUTION LANDSCAPES AND TRAJECTORIES

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

Gabriela Ochoa
MOTIVATION

Visualisation & Analysis Tools

Increased Understanding (Explainability)

Informed Algorithm Selection & Configuration
OUTLINE

G. Theory ~1700, ~2000
Complex Networks

Local Optima Networks (LONs)
2008, NAS 2022

Search Trajectory Networks (STNs)
2020, NE 2021, 2022
WHAT IS A (COMPLEX) NETWORK

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

Graph
Mathematical Object

Network
Data Driven Instantiation

Protein Homology
32,727 Proteins
1.2 M Edges

https://mathworld.wolfram.com/WeaklyRegularGraph.htm
http://www.visualcomplexity.com/
NETWORK VISUALISATION

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

Petersen graph

Node-edge diagram graph layouts

https://kateto.net/network-visualization
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)
Many Algorithms

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

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

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

Biology, Computational Chemistry
Evolutionary Computation & Metaheuristics
Machine Learning

Multimodality
Ruggedness
Deceptiveness
Neutrality

2-d slice, (9-d) landscape in Evol. Robotics (Barnet, 2002)
WHAT IS A FUNNEL?

Protein Folding

(Kerschke et al., GECCO 2015)

T. Splettstoesser
www.scistyle.com

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)


Characterising Funnels with LONs

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

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

Chained Lin-Kernighan
(Martin, Otto, Felten, 1992)

- Form of Iterated Local Search
- Diversification & Intensification

Algorithm 1 ILS sampling

| L ← {}; E ← {} (initialise LON data) |
| repeat |
| s₀ ← RandomInitialSolution |
| s* ← HillClimber(s₀) |
| L ← L ∪ {s*} (record LON data) |
| repeat |
| s' ← RandomMutation(s*) |
| s*' ← HillClimber(s') |
| if f(s'*') ≤ f(s*) then |
| L ← L ∪ {s*''} (record LON data) |
| E ← E ∪ {(s*, s'*')} (record LON data) |
| s* ← s*' |
| end if |
| until ILS termination condition is met |
| until sampling termination condition is met |

Nodes: LK local optima
Edges: double-bridge escapes
A Chained-LK run on instance lin318
318 points, drilling application
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
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 followed 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.
ALTERNATIVE NAS SEARCH SPACES

Chain Structure

Multi-branch Structure

Two cells: Normal & Reduction

Cell (Block) Search Space
EXAMPLE OF SEARCH PROGRESS

**CELL BASED TABULAR BENCHMARK**

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

- **Image**
- **Conv**
- **Cell** $\times$ N
- **Residual block (stride = 2)**
- **Cell** $\times$ N
- **Residual block (stride = 2)**
- **Cell** $\times$ N
- **Global avg. pool**

### Genotype

- **DAG**

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

---

**Cell (DAG)**

Discrete optimisation problem

$$\max_{a \in A} f(a)$$

$\tilde{f}(a)$ validation accuracy

---

IMAGE CLASSIFICATION DATASETS

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

- airplane
- automobile
- bird
- cat
- deer
- dog
- frog
- horse
- ship
- truck

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

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

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

Algorithms with no Local Search

Generalise LONs

Population-based Algorithms
**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 $STN = (N, E)$. 

---

33
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, $D = 10$

**PSO** – Particle Swarm Optimisation

**DE** – Differential Evolution

<table>
<thead>
<tr>
<th></th>
<th>PSO</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>93</td>
<td>154</td>
</tr>
<tr>
<td>Path length</td>
<td>9.2</td>
<td>15.1</td>
</tr>
<tr>
<td>Success</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Unimodal function
Michalewicz, D = 5

Multimodal, large plateaus close to the global optimum

**PSO** – Particle Swarm Optimisation

**DE** – Differential Evolution

<table>
<thead>
<tr>
<th></th>
<th>PSO</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>76</td>
<td>61</td>
</tr>
<tr>
<td>Path length</td>
<td>7.5</td>
<td>7.0</td>
</tr>
<tr>
<td>Success</td>
<td>40%</td>
<td>80%</td>
</tr>
</tbody>
</table>
NEUROEVOLUTION

NEAT – Neuroevolution of Augmented Topologies

Evolve weights, fixed topology

Evolve topology, learn weights with gradient-based methods

Evolve both weights and topology

Evolve components, hyperparameters, learning rules

Genome (Genotype)

<table>
<thead>
<tr>
<th>Node</th>
<th>Node 1</th>
<th>Node 2</th>
<th>Node 3</th>
<th>Node 4</th>
<th>Node 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor</td>
<td>Sensor</td>
<td>Sensor</td>
<td>Output</td>
<td>Hidden</td>
<td></td>
</tr>
<tr>
<td>In 1</td>
<td>Out 4</td>
<td>In 2</td>
<td>Out 4</td>
<td>In 3</td>
<td>In 2</td>
</tr>
<tr>
<td>Weight 0.7</td>
<td>Weight 0.5</td>
<td>Weight 0.5</td>
<td>Weight 0.2</td>
<td>Weight 0.4</td>
<td>Weight 0.6</td>
</tr>
<tr>
<td>Enabled</td>
<td>Enabled</td>
<td>Disabled</td>
<td>Enabled</td>
<td>Enabled</td>
<td>Enabled</td>
</tr>
<tr>
<td>In 1</td>
<td>In 2</td>
<td>In 3</td>
<td>In 4</td>
<td>In 5</td>
<td>In 1</td>
</tr>
<tr>
<td>In 1</td>
<td>Out 5</td>
<td>Out 4</td>
<td>Out 5</td>
<td>Out 4</td>
<td>Out 5</td>
</tr>
<tr>
<td>Weight 0.6</td>
<td>Weight 0.4</td>
<td>Weight 0.6</td>
<td>Enabled</td>
<td>Enabled</td>
<td>Enabled</td>
</tr>
<tr>
<td>Enabled</td>
<td>Enabled</td>
<td>Enabled</td>
<td>In 6</td>
<td>In 11</td>
<td>In 4</td>
</tr>
</tbody>
</table>

Network (Phenotype)

(Stanley & Miikkulainen, 2002)
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)

Many different genomes produce the same behaviour

Genetic diversity is not sufficient in Neuroevolution

Idea: directly reward a diversity of behaviours
MAZE NAVIGATION

Virtual Robot

Key
- Rangefinder sensor
- Agent’s orientation
- Pie-slice radar sensors

Medium Maze

Fitness-based search

Novelty search

Crossover vs. no Crossover

Hard Maze
Deceptive problem

Goal
Start
Deceptive traps
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
STNs IN BEHAVIOUR SPACE

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

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


STNs for NEAT

