Artificial Neural Networks
NeuroEvolution = ANN + EA

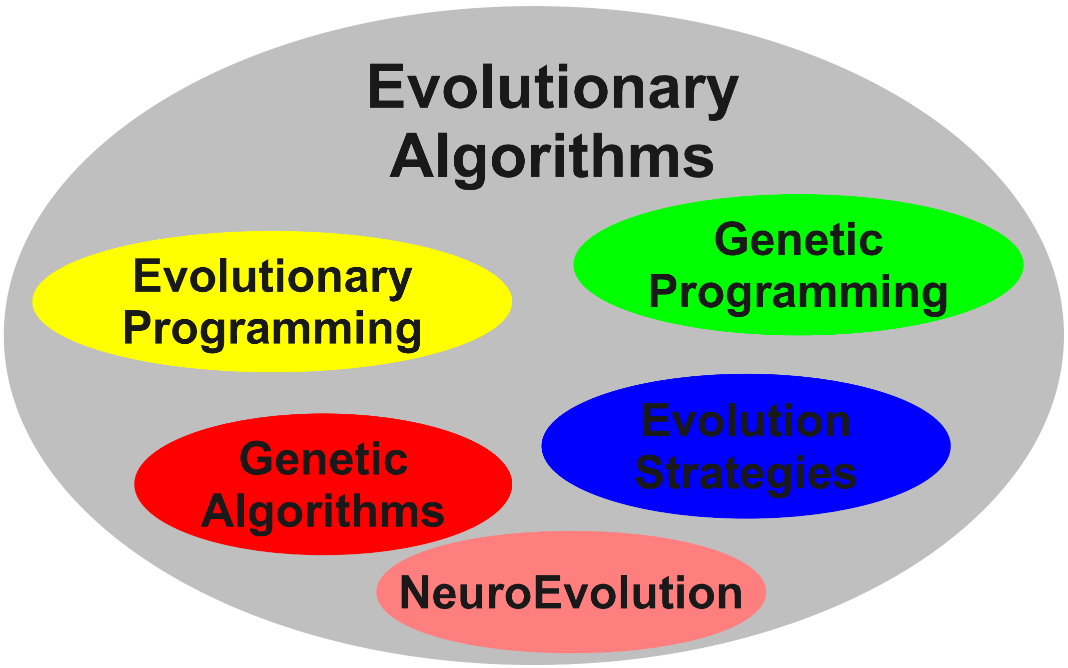
Jan Drchal
`drchajan@fel.cvut.cz`

Computational Intelligence Group
Department of Computer Science and Engineering
Faculty of Electrical Engineering
Czech Technical University in Prague

Motivation

- Learning ANNs = optimization of weights (or potentially structure).
- Problem of local extremes → unable to learn hard task/large networks.
- Use of **Evolutionary Algorithms** → slower, but more robust than classic gradient methods like Back-Propagation.

Evolutionary Algorithms (EAs)



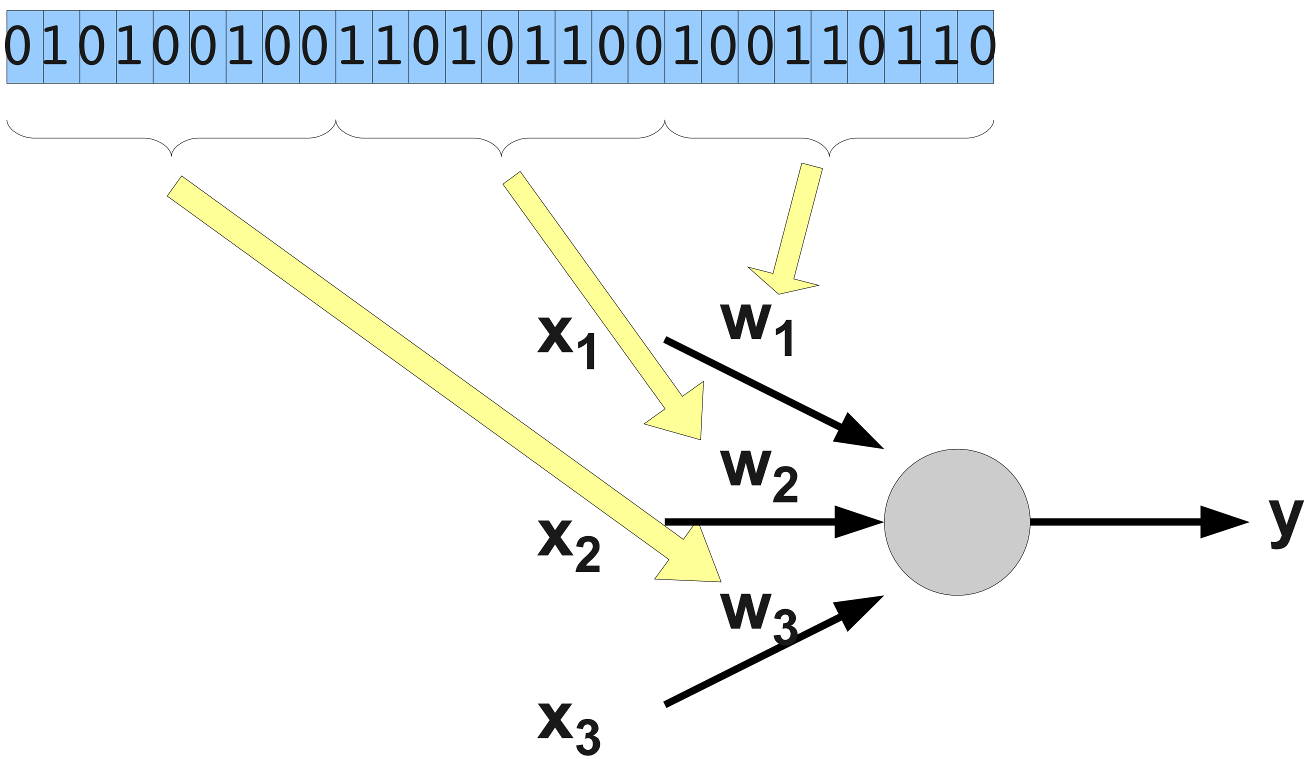
- **Genetic Algorithms:** binary strings
- **Evolutionary Strategies:** real vectors, only mutation.
- **Genetic Programming:** evolution of program trees.
- **Evolutionary Programming:** evolving FSMs.
- **NeuroEvolution**

What is Neuro Evolution?

- **Neuro-evolutionary algorithm is just another special kind of EA** → the task is to evolve (learned) neural networks.
- Both parameters (weights) and topology can be optimized by evolution.
- **But how to encode a network into a genome?** → A network with fixed topology is described by a vector of all its weights (real numbers)...

Direct Encoding of Neural Network

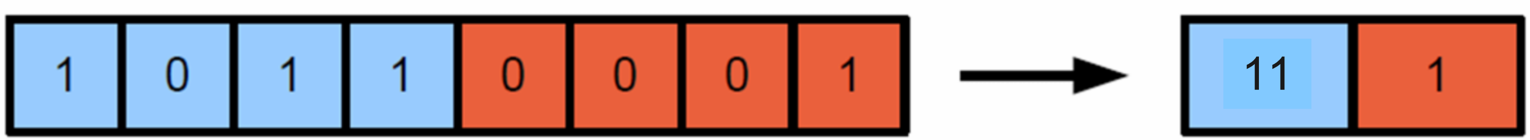
- Directly encode the weights as a bit string:



- Can we do it better? Yes.

Floating-Point Encoding

- Motivation: simplicity, precision

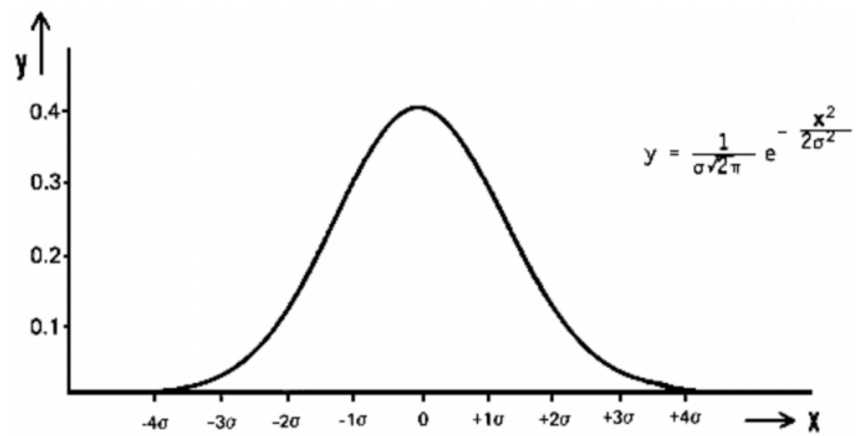


Binary string encodes vector of 2 numbers .

The same encoded using floating-point encoding.

What about mutation? -> Gaussian noise.

Idea:
small changes with *higher* probability,
large changes with *lower*.

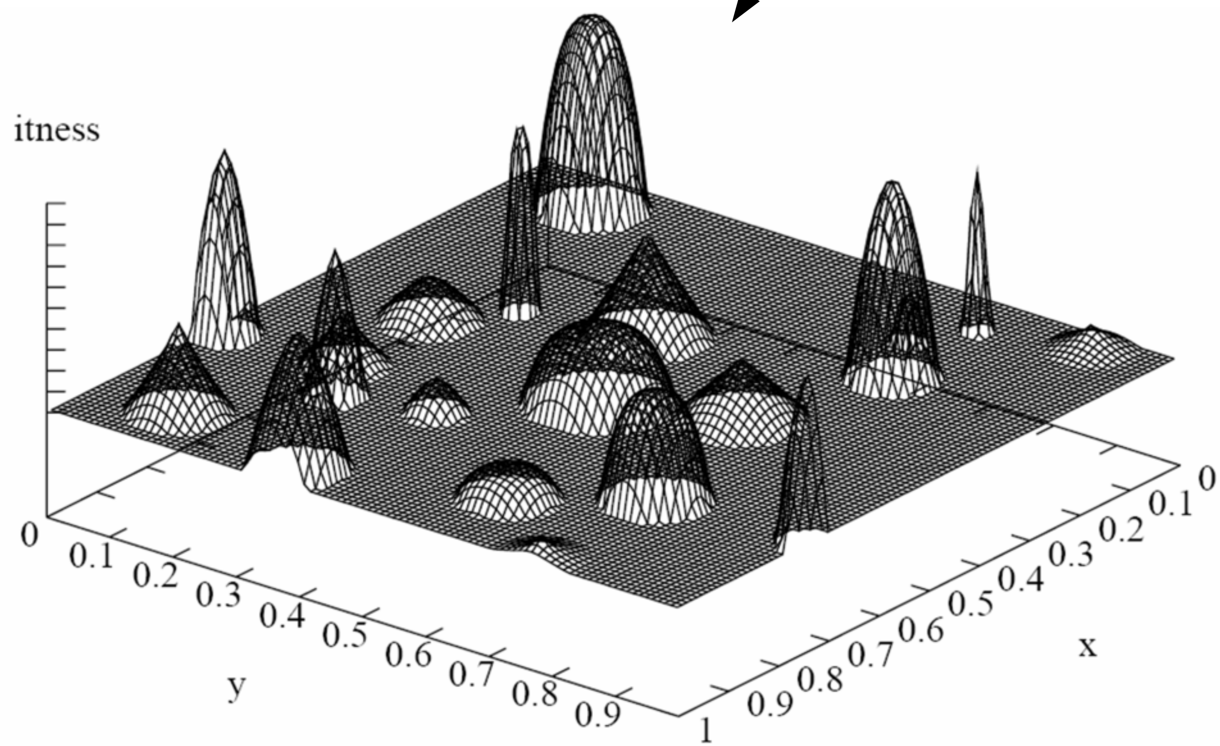


- Useful for integers and floats ...

Multimodal Domains

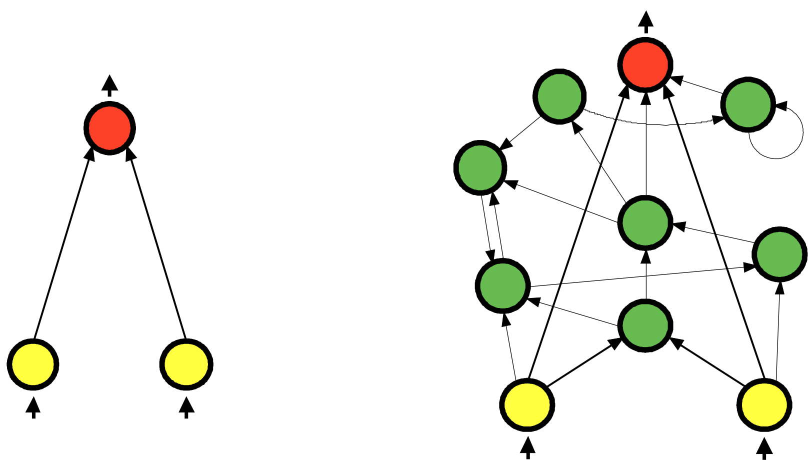
- Multimodal functions:
  - multiple optima,
  - many local.
- Too many attractors → hard optimization :(
- ANN fitness/error landscapes look like this.

This is our fitness landscape



NEAT

- **N**euro**E**volution of **A**ugmenting **T**opologies: Kenneth O. Stanley, 2001, The University Of Texas at Austin
- Complexification – start from small topologies: evolution add neurons/links as needed by task.



Kenneth O. Stanley and Risto Miikkulainen: **Evolving Neural Networks Through Augmenting Topologies**

NEAT 2

- Topology is augmented by adding neurons and links between.

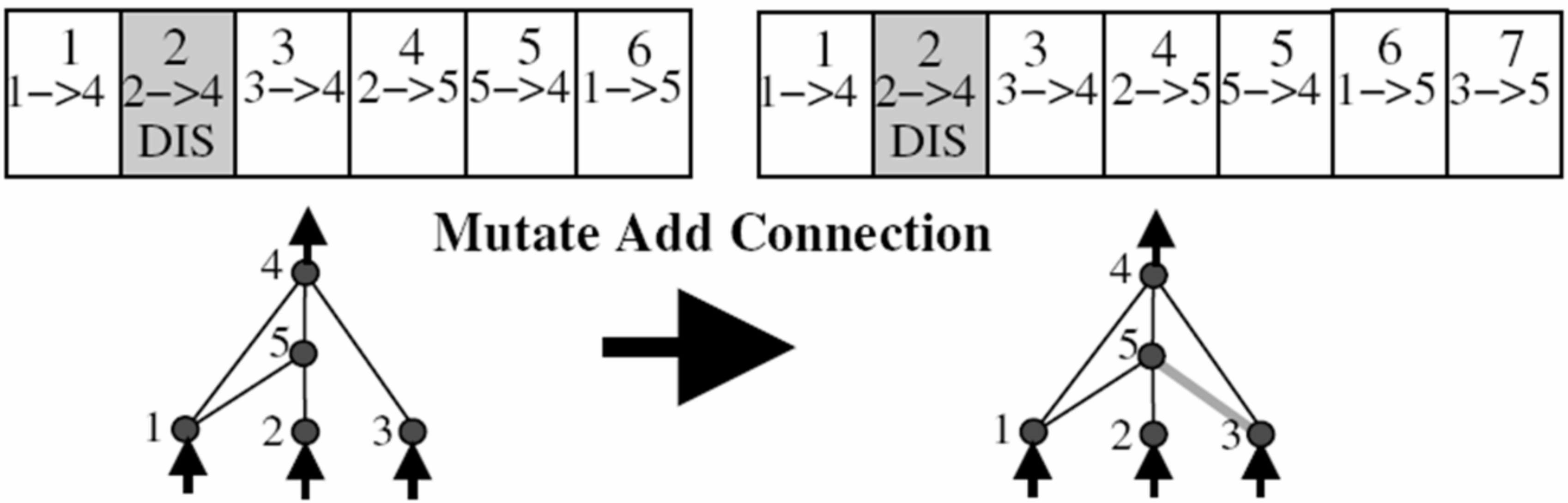
  → Variable genome length.

- Mutations:
  - parametric – Gaussian noise,
  - structural – adding neurons & links (no pruning), switch on/off links.

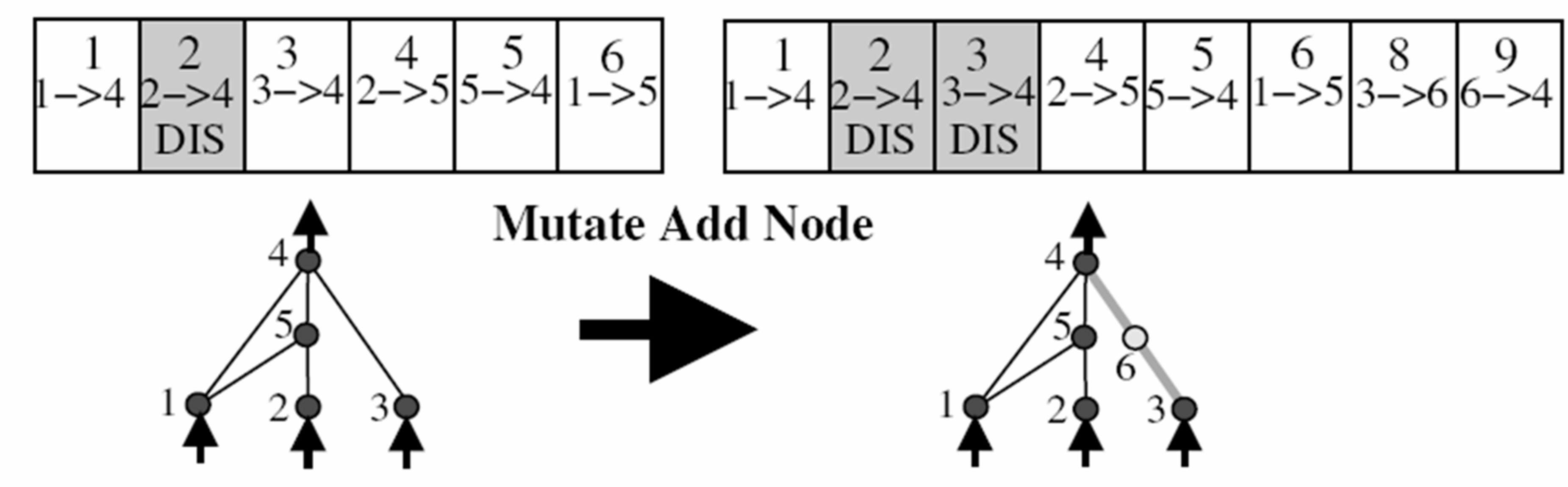
- Mating – special crossover two parents → single child.

Note, some newer implementations use pruning However, it is not essential.

Add Link Mutation



Add Neuron Mutation



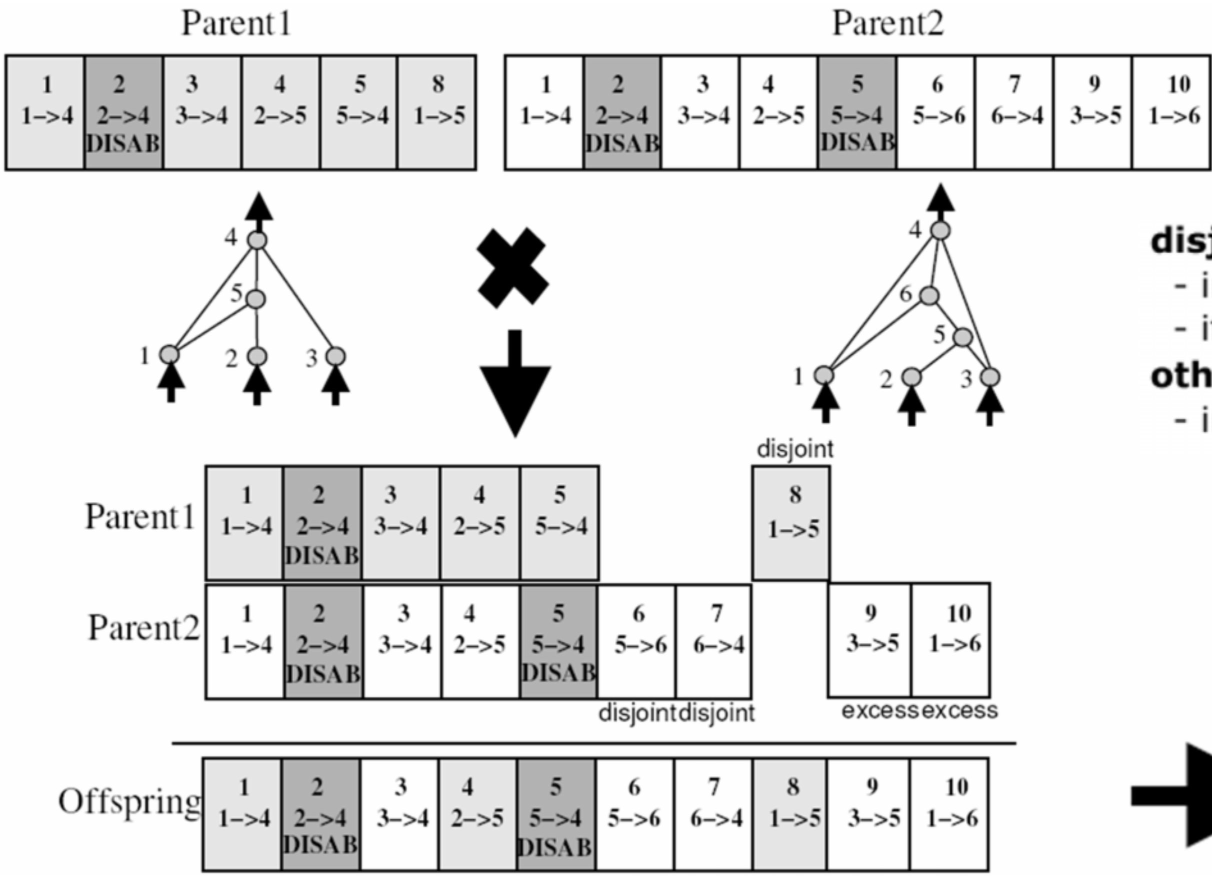
The weights of new neuron's incoming/outgoing links are set in a way which minimizes the difference between original and mutated networks.

Historical Markings

- **Q**: How to *align* two genomes of different size representing two different networks?
- **A**: let's use “the creation date” of a particular gene (caused by a structural mutation) – **historical marking (innovation number)**.
- Aligning two genomes:
  - when two genes with matching HMs are found, it is likely they have similar function in the network.
- HM is a counter, the same value is assigned for the same innovation within a single generation (i.e. adding a link between neurons #3 and #4).

Mating

Let's use historical marking.

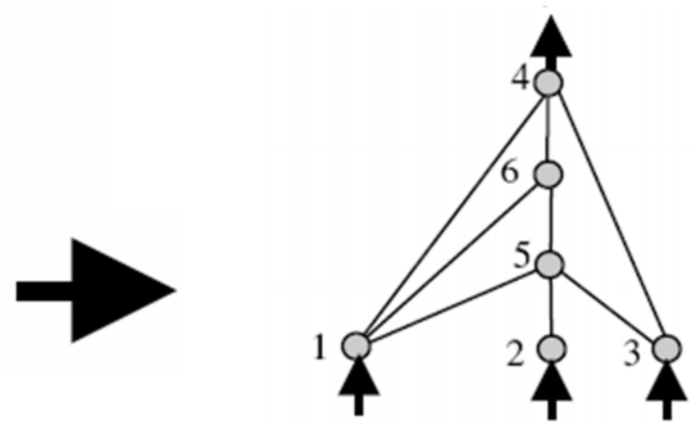


disjoint or **excess**

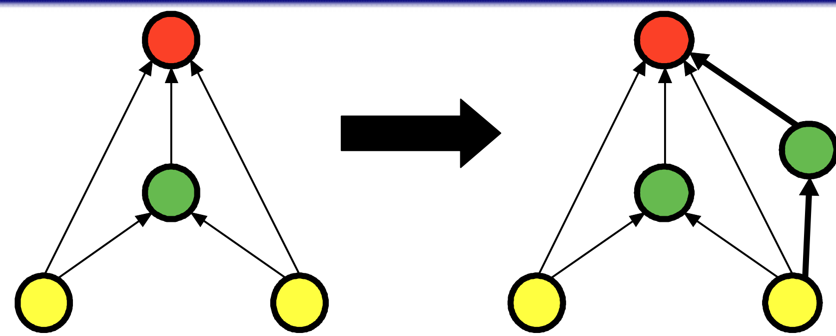
- inherit more fit
- if equal fitness inherit randomly

other

- inherit randomly



Niching



- There are networks of different sizes in the population.
- Adding a new structure:
  - likely lowers the fitness,
  - larger networks → longer genome → more time needed to optimize parameters.
- **New topologies must be protected** → niching.
- Here we use Explicit Fitness Sharing:

  Separate the population into species → selection and reproduction only among similar individuals → HMs again used to compute similarity of two genomes.

The Three Most Important Ideas Behind NEAT

- **Complexification** – start with small networks, gradually add neurons/links (reminds GMDH or GAME approaches).
- Concept of **historical markings** - cross/match only corresponding genes → **deals with competing conventions.**
- Use of **niching** - allows the survival of larger, recently structurally innovated networks → gives them time to optimize their weights and “show” that the structural innovation was beneficial.