Lecture 1: Introduction to Evolutionary Computation

CSE5012: Evolutionary Computation and Its Applications

Xin Yao

CSE, SUSTech

18 Feb 2022

Lecture 1: Introduction to Evolutionary Computation

Outline of This Lecture

Current Computing Systems

- **Rigid (死板的) and inflexible (不灵活)**
- **Brittle (脆弱的)**
- **Non-adaptive (自适应性差)**
- **Poor at coping with noise (抗噪声能力低)**
- **Reliant on human intervention (太依赖于人)**
- **Boring**
- **Doesn't learn and generalise**
- **Not autonomous**
- **Slow**
- **...**

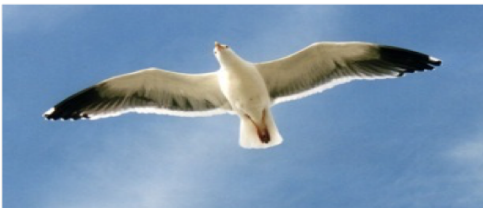
Brittle
(脆弱性)

Figure 1: Image source:
http://icdn5.digitaltrends.com/image/spilled-liquid-on-laptop-625x400.jpg

Natural Systems
(自然系统)

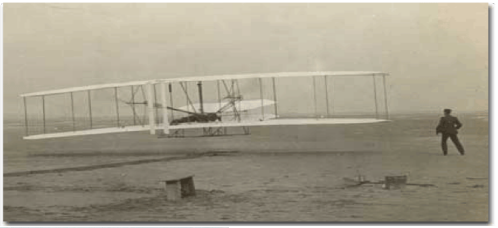
In contrast, natural systems are often (对比而言，自然系统通常是):

- **Very resilient (恢复力强)**
- **Exceptionally well adapted (自适应性极强)**
- **Always learning (总在学习)**
- **Unsupervised (无需监督的)**
- **Creative & Ingenious (充满创造性&奇妙性)**



Inspiration, Not Copying

Nature **Inspired** Computation



Why bother?

(为何要向大自然借鉴?)

- **Nature, through evolution, has solved some extraordinarily complex problems (自然善于解决复杂问题)**
- **It is a highly efficient system (高效率)**
- **By necessity (必需的)!**
- **They are (mostly) intuitively simple (通常很简单)**
- **They give (mostly) comprehensible answers (通常给出可解释的答案)**

Nature Inspired Computing Techniques
(受自然启发的计算技术)

Technique (计算技术)	Inspiration (自然界灵感)
Evolutionary algorithms (演化算法)	**Biological process of evolution** (生物进化过程)
Artificial neural networks (人工神经网络)	**Neurons in the brain** (大脑神经元)
Agent-based techniques (多智能体系统)	**Human social interaction** (人类社交活动)
Ant / Swarm techniques (蚁群技术)	**Social insects** (群居昆虫)
Quantum computing (量子计算)	**Quantum physics** (量子物理)

Outline of This Lecture

Evolutionary Computation
(演化计算)

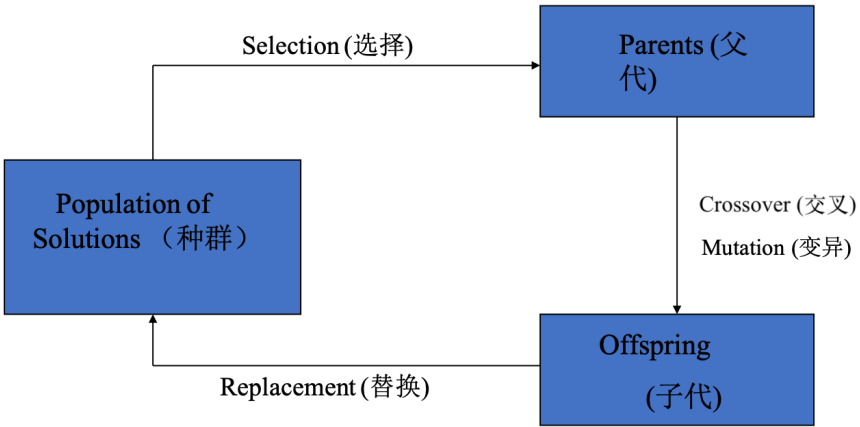


Figure 2: Survival of the fittest (适者生存).

What Is Evolutionary Computation

1. **It is the study of computational systems which use ideas and get inspirations from natural evolution.**

2. **One of the principles borrowed is survival of the fittest.**

3. **EC techniques can be used in optimisation, learning and design.**

4. **EC techniques do not require rich domain knowledge to use. However, domain knowledge can be incorporated into EC techniques.**

EA as Population-Based Generate-and-Test

Generate **Mutate and/or recombine individuals in a population.**

Test **Select the next generation from the parents and offspring.**

While drawing analogy to biology may be sexy, it is important to understand the underlying links between "new" AI and "old" AI (GOFAI).

A Simple Evolutionary Algorithm (EA)

1. **Generate the initial population $P(0)$ at random;**
2. **Set $i \leftarrow 0$;** ***// Generation counter***
3. **REPEAT**
   - 3.1 **Evaluate the fitness of each individual in $P(i)$;**
   - 3.2 **Select parents from $P(i)$ based on their fitness in $P(i)$;**
   - 3.3 **Generate offspring from the parents using crossover and mutation to form $P(i+1)$;**
   - 3.4 **$i \leftarrow i+1$;**
4. **UNTIL halting criteria are satisfied**

Evolutionary Computation - Conceptually An Important Means to AI

Turing's conceive: realising intelligence with evolution

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child brain is something like a notebook as one buys it from the stationer's. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed. The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.

[1] A. M. Turing, "Computing machinery and intelligence." Mind, 49:433-460, 1950.

Evolutionary Computation - Conceptually An Important Means to AI

Turing's conceive: realising intelligence with evolution

We have thus divided our problem into two parts. The child programme and the education process. These two remain very closely connected. We cannot expect to find a good child machine at the first attempt. One must experiment with teaching one such machine and see how well it learns. One can then try another and see if it is better or worse. There is an obvious connection between this process and evolution, by the identifications

Structure of the child machine = hereditary material

Changes of the child machine = mutation,

Natural selection = judgment of the experimenter

One may hope, however, that this process will be more expeditious than evolution. The survival of the fittest is a slow method for measuring advantages. The experimenter, by the exercise of intelligence, should he able to speed it up. Equally important is the fact that he is not restricted to random mutations. If he can trace a cause for some weakness he can probably think of the kind of mutation which will improve it.

[1] A. M. Turing, "Computing machinery and intelligence." Mind, 49:433-460, 1950.

How Does a Simple EA Work

Example

Let's use the simple EA to maximise the function

$$f(\mathbf{x}) = \mathbf{x}^2$$

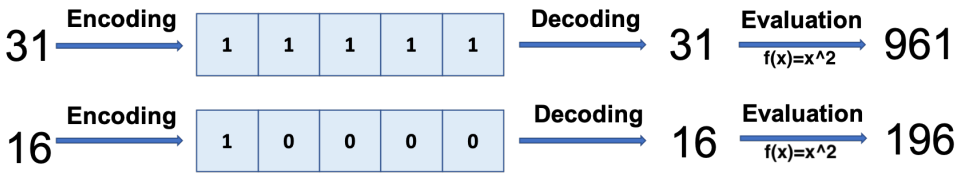
with $\mathbf{x}$ in the *integer* interval $[0, 31]$, i.e., $\mathbf{x} = 0, 1, \cdots, 30, 31$.

How Does a Simple EA Work

Representation

The first step of EA applications is encoding, i.e., the representation of chromosomes:

- **We adopt binary representation for integers.**
- **Five bits are used to represent integers up to 31.**



How Does a Simple EA Work I

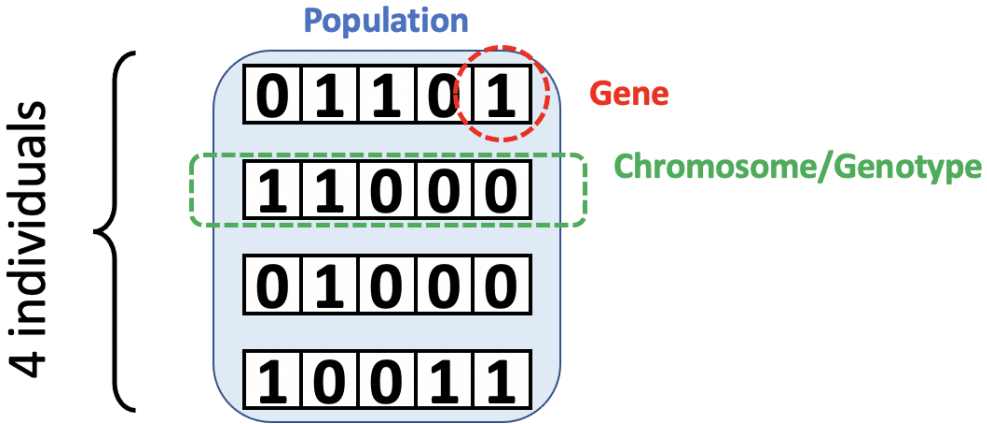
Assume that the population size is 4.

1. **Generate initial population at random, e.g.,**

$$01101, 11000, 01000, 10011.$$

These are *chromosomes* (染色体) or *genotypes* (基因型).

How Does a Simple EA Work II



How Does a Simple EA Work III

2. **Calculate fitness value for each individual.**

   2.1 **Decode the individual into an integer (called *phenotypes* 表现型),**

$$01101 \rightarrow 13$$

$$11000 \rightarrow 24$$

$$01000 \rightarrow 8$$

$$10011 \rightarrow 19$$

   2.2 **Evaluate the fitness according to $f(\mathbf{x}) = \mathbf{x}^2$,**

$$01101 \rightarrow 13 \rightarrow 169$$

$$11000 \rightarrow 24 \rightarrow 576$$

$$01000 \rightarrow 8 \rightarrow 64$$

$$10011 \rightarrow 19 \rightarrow 361$$

How Does a Simple EA Work IV

3. **Select two individuals for crossover based on their fitness.**

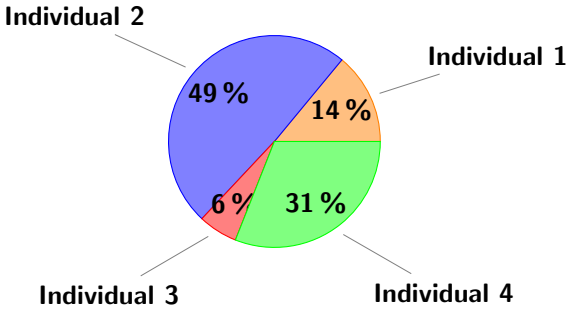
   3.1 **If roulette-wheel selection is used, then**

$$p_i = \frac{f_i}{\sum_j f_j}.$$

Two offspring are often produced and added to an intermediate population. Repeat this step until the intermediate population is filled. In our example,

$$p_1(13) = 169/1170 = 0.14 \quad p_2(24) = 576/1170 = 0.49$$
$$p_3(8) = 64/1170 = 0.06 \quad p_4(19) = 361/1170 = 0.31$$

How Does a Simple EA Work V



3.2 **Assume we have** $crossover(01101, 11000)$ **and** $crossover(10011, 11000)$**. We may obtain offspring** 01100 **and** 11001 **from** $crossover(01101, 11000)$ **by choosing a random crossover point at** 4**, and obtain** 10000 **and** 11011 **from** $crossover(10011, 11000)$ **by choosing a random crossover point at** 2**. Now the intermediate population is 01100, 11001, 10000, 11011.**

How Does a Simple EA Work VI

4. **Apply mutation to individuals in the intermediate population with a *small* probability. A simple mutation is bit-flipping. For example, we may have the following new population $P(1)$ after random mutation:**

$$0110\mathbf{1}, 11001, \mathbf{0}0000, 11011$$

The initial population is:

$$01101, 11000, 01000, 10011.$$

5. **Goto Step 2 if not stop.**

Remarks

- **Population-based algorithm.**
- **Stochastic algorithm.**
- **No restriction on the fitness or objective function. It can be nondifferentiable, nonsmooth or even discountinuous.**
- **No need to know the exact form of the objective function. If the objective function is too complex to express explicitly, they can be simulated since EAs only use fitness values.**
  → *Black-box optimisation.*
- **There can be many different evolutionary operators and selection mechanisms.**
  → ***We will learn more in future lectures.***
- **The initial population does not have to be generated at random.**
  → ***We will learn more in future lectures.***
- **The representation of chromosomes does not have to be binary.**
  → ***We will learn more in future lectures.***

Outline of This Lecture

Different Evolutionary Algorithms

There are several well-known EAs with different

- **historical backgrounds,**
- **representations,**
- **variation operators, and**
- **selection schemes.**

In fact, EAs refer to a whole family of algorithms, not a single algorithm.

EA Families

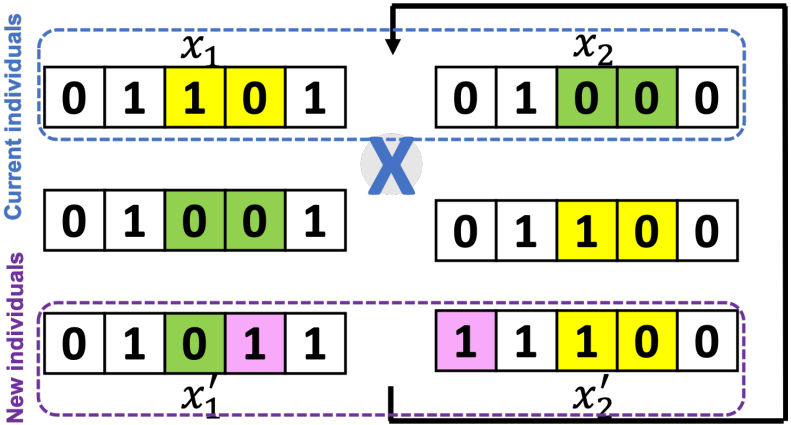
- **Genetic Algorithms (GAs)**
- **Evolutionary Programming (EP)**
- **Evolution Strategies (ES)**
- **Genetic Programming (GP)**
- **Differential Evolution (DE)**
- **Particle Swarm Optimisation (PSO)**
- **...**

Genetic Algorithms (GAs)

First formulated by Holland and by his students for both numerical optimisation and adaptive system design from mid 1960s to mid 1970s [2].

1. **Binary strings have been used extensively as individuals (*chromosomes*).**
2. **Simulate Darwinian evolution.**
3. **Search operators are only applied to the *genotypic* representation (chromosome) of individuals.**
4. **Emphasise the role of recombination (*crossover*). Mutation is only used as a background operator.**
5. **Often use roulette-wheel selection.**

Crossover and mutation



Evolutionary Programming (EP)

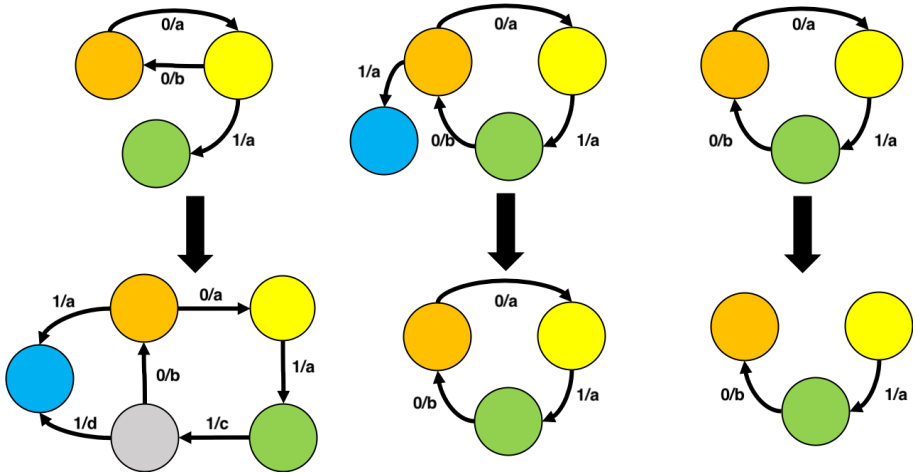
First proposed by Fogel for simulating intelligence in 1960s [3], [4].

1. **Finite-state machines (FSMs) were used to represent individuals, although real-valued vectors have always been used in numerical optimisation.**
2. **It is closer to Lamarckian evolution.**
3. **Search operators (mutations only) are applied to the phenotypic representation of individuals.**
4. **No encoding, not necessarily using binary strings.**
5. **It does not use any recombination.**
6. **Self-adaptive mutation is used.**

Evolutionary Programming (EP)

Illustration

Mutation



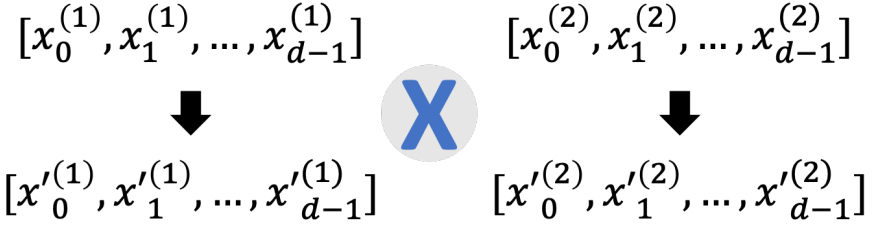
Evolution Strategies (ES)

First proposed by Rechenberg and Schwefel in mid 1960s for numerical optimisation [5], [6].

1. **Real-valued vectors are used to represent individuals.**
2. **They are closer to Larmackian evolution.**
3. **They do have recombination.**
4. **They use self-adaptive mutations.**

Evolution Strategies (ES)

Illustration



Notations:

- **Each vector is a candidate solution.**
- d **denotes the dimension of the solution.**
- $\forall i \in \{0, 1, \ldots, d-1\}$ **and** $\forall j \in \{1, 2\}$**,**
  - $x_i^{(j)}$ **is the coordinate** $(i+1)$ **of the parent** j**,**
  - $x'^{(j)}_i$ **is the coordinate** $(i+1)$ **of the offspring** j**.**

Genetic Programming (GP)

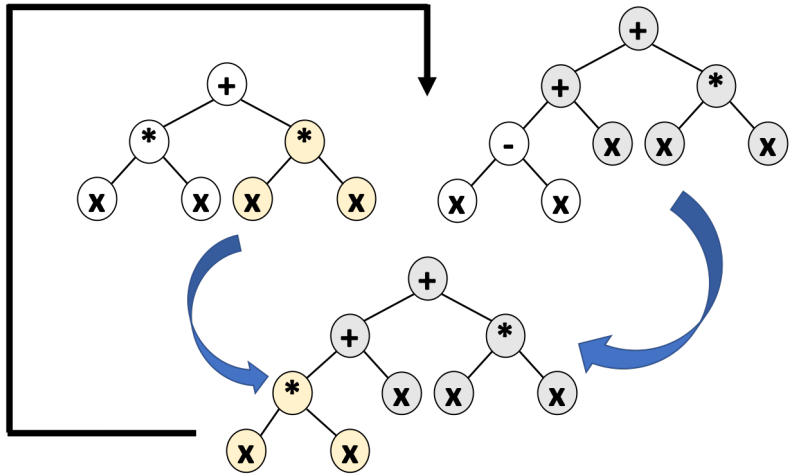
First used by de Garis to indicate the evolution of artificial neural networks [7], but used by Koza to indicate the application of GAs to the evolution of computer programs [8].

1. **Trees (especially Lisp expression trees) are often used to represent individuals.**
2. **Both crossover and mutation are used.**

Genetic Programming (GP)

Illustration

Crossover



Differential Evolution (DE)

- **Continuous search space**
- **Gradient-free**
- **Ill-conditioning**

DE variants [9]:

- **DE/best/1:** $p_i' = p_{best} + F(p_b - p_c)$
- **DE/best/2:** $p_i' = p_{best} + F(p_b - p_c) + F(p_d - p_e)$
- **DE/rand/1:** $p_i' = p_a + F(p_b - p_c)$
- **DE/rand/2:** $p_i' = p_a + F(p_b - p_c) + F(p_d - p_e)$

where p_{best} is the best point in the current population, p_a, p_b, p_c, p_d and p_e are distinct points randomly chosen in the current population.

Differential Evolution (DE)

Generalised Framework

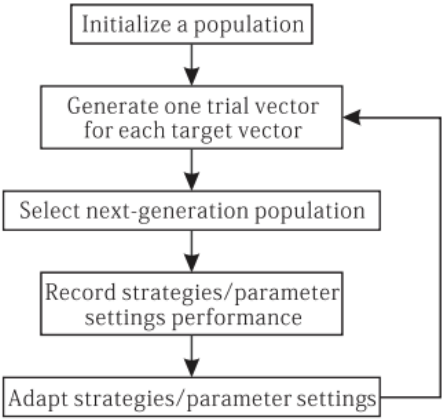


Figure 3: Screenshot of Figure 1.(a) in "X. Lu, K. Tang, B. Sendhoff and X. Yao, *A New Self-adaptation Scheme for Differential Evolution*, Neurocomputing, 146:2-16, 2014".

Particle Swarm Optimisation (PSO)

- **Proposed by Kennedy and Eberhart in 1995 [10]**
- **No crossover & mutation defined through a vector addition**
- **Consider an individual as a point in space with**
  - **a position $\mathbf{x}$ and**
  - **a velocity $\mathbf{v}$: used to determine a new position (and a new velocity)**
- **Differences compared to DE: every candidate solution $\mathbf{x} \in \mathbb{R}^d$ carries its own perturbation vector $\mathbf{v} \in \mathbb{R}^d$**

Preferred Term: Evolutionary Algorithms

- **EAs face the same fundamental issues as those classical AI faces, i.e., representation, and search.**
- **Although GAs, EP, ES, GP, DE, PSO, etc., are different, they are all different variants of population-based generate-and-test algorithms. They share more similarities than differences!**
- **A better and more general term to use is evolutionary algorithms (EAs).**

Variation Operators and Selection Schemes

Crossover/Recombination: k**-point crossover, uniform crossover, intermediate crossover, global discrete crossover, etc.**

Mutation: **bit-flipping, Gaussian mutation, Cauchy mutation, etc.**

Selection: **roulette wheel selection (fitness proportional selection), rank-based selection (linear and nonlinear), tournament selection, elitism, etc.**

Replacement Strategy: **generational, steady-state (continuous), etc.**

Specialised Operators: **multi-parent recombination, inversion, order-based crossover, etc.**

Key points about operators and selections?

Outline of This Lecture

Major Areas in Evolutionary Computation

- **Optimisation**
- **Learning**
- **Design**
- **Theory**

Evolutionary Optimisation

1. **Numerical (global) optimisation.**
2. **Combinatorial optimisation (of NP-hard problems).**
3. **Mixed optimisation.**
4. **Constrained optimisation.**
5. **Multi-objective optimisation.**
6. **Optimisation in a dynamic and/or uncertain environment**

What Can Evolutionary Computation Bring to Us?

Intelligent Optimisation as An Example

- **Optimisation lies in the foundation of AI.**
- **Optimisation is to choose the best solution from a set of candidates as quickly as possible.**

$$\underbrace{x^*}_{\text{The optimal solution}} = \operatorname*{argmin}_{x \in \mathcal{S}} \underbrace{f}_{\text{The problem}}(x)$$

where

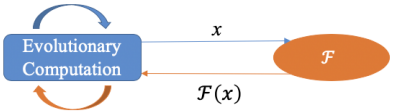
- $\mathcal{S}$ **is solution space and** d **is dimensionality,**
- $f : \mathcal{S} \mapsto \mathbb{R}$ **and** $\mathcal{S} \subseteq \mathcal{X}^d$**.**

- **Evolutionary Computation is featured to realise intelligent optimisation.**

The Features of Evolutionary Computation

- **Grey-box Model. Domain knowledge is not a must, but a boost.**
  - **No need for the mathematical definitions of problems.**
  - **Open to various types of search space.**
    - **Continuous optimisation.**
    - **Combinatorial optimisation.**
    - **Hybrid optimisation.**

- **Population-based Randomised Iterative Search.**
  - **Resistant to data noise.**
  - **Resistant to local optima.**
  - **Providing multiple solutions.**
    - **Multi-objective optimisation.**
    - **Finding multiple solutions.**



Evolutionary Learning

Evolutionary learning can be used in supervised, unsupervised and reinforcement learning.

1. **Learning classifier systems (rule-based systems).**
2. **Evolutionary artificial neural networks.**
3. **Evolutionary fuzzy logic systems.**
4. **Co-evolutionary learning.**
5. **Automatic modularisation of machine learning systems by speciation and niching.**

Evolutionary Design

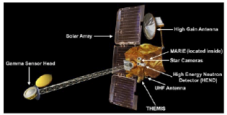
EC techniques are particularly good at exploring unconventional designs which are very difficult to obtain by hand.

1. **Evolutionary design of artificial neural networks.**
2. **Evolutionary design of electronic circuits.**
3. **Evolvable hardware.**
4. **Interactive creative design using evolutionary approaches**

Evolutionary Computation Theory

It explains how, when and why EAs work.

Some Successful Examples of Evolutionary Computation



Aerospace

G. S. Hornby et al., Automated antenna design with evolutionary algorithms. American Institute of Aeronautics and Astronautics, 2006.

Logistic

Thomas Weise, Alexander Podlich, Kai Reinhard, Christian Gorldt, and Kurt Geihs (2009): "Evolutionary Freight Transportation Planning," in Applications of Evolutionary Computing

Architecture

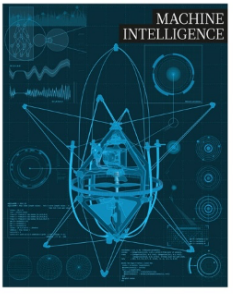
Ludger Hovestadt. Beyond the Grid - Architecture and Information Technology. Applications of a Digital Architectonic. Birkhäuser Basel / Boston 2009.

Robotics

Zykov V., Mytilinaios E., Adams B., Lipson H. (2005) "Self-reproducing machines", Nature Vol. 435 No. 7038, pp. 163-164

Become an Important Branch of AI

- **EC has been listed as one out of six representative techniques of machine intelligence by *Nature* in 2015.**



EDITORIAL Top

Machine intelligence
Tanguy Chouard & Liesbeth Venema
Nature **521**, 435 (28 May 2015)

ARTICLES Top

Deep learning
Yann LeCun, Yoshua Bengio & Geoffrey Hinton
Nature **521**, 436–444 (28 May 2015)

Probabilistic machine learning and artificial intelligence
Zoubin Ghahramani
Nature **521**, 452–459 (28 May 2015)

Design, fabrication and control of soft robots
Daniela Rus & Michael T. Tolley
Nature **521**, 467–475 (28 May 2015)

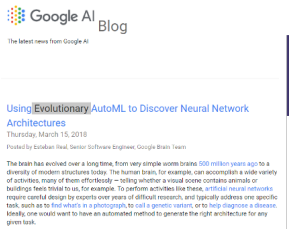
Reinforcement learning improves behaviour from evaluative feedback
Michael L. Littman
Nature **521**, 445–451 (28 May 2015)

Science, technology and the future of small autonomous drones
Dario Floreano & Robert J. Wood
Nature **521**, 460–466 (28 May 2015)

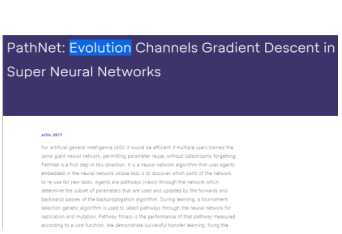
From evolutionary computation to the evolution of things
Agoston E. Eiben & Jim Smith
Nature **521**, 476–482 (28 May 2015)

Attentions from Government And Leading IT Companies

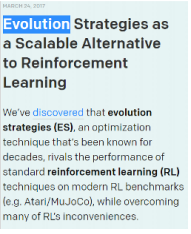
- **EC has been listed in the Commerce Control List of US Department of Commerce at Nov. 2018, placed the 2^{nd} technique in AI field.**
  (From "美国商务部受出口管制的代表性新兴技术清单征求意见稿"*)
- **Recently, leading IT companies have put a lot on EC.**



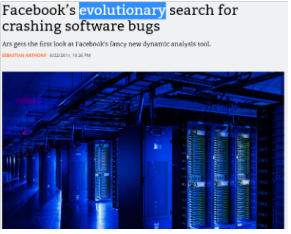
(1) Google AI



(2) DeepMind



(3) Open AI



(4) Facebook

(*) https://www.govinfo.gov/content/pkg/FR-2018-11-19/pdf/2018-25221.pdf
(1) https://ai.googleblog.com/2018/03/using-evolutionary-automl-to-discover.html
(2) https://deepmind.com/research/publications/pathnet-evolution-channels-gradient-descent-super-neural-networks/
(3) https://blog.openai.com/evolution-strategies/#content
(4) https://arstechnica.com/information-technology/2017/08/facebook-dynamic-analysis-software-sapienz/

Outline of This Lecture

Summary

1. **EC is a field of study that includes EAs and other areas. EAs include many types of algorithms.**
2. **EAs can be regarded as population-based generate-and-test algorithms.**
3. **EC techniques can be used in optimisation, learning and design.**
4. **EC techniques are flexible and robust.**
5. **EC techniques are definitely useful tools in your toolbox, but there are problems for which other techniques might be more suitable.**

Outline of This Lecture

Essential Reading for This Lecture

1. **D. B. Fogel (ed.), *Evolutionary Computation: The Fossil Record*, IEEE Press, Piscataway, NJ, USA, 1998. (ISBN-10: 0780334817, ISBN-13:978-0780334816)**
   - **It is essential that you read the introduction and comments, if you do not have time to read all the classical papers there.**
   - **It is extremely useful to know the history and origin. Not only do we want to learn brilliant ideas, we also need to learn how they were first conceived and generated.**

2. **X. Yao, *Evolutionary computation: A gentle introduction*, In *Evolutionary Optimization*, R. Sarker, M. Mohammadian and X. Yao (eds.), Chapter 2, pp.27-53, Kluwer Academic Publishers, Boston, 2002. (ISBN 0-7923-7654-4)**

Other References I

[1] A. M. Turing, "Computing machinery and intelligence," *Mind*, vol. LIX, no. 49, pp. 433–460, 1950.

[2] J. H. Holland, "Genetic algorithms," *Scientific American*, vol. 267, no. 1, pp. 66–73, 1992.

[3] L. J. Fogel, "Autonomous automata," *Industrial Research*, vol. 4, pp. 14–19, 1962.

[4] L. J. Fogel, A. J. Owens, and M. J. Walsh, "Artificial intelligence through simulated evolution,", 1966.

[5] I. Rechenberg, "Cybernetic solution path of an experimental problem," *Royal Aircraft Establishment Library Translation 1122*, 1965.

[6] H.-P. Schwefel, "Kybernetische evolution als strategie der experimentellen forschung in der stromungstechnik," *Diploma thesis, Technical Univ. of Berlin*, 1965.

Other References II

[7] H. De Garis, "Genetic programming: Building artificial nervous systems using genetically programmed neural network modules," in *Machine Learning Proceedings 1990*, Elsevier, 1990, pp. 132–139.

[8] J. R. Koza, "Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems,", vol. 34, 1990.

[9] R. Storn, "On the usage of differential evolution for function optimization," in *Fuzzy Information Processing Society, 1996. NAFIPS. 1996 Biennial Conference of the North American*, IEEE, 1996, pp. 519–523.

[10] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of IEEE International Conference on Neural Networks*, 1995, 1942–1948.