Computational Intelligence

LMU Munich
winter term 2024/2025

Thomas Gabor
Claudia Linnhoff-Popien

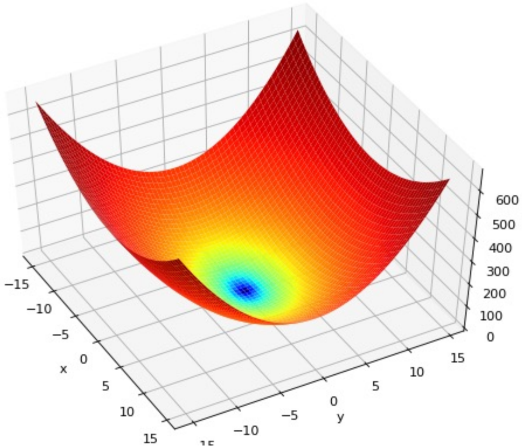
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

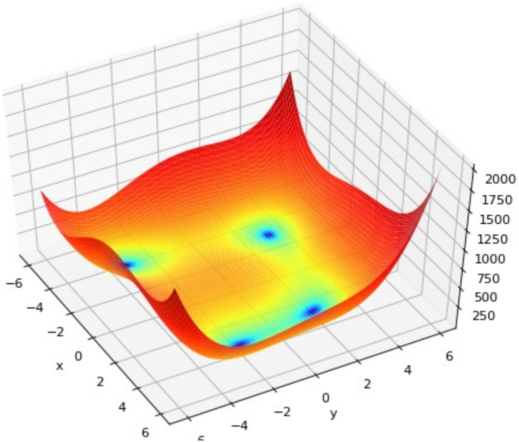
Definition 2 (optimization)**.** Let $\mathcal{X}$ be an arbitrary state space. Let $\mathcal{T}$ be an arbitrary set called target space and let $\leq$ be a total order on $\mathcal{T}$. A total function $\tau : \mathcal{X} \to \mathcal{T}$ is called target function. Optimization (minimization/maximization) is the procedure of searching for an $x \in \mathcal{X}$ so that $\tau(x)$ is optimal (minimal/maximal). Unless stated otherwise, we assume minimization. An optimization run of length $g+1$ is a sequence of states $\langle x_t \rangle_{0 \leq t \leq g}$ with $x_t \in \mathcal{X}$ for all t.

Let $e : \langle \mathcal{X} \rangle \times (\mathcal{X} \to \mathcal{T}) \to \mathcal{X}$ be a possibly randomized or non-deterministic function so that the optimization run $\langle x_t \rangle_{0 \leq t \leq g}$ is produced by calling e repeatedly, i.e., $x_{t+1} = e(\langle x_u \rangle_{0 \leq u \leq t}, \tau)$ for all t, $1 \leq t \leq g$, where x_0 is given externally (e.g., $x_0 =_{def} 42$) or chosen randomly (e.g., $x_0 \sim \mathcal{X}$). An optimization process is a tuple $(\mathcal{X}, \mathcal{T}, \tau, e, \langle x_t \rangle_{0 \leq t \leq g})$.

Definition 3 (optimization (policy))**.** Let $\mathcal{X} = \Pi$ be a policy space. Let $\mathcal{D} = (\Pi, \mathcal{T}, \tau, e, \langle x_t \rangle_{0 \leq t \leq g})$ be an optimization process according to Definition 2. $\mathcal{D}$ is called a policy optimization process.

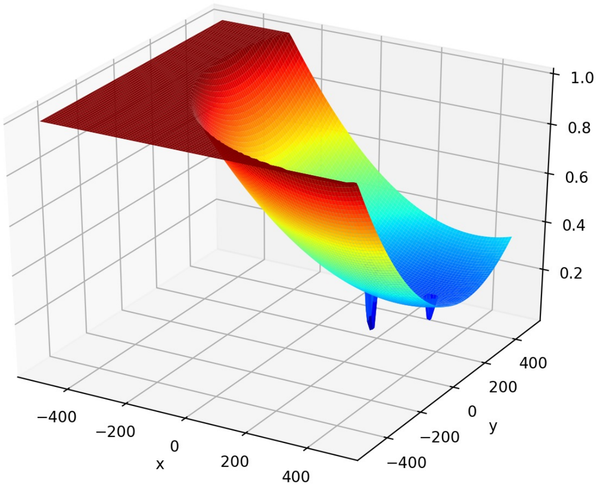
The Solution Landscape Metaphor

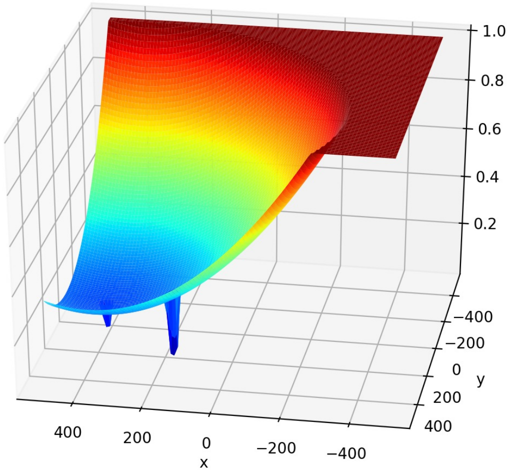




 source: deap.readthedocs.io/en/master/api/benchmarks.html

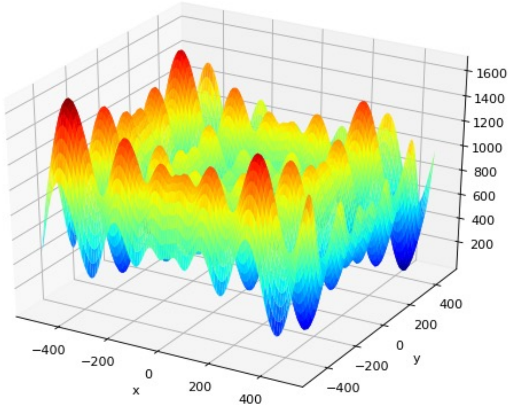
The Solution Landscape Metaphor

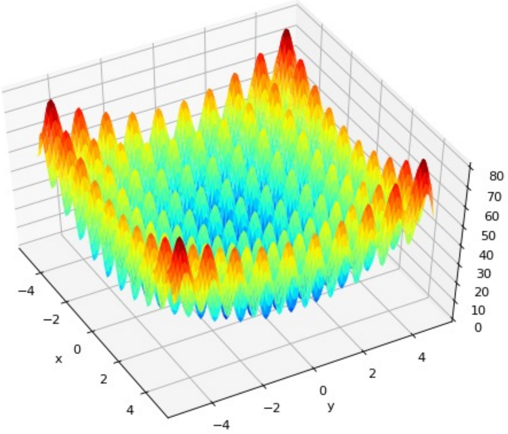




 source: deap.readthedocs.io/en/master/api/benchmarks.html

The Solution Landscape Metaphor





 source: deap.readthedocs.io/en/master/api/benchmarks.html

Algorithm 3 (simulated annealing)**.** Let $\mathcal{D} = (\mathcal{X}, \mathcal{T}, \tau, e, \langle x_u \rangle_{0 \leq u \leq t})$ be an optimization process. Let *neighbors* : $\mathcal{X} \to \wp(\mathcal{X})$ be a function that returns a set of neighbors of a given state $x \in \mathcal{X}$. Let $\kappa : \mathbb{N} \to \mathbb{R}$ be a temperature schedule, i.e., a function that returns a temperature value for each time step. Let $A : \mathcal{T} \times \mathcal{T} \times \mathbb{R} \to \mathbb{P}$ with $\mathbb{P} = [0; 1] \subset \mathbb{R}$ be a function that returns an acceptance probability given two target values and a temperature. Commonly, we use

$$A(Q, Q', K) = e^{\frac{-(Q'-Q)}{K}}$$

for $\mathcal{T} \subseteq \mathbb{R}$. The process $\mathcal{D}$ continues via simulated annealing if e is of the form

$$e(\langle x_u \rangle_{0 \leq u \leq t}, \tau) = x_{t+1} = \begin{cases} x'_t & \text{if } \tau(x'_t) \leq \tau(x_t) \text{ or } r \leq A\big(\tau(x_t), \tau(x'_t), \kappa(t)\big), \\ x_t & \text{otherwise,} \end{cases}$$

where $x'_t \sim neighbors(x_t)$ and $r \sim \mathbb{P}$ are drawn at random for each call to e.

other means of biased sampling

Definition 4 (population-based optimization)**.** Let $\mathcal{X}$ be a state space. Let $\mathcal{T}$ be a target space with total order $\leq$. Let $\tau : \mathcal{X} \to \mathcal{T}$ be a target function. A tuple $\mathcal{E} = (\mathcal{X}, \mathcal{T}, \tau, E, \langle X_t \rangle_{0 \leq t \leq g})$ is a population-based optimization process iff $X_t \in \wp^*(\mathcal{X})$ for all t and $E : \langle \wp^*(\mathcal{X}) \rangle \times (\mathcal{X} \to \mathcal{T}) \to \wp^*(\mathcal{X})$ is a possibly randomized, non-deterministic, or further parametrized function so that the population-based optimization run is produced by calling E repeatedly, i.e., $X_{t+1} = E(\langle X_u \rangle_{0 \leq u \leq t}, \tau)$ where X_0 is given externally or chosen randomly.

Definition 5 (population-based optimization (alternate))**.** An optimization process $\mathcal{E} = (\mathcal{X}, \mathcal{T}, \tau, E, \langle X_t \rangle_{0 \leq t \leq g})$ is called population-based iff $\mathcal{X}$ has the form $\mathcal{X} = \wp^*(\mathcal{Y})$ for some other state space $\mathcal{Y}$.

Multi-Sets

Let X be a set with members from a space $\mathcal{X}$.
We can identify a set X with its membership function $\mu_X : \mathcal{X} \to \mathbb{B}$ such that

$$\mu_X(x) = 1 \iff x \in X.$$

For a multiset X' we generalize the membership function so that

$$\mu_{X'}(x) \geq 1 \iff x \in X'.$$

In both cases, the subset relation between (multi-)sets X, Y can be defined as

$$X \subseteq Y \iff \forall x \in X : \mu_X(x) \leq \mu_Y(x).$$

$$\begin{aligned} \text{We then define } \wp(X) = 2^X = {} & \mathbb{B}^X \mathrel{\widehat{=}} X \to \mathbb{B}, \\ \text{and } \wp^*(X) = {} & \mathbb{N}^X \mathrel{\widehat{=}} X \to \mathbb{N}, \end{aligned}$$

to be the powerset functions for sets and multisets, respectively.

Multi-Sets

Evolutionary Algorithms

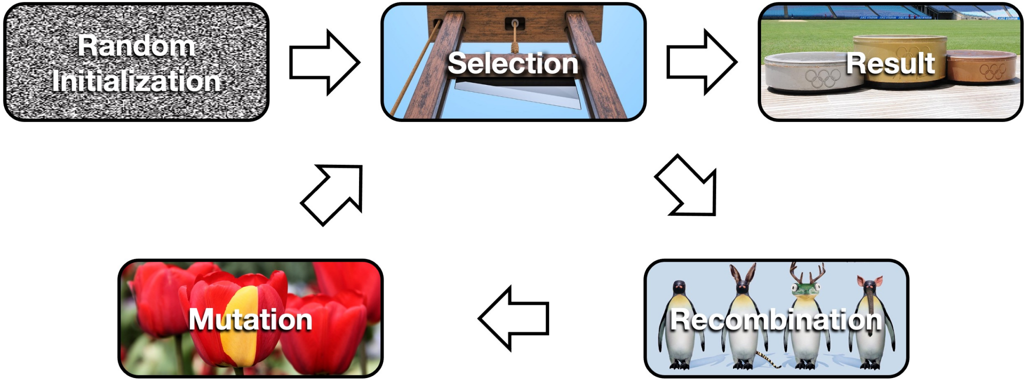


image sources:
www.bostonmagazine.com/news/2015/07/30/boston-2024-winners-losers
en.wikipedia.org/wiki/Mutation#/media/File:Darwin_Hybrid_Tulip_Mutation_2014-05-01.jpg
www.heise.de/ct/artikel/Die-Woche-Microsoft-und-Linux-1283059.html
https://phys.org/news/2019-10-guillotine-cruel-poisoning.html

Algorithm 4 (basic evolutionary algorithm)**.** Let $\mathcal{E} = (\mathcal{X}, \mathcal{T}, \tau, E, \langle X_u \rangle_{0 \leq u \leq t})$ be a population-based optimization process. The process $\mathcal{E}$ continues via an evolutionary algorithm if E has the form

$$E(\langle X_u \rangle_{0 \leq u \leq t}, \tau) = X_{t+1} = \mathit{selection}\big(X_t \uplus \mathit{variation}(X_t)\big)$$

where *selection* and *variation* are possibly randomized or non-deterministic functions so that for any $X \in \wp^*(\mathcal{X})$ it holds that $|\mathit{selection}(X)| \leq |X|$ and $\big|\mathit{selection}\big(X \uplus \mathit{variation}(X)\big)\big| = |X|$.

Which optimization algorithm
is the best?