Stochastic Optmization - 1

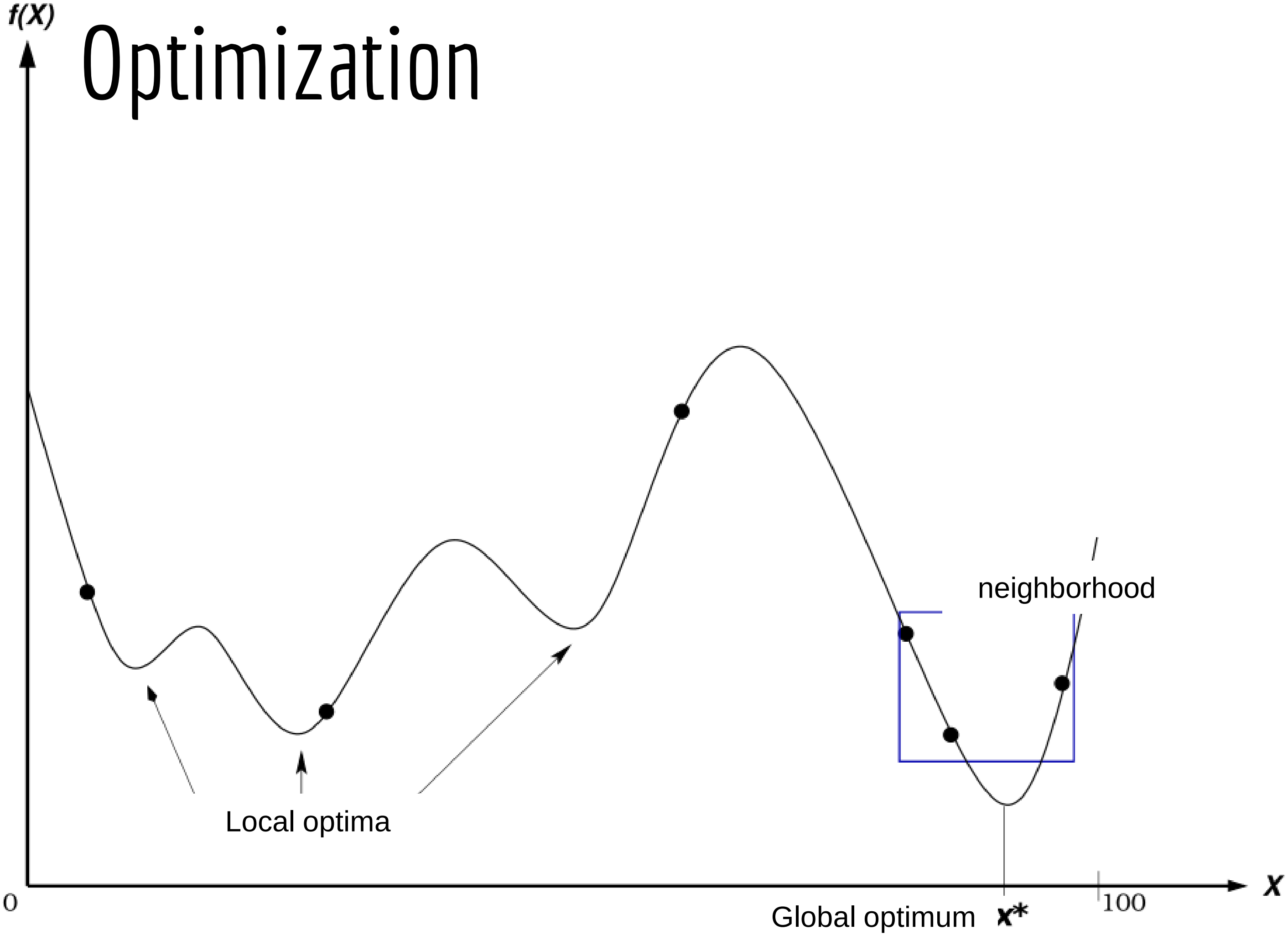
Simulated Annealing

EISTI
Cergy - Pau

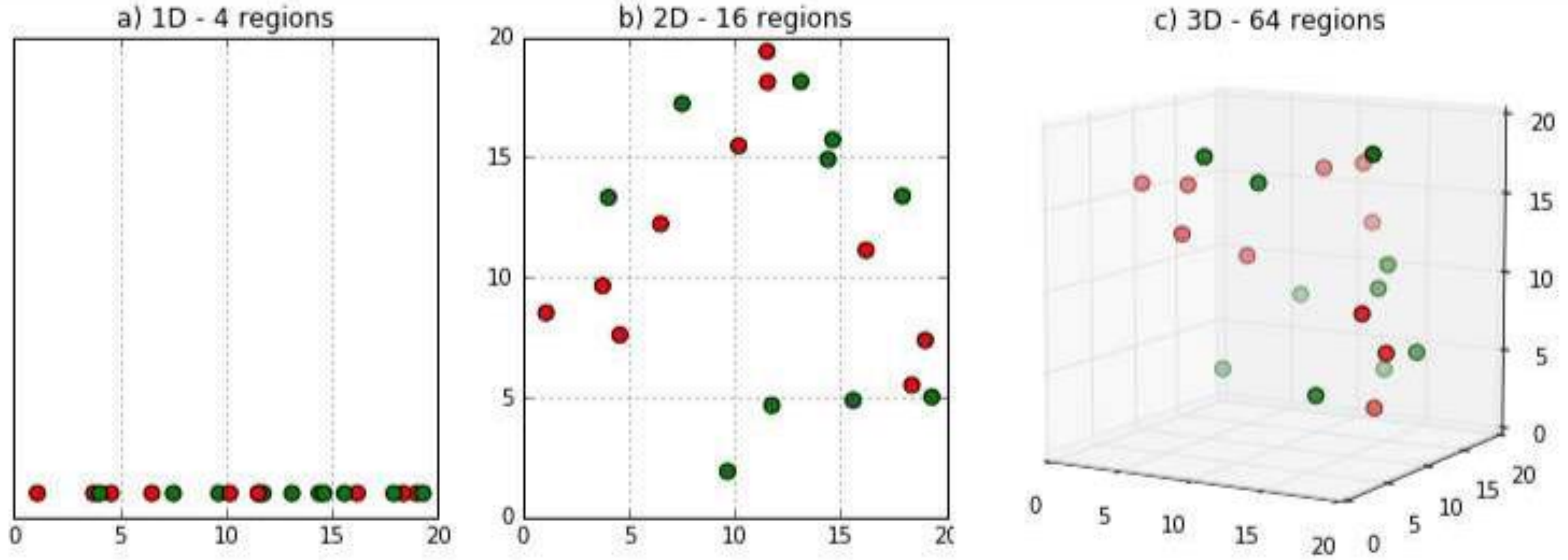
Mathematical modeling

- Define the decision variables
  - set of variables that govern the situation to be modeled
  - real, integer, binary variables
- Specify the objective function
  - mathematical function composed of decision variables that represents the modeled physically system
  - linear, non-linear function
- Specify the constraints of the problem
  - set of parameters that limit the achievable (feasible) model
  - equations or inequations composed of decision variables
- Specify the parameters of the model
  - constants associated with the constraints and the objective function

Optimization



The curse of dimension



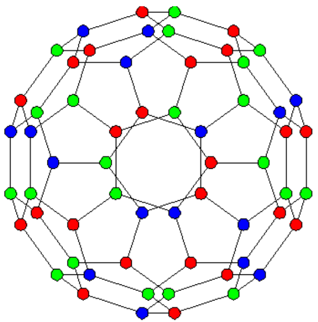
http ://cleverowl.uk/2016/02/06/curse-of-dimensionality-explained/

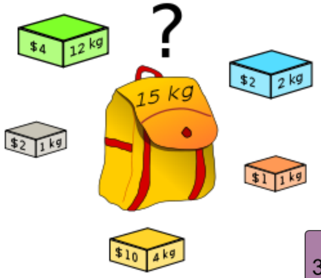
Typology of problems

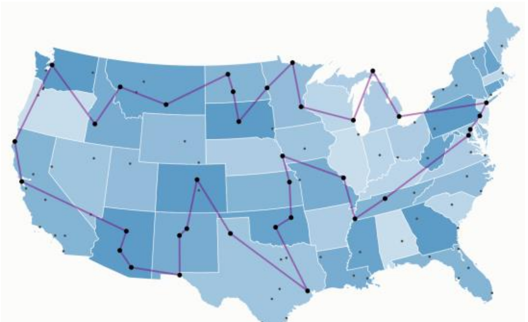
- Hard optimization
- Minimization / maximization problem;
- Single-objective, multi-objective, multi-objective problems;
- Problems with discrete / continuous / mixed variables;
- Problems under constraints;
- high-dimensional problems
- dynamic problems.

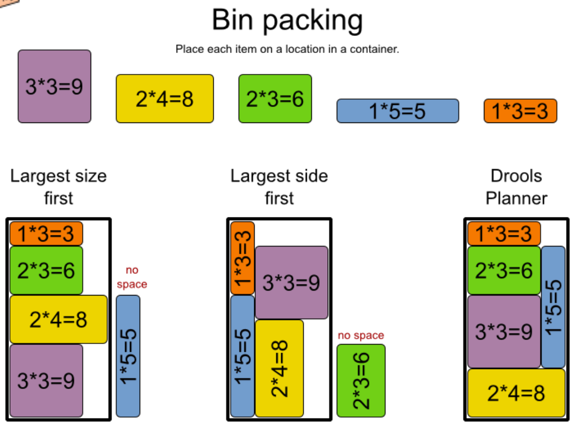
Typology of problems with discrete variables

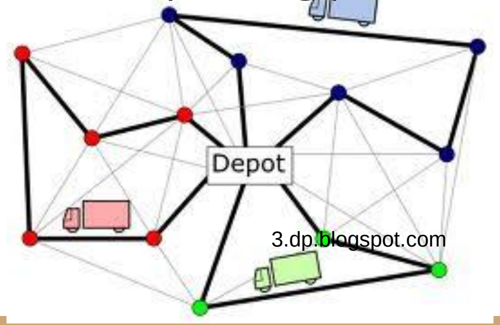
- Graph coloring problem
- Backpack problem
- Traveler salesman problem
- Vehicle routing problem
- Planning problem
- Bin packing problem


mathworks.com

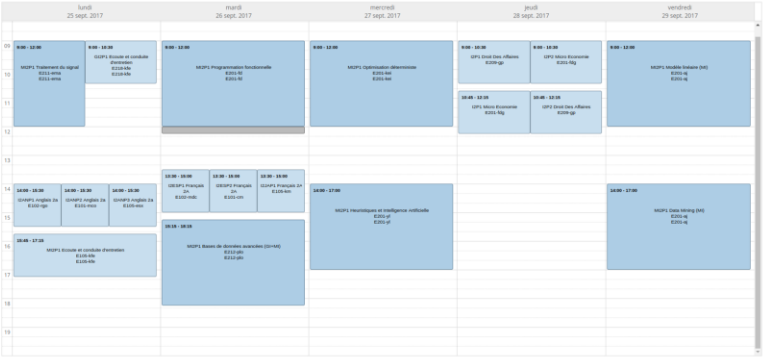



gurobi.com



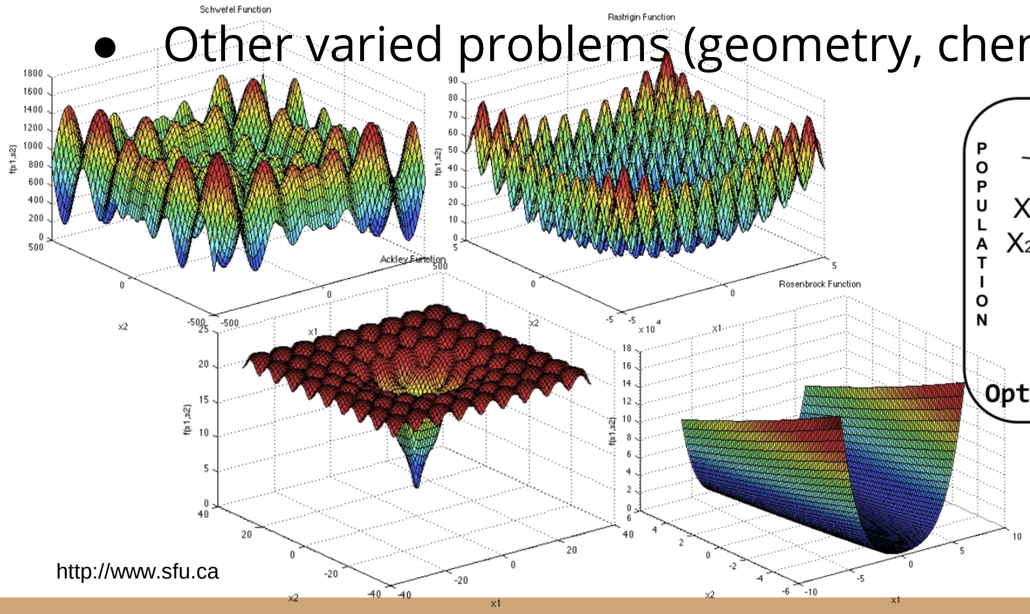


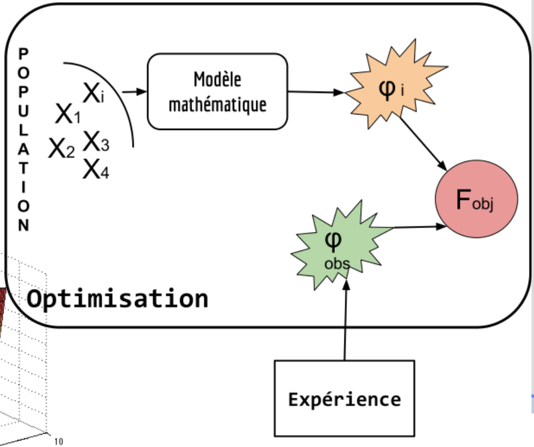
3.dp.blogspot.com

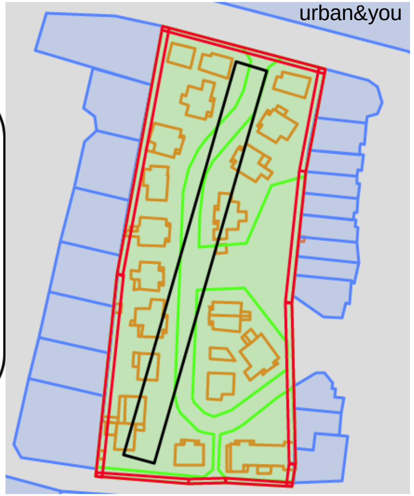


Typology of problems with continues variable

- Analytical function (benchmarks);
- Non-analytical expression (calculation code, experience résults, inverse problem);
- Other varied problems (geometry, chemistry, physics) ...







Metaheuristics

These are:

- global;
- stochastic;
- generic even if dependent on the continuous / discrete context;
- that do not guarantee optimality;
- based on analogies with biology, physics, ethology

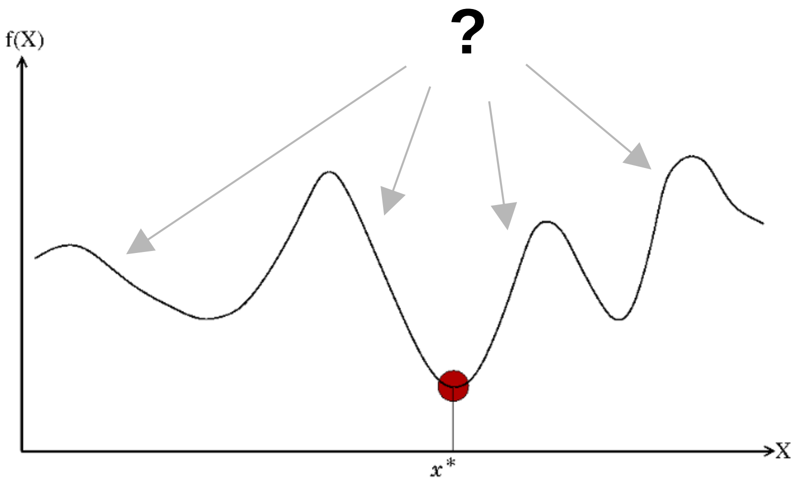
Components and strategies

They are generally composed of:

- One or more (population) initial solutions;
- A generating strategy of new solution based on randomness;
- A criterion of acceptance of a new solution;
- behavioral control coefficients of the algorithm;
- Information sharing, the use of a memory of explored solutions;
- A convergence criterion.

initialization

- Randomly
- Smart way
  - sampling
  - Latin hypercubes
  - Halton sequence
  - entropy
  - ...



Research

Exploration

VS

Exploitation

neighborhood

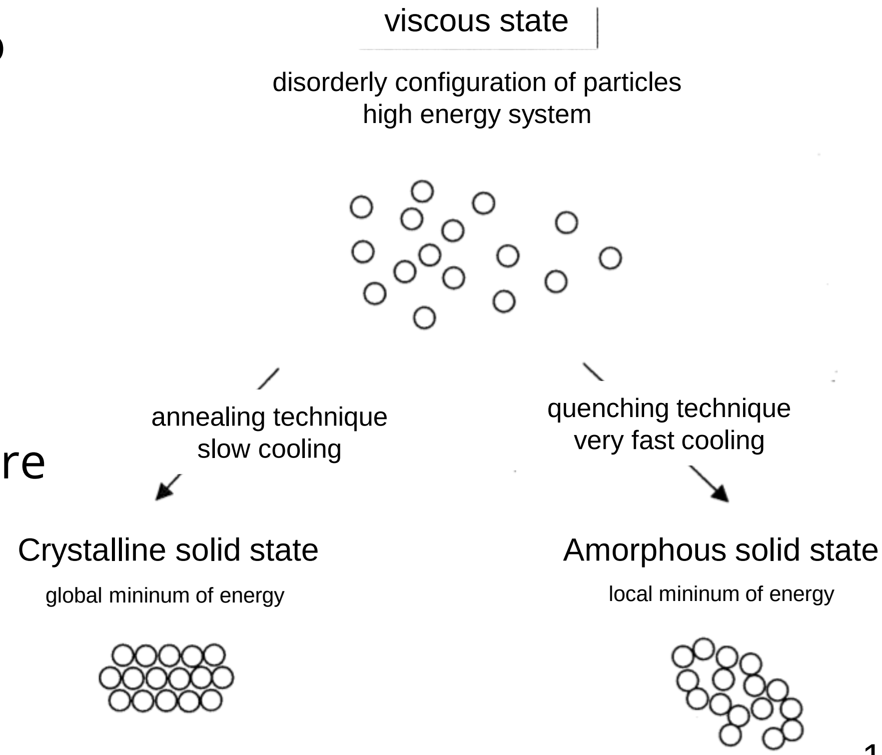
- Definition:
  - Case of discrete variables: offset of data components,
  - Case of continuous variables: hyper (sphere | cube) centered around the data.
- Strategies
  - Geographical neighborhood,
  - Social neighborhood,
  - Random neighborhood.
- Gradient Methods, Newton-Raphson, secant method, Dichotomy, Nelder-Mead Polytope…
- Hybridization with local search methods

The simulated annealing

For history ... a little physical!

Annealing is a technique that allows to improve the quality of a material according to the following method:

- It is heated to a very high temperature for liquefying
- Gradually lowering the temperature to stabilize the structure of the material



[Métaheuristiques pour l'optimisation difficile, P. Siarry et al.]

Simulated Annealing (SE)

- Metaheuristic variant of the Metropolis-Hastings algorithm;
- Proposed in 83 by Kirkpatrick, Gelatt and Vecchi and in 85 by Cerny;
- First proposed metaheuristic;
- Adapted to discrete problems (originally electronic component placement on a plate).

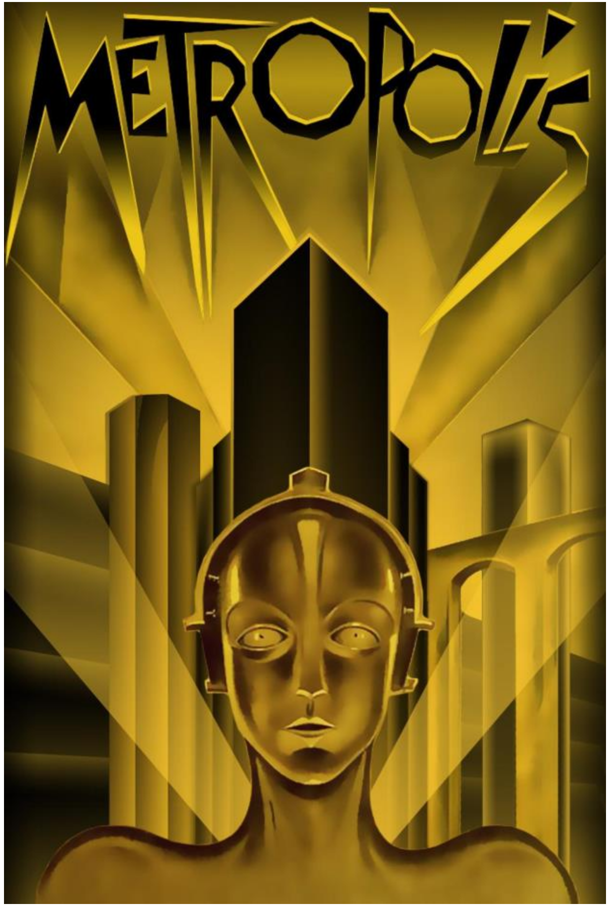
Simulated annealing - the analogy

- The function f to be minimized is the energy of the system;
- A solution X represents a state of the material;
- The thermodynamic equilibrium is reached during a temperature step;
- At temperature T, a perturbation of the current solution is accepted with a probability based on the Metropolis criterion.

Criterion of Metropolis

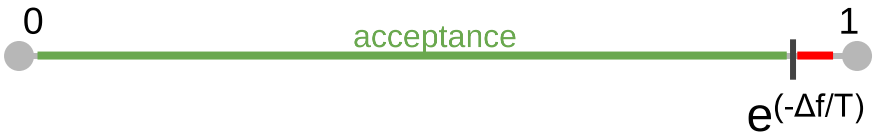
```
MetropolisCriterion(Δf,T) :
     if  Δf ≤ 0 then
         return TRUE
    else
         return rand(0,1) < e^(-Δf/T)
```

- $\Delta f \leq 0$, the neighbor is accepted;
- A small variation has more chance to be accepted than an important one;
- This function is stochastic.

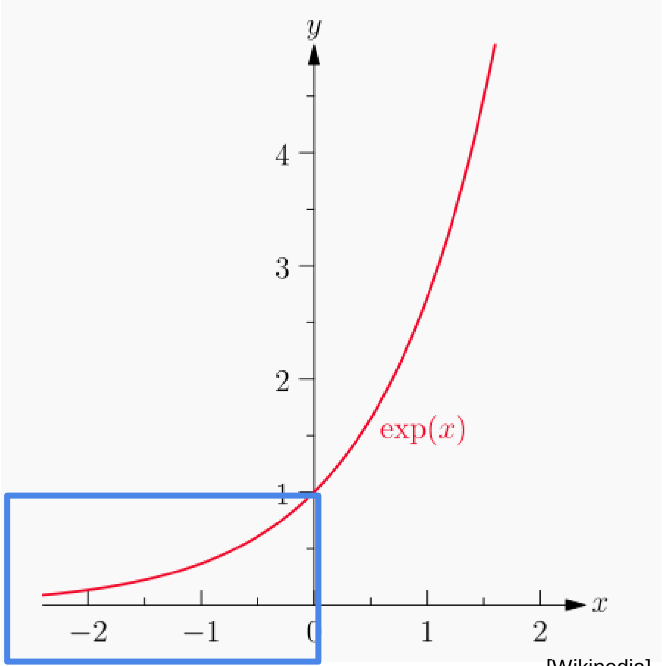


Criterion of Metropolis

- At high temperature, the Metropolis criterion value is close to 1, most degrading solutions are accepted:



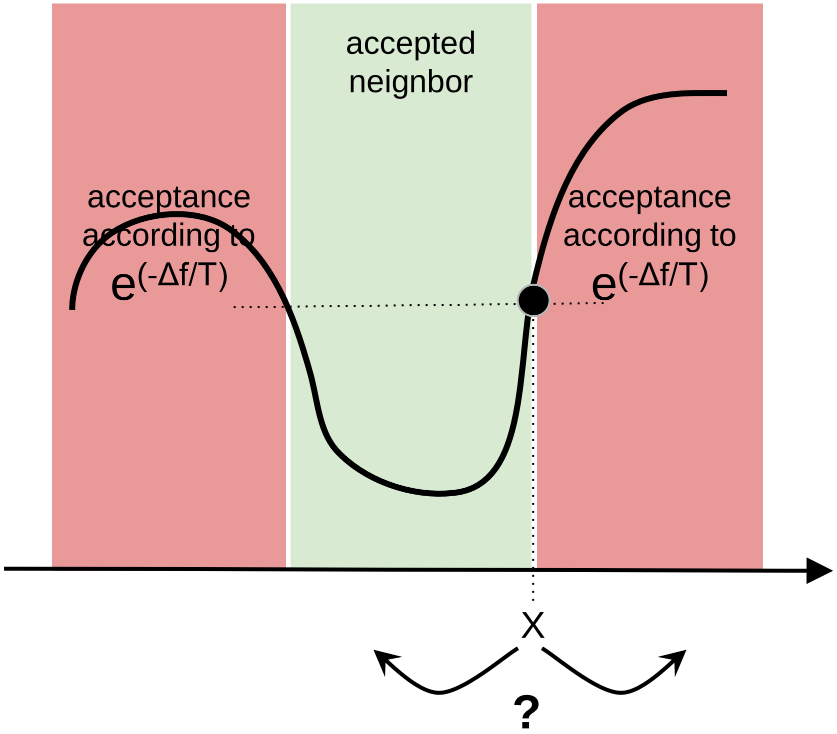
- At low temperatures, the Metropolis criterion value tends to 0, most degrading solutions are rejected:

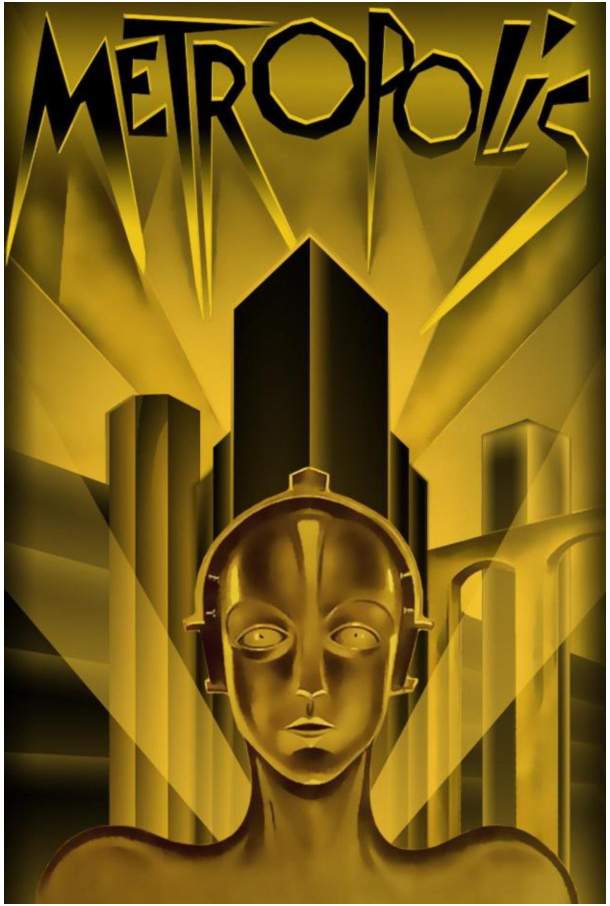


[Wikipedia]

Variation space of the Metropolis criterion

Criterion of





Algorithm of simulated annealing

```
X, one solution, f(X) energy of the system, T initial temperature
f_min <- f(X)
X_min <- X
while T > T_min or not convergenceCriterion()
       while not thermodynamicEquilibrium () // step of temperature
                    X_neigh <- disturbance (X)
                    Δf = f(X_neigh) - f(X)
             if Δf < 0 or acceptanceMetropolisCriterion (Δf,T) then
                    if Δf < 0 et f(X_neigh) < f_min then
                       f_min <- f(X_neigh)
                        X_min <- X_neigh
                    end of if
                    f(X) <- f(X_neigh)
                     X <- X_neigh
              end of if
       end of while
       T <- cooling(T)
End of while
```

Simulated Annealing: Convergence Criteria

Allows to complete the algorithm according to several conditions:

- The temperature reaches a minimum value;
- The number of evaluations reaches a limit;
- There has been no improvement since a number of iterations.

Simulated Annealing: Thermodynamic Equilibrium

Allows to “search” around an actual good solution :

- The number of evaluations reaches a limit;
- There has been no improvement since a number of iterations.

Simulated Annealing: Cooling

Allows to reduce the probability of acceptance of a non-improvement solution according to the Metropolis criterion:

- A strong decrease prevents suitable exploration, the algorithm converges rapidly to a local minimum;
- A small decrease allows a strong exploration during the first iterations of the algorithm;
- At $T = 0$, no degrading solution is selected; at $T = \infty$, any solution is accepted..

Simulated Annealing: Cooling

Different patterns of temperature reduction:

- Geometric (if we do not consider the temperature steps ...): $T_{k+1} = \alpha . T_k$, the most commonly used;
- Logarithmic: $T_k = \mu / \log(1 + k)$, where k: number of steps and μ a constant. Very expensive in computing time, little used;
- Exponential: $T_k = T_0 . \exp(-k / \tau)$, where k: nb of temperature steps and τ a constant;
- Esoteric: it can increase the temperature according to a particular criterion.

Simulated Annealing: Disturbance

- Generates a new solution in the neighborhood of the current solution;
- Influence of the distance between these two solutions (hypervolume of the neighborhood);
- Specific to the problem to be solved (discrete / continuous).

That’s all folks !