15-382 Collective Intelligence – S18

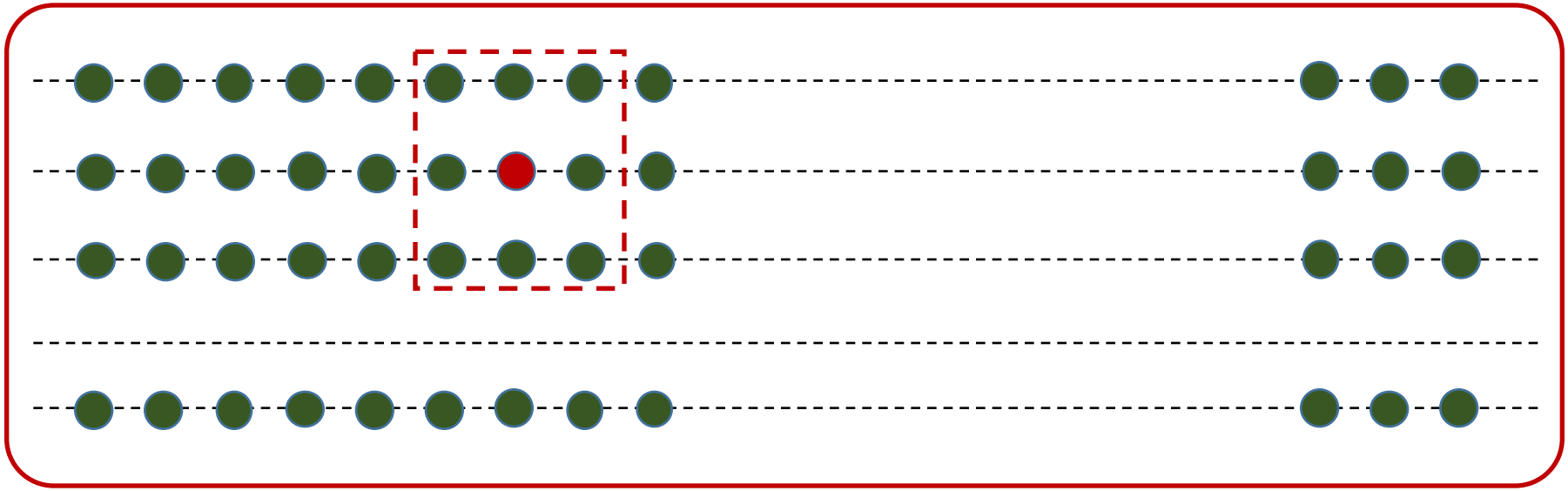
Lecture 15:
Swarm Intelligence 1 / Particle Swarm Optimization 1

Instructor:
Gianni A. Di Caro

جامعة كارنيجي ميلون في قطر
Carnegie Mellon University Qatar

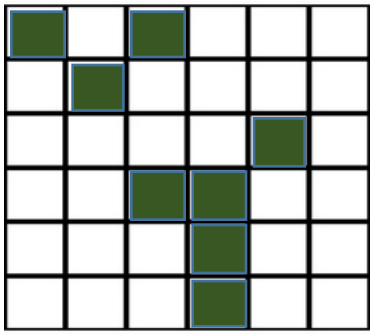
Limitations of the (classic) CA model

2D Environment

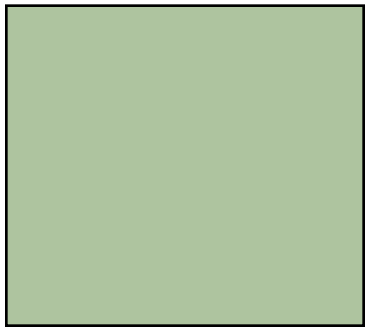


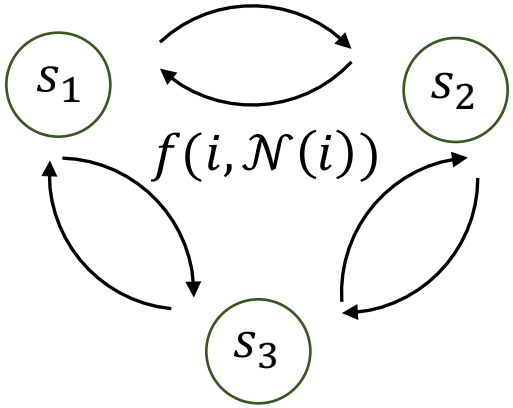
- **Space is discretized** according to an $n \times m$ regular lattice
- The CA is equivalent to $\boldsymbol{n \times m}$ ***coupled*** **iterated maps**
- Coupling is determined by the definition of **neighborhoods**, and by **boundary conditions** and **synchronization**
- Neighborhoods are statically defined based on the lattice's topology, and capture some notion of meaningful **spatial proximity**
- The CA is useful as a **simulation model** for a dynamical system, but has some obvious limitations for different types of use

From CA to PSO / SI



Discrete → **Continuous**

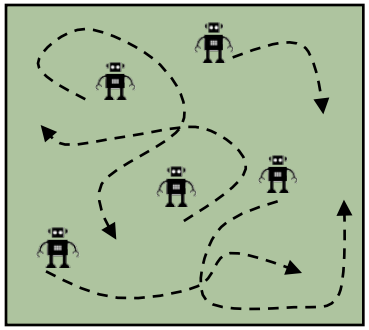


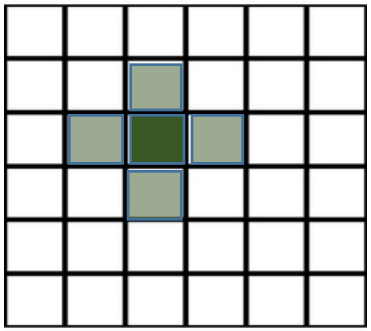


Spatially-fixed cell state ~ FSM

→ **Agent-model:**

- *internal state*
- *mobile*

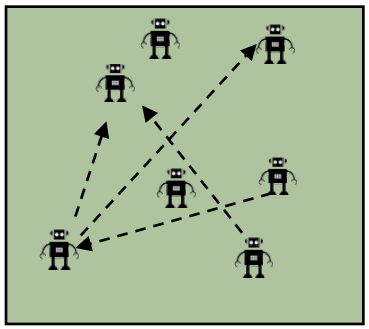




Spatially-related neighborhood (static)
Physical topology induced by the lattice

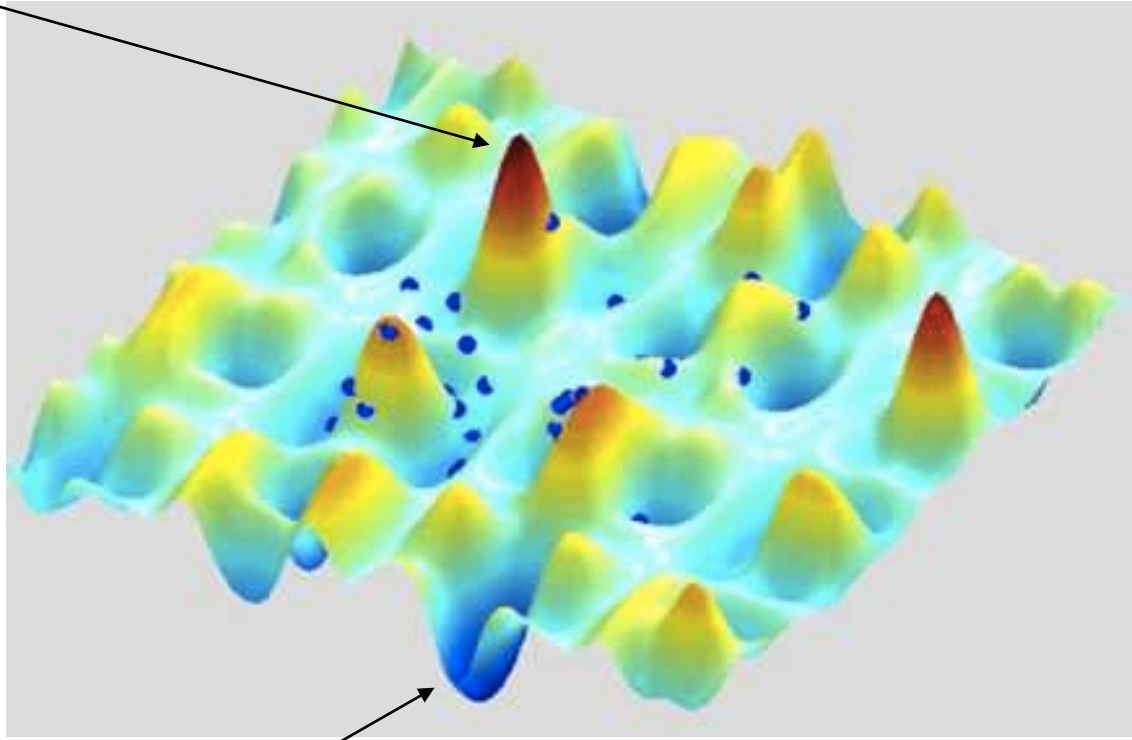
→ **Relational neighborhood**

- *Agents form a network*
- *Logical topology*
- *Can be dynamic*



Reference task: Global Function Optimization

Find the **global maximum** of the function $f(\boldsymbol{x})$
(and the point $\boldsymbol{x}$ where it happens)



Find the **global minimum** of the function

Optimization Problems

- Optimization problems expressed in mathematical form:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x})$$
$$\text{subject to } \boldsymbol{x} \in \mathcal{F}$$

  - $f\colon \mathbb{R}^n \longmapsto \mathbb{R}^m$ is the **objective function** (for now, m=1)
  - $\boldsymbol{x} \in X^n \subseteq \mathbb{R}^n$ is the **optimization vector variable**
  - $\mathcal{F} \subseteq X^n$ is the **feasible set** (constraints = values the variables can feasibly take)

- $\boldsymbol{x}^* \in X^n$ is an **optimal solution (global minimum)** if $\boldsymbol{x}^* \in \mathcal{F}$ and $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{F}$

- ***Mathematical programming* problem**

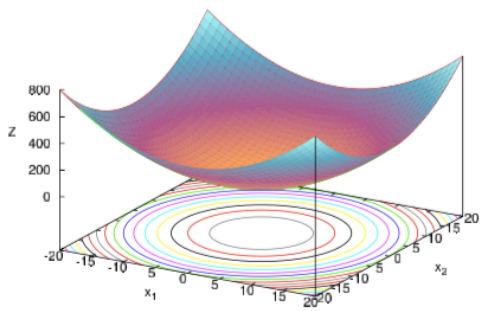
Basic properties

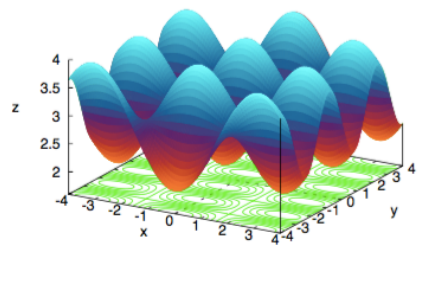
Given an optimization problem:

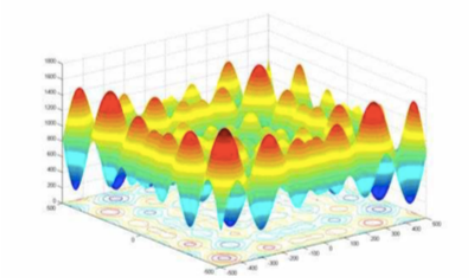
$$\min_{\boldsymbol{x}} f(\boldsymbol{x})$$
$$\text{subject to } \boldsymbol{x} \in \mathcal{F}, \qquad \boldsymbol{x} \in X^n \subseteq \mathbb{R}^n$$

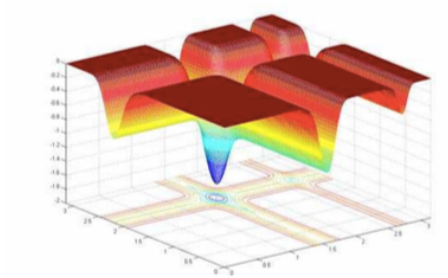
- $\min_{\boldsymbol{x}} f(\boldsymbol{x})$ is equivalent to $\max_{\boldsymbol{x}} -f(\boldsymbol{x})$
- If $\mathcal{F} = \emptyset$ the problem has no solution (unfeasible)
- If $\mathcal{F}$ is an open set, only the inf (sup) is guaranteed but not min (max)
- The problem is unbounded if $f \to -\infty$

Unconstrained vs. Constrained Opt







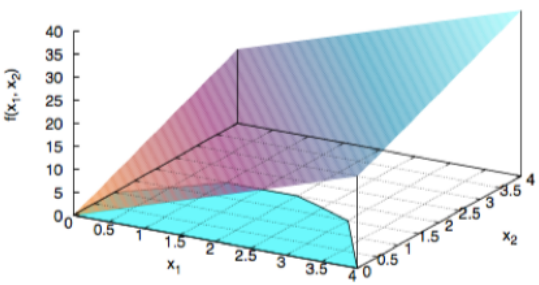


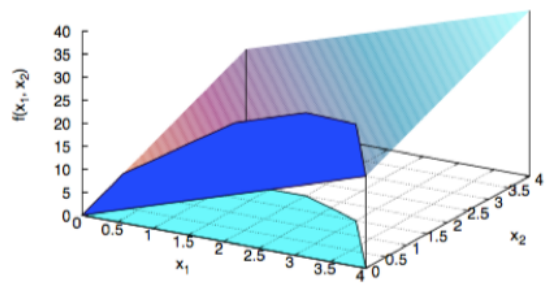
- Sublevel sets (isolines):

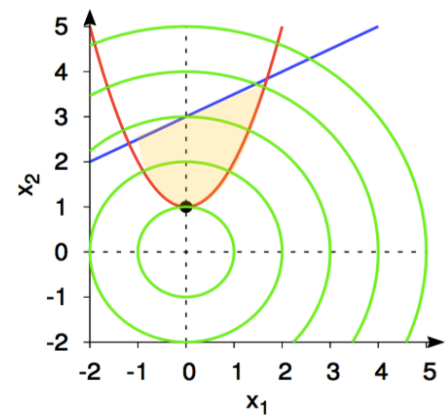
$$\{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) = c\}$$

$$\begin{aligned} \min_{\bar{x}} \quad & z = x_1^2 + x_2^2 \\ s.t. \quad & x_1 - 2x_2 + 6 \geqslant 0 \\ & -x_1^2 + x_2 - 1 \geqslant 0 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

$$\max\ z = f(x_1, x_2) = 5x_1 + 4x_2, \quad x_1, x_2 \in X = \text{set of linear constraints}$$





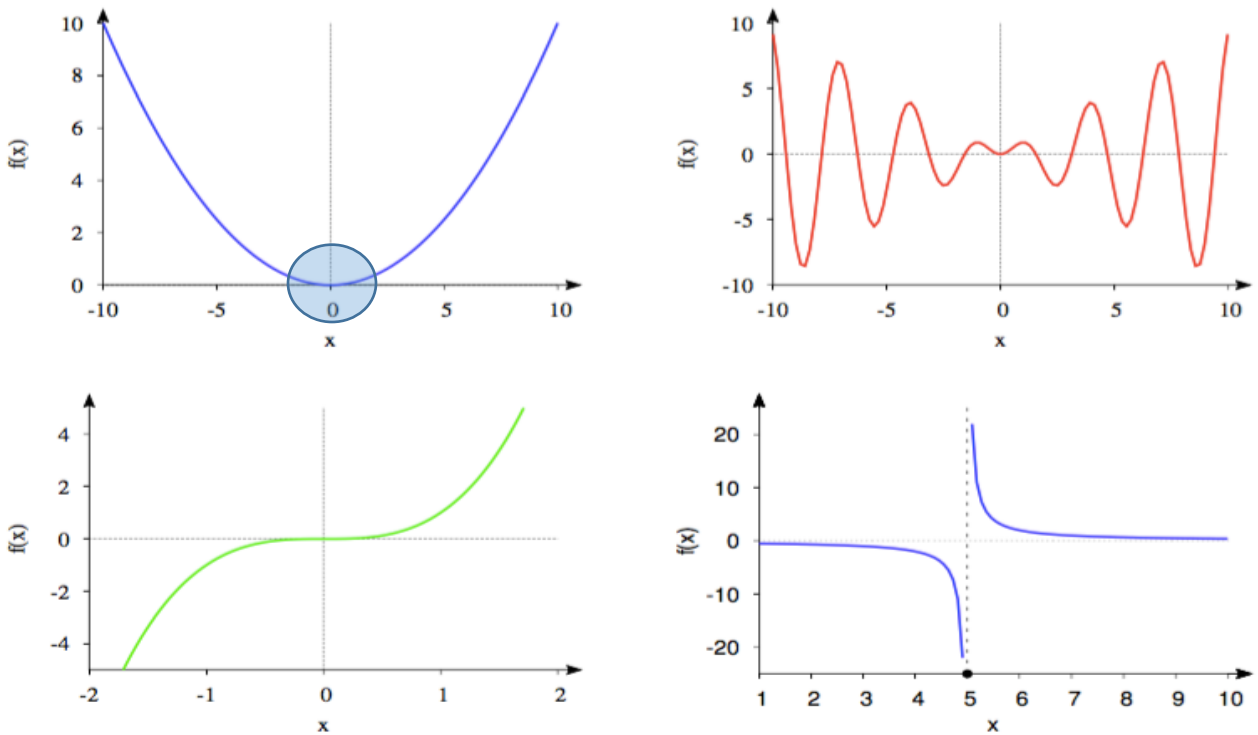


✓ Any constrained optimization problem can be formulated as an unconstrained one by including constraint violations as <u>penalty terms</u> in the objective function

GLOBAL AND LOCAL OPTIMALITY

- A point $\boldsymbol{x} \in X^n \subseteq \mathbb{R}^n$ is **globally optimal** (global minimum) if $\boldsymbol{x} \in \mathcal{F}$ and for all $\boldsymbol{y} \in \mathcal{F}$, $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$

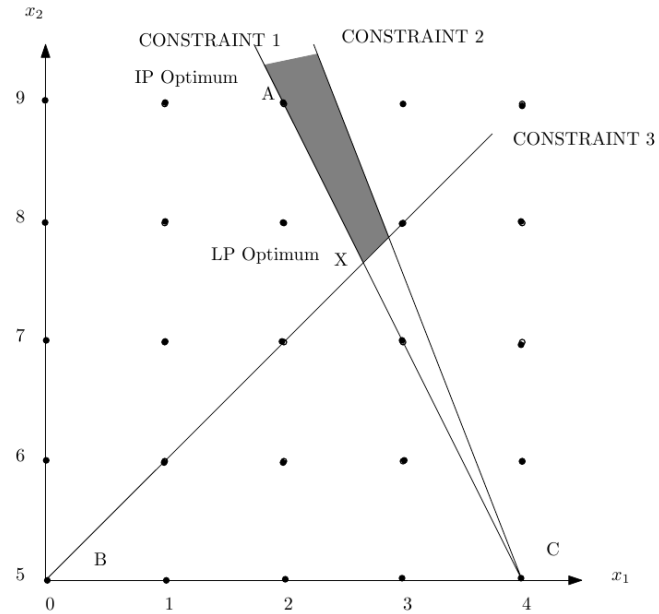
- A point $\boldsymbol{x} \in \mathbb{R}^n$ is **locally optimal** (local minimum) if $\boldsymbol{x} \in \mathcal{F}$ and there exists $\varepsilon > 0$ small such that for all $\boldsymbol{y} \in \mathcal{F}$ with $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon$, $f(\boldsymbol{x}) \leq f(\boldsymbol{y})$

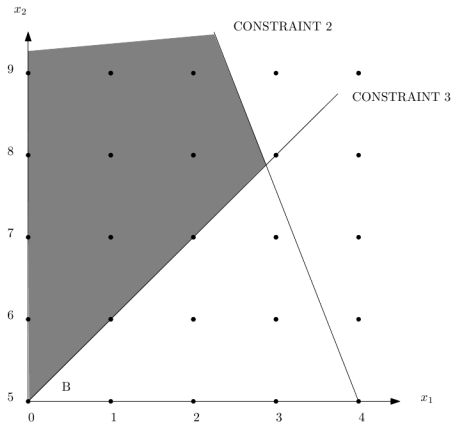


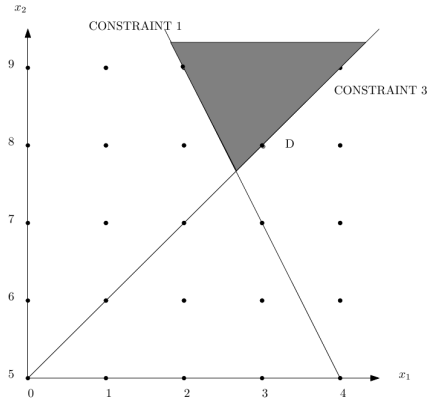
GLOBAL AND LOCAL OPTIMALITY

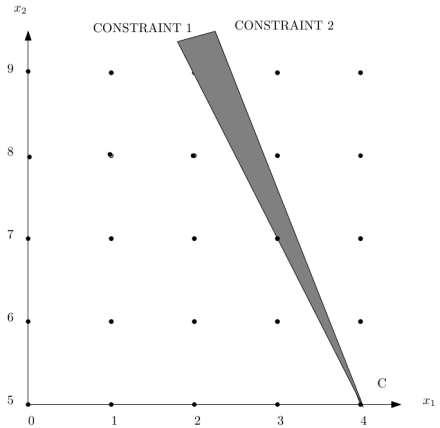
- What about discrete spaces? $\boldsymbol{x} \in X^n \subseteq \mathbb{Z}^n$

$$
\begin{aligned}
\min \quad & Z_{ILP} = x_2 \\
s.t. \quad & 2x_1 + x_2 \geqslant 13 \\
& 5x_1 + 2x_2 \leqslant 30 \\
& -x_1 + x_2 \geqslant 5 \\
& x_1,\ x_2 \in \mathbb{Z}^+
\end{aligned}
$$





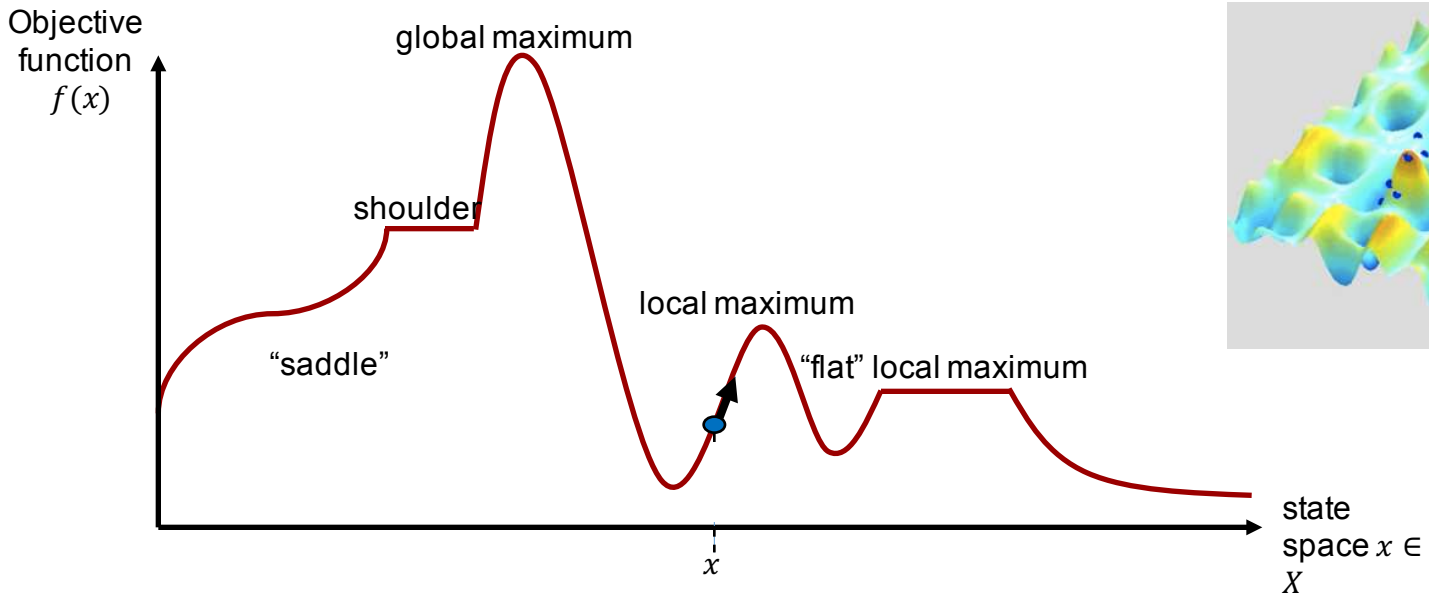


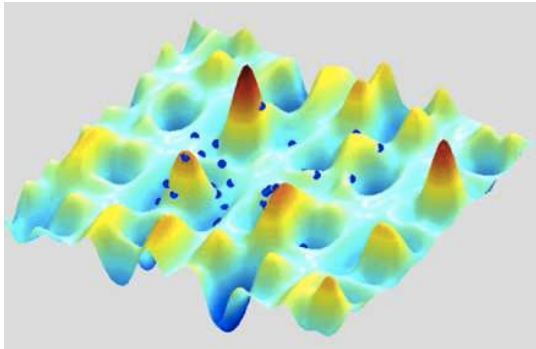


Black-box Optimization

- The function to optimize is *not given in algebraic form*, or
- The function is given, but it's not amenable to analytical treatment in terms of using its *derivatives* for finding min/max
- All we can do is to query the black box and observe $(\boldsymbol{x}, BB(\boldsymbol{x}))$ pairs...
- **Black-box / Derivative-free optimization** vs.
  White box optimization

How do we find minima / maxima?

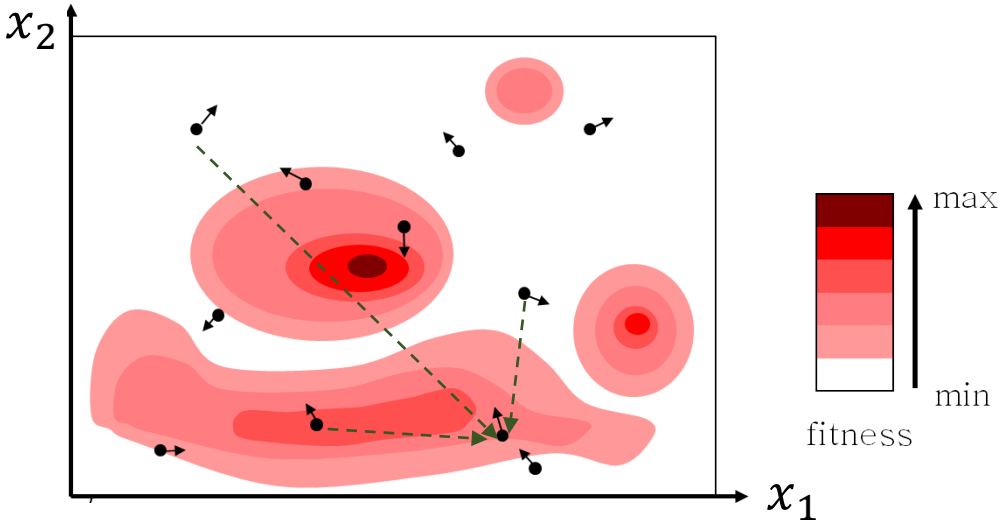




- Using <u>rates of change:</u> derivatives / gradients / Jacobians / Hessians / ...
- <u>Sampling / Searching in the $X^n \subseteq \mathbb{R}^n$ domain</u>, the input state space:
  - Figure out where to search / sample next, from the values that are returned from the function, *without generating a model of the function*
  - Using the sampled data to *generate a model of the function* and in turn, using it to iteratively direct the search
- Iteratively <u>constructing a solution</u>, by adding / trying out assignment to solution components $x_1, x_2, \ldots\ x_n$
- ... many, many variants and combinations of these two basic approaches...

Particle Swarm Optimization (PSO)

J. Kennedy, R. Eberhart, *Particle Swarm Optimization*. Proc. 4th IEEE Int. Conf. on Neural Networks, 1995.



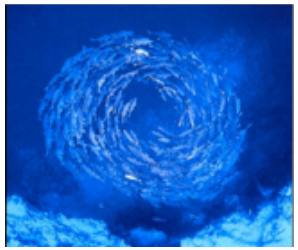
- **Multi-agent black-box optimization inspired by social and roosting behavior of flocking birds:**
  - Each agent (a *particle*) encodes a solution point $\boldsymbol{x}$
  - Agents move in $X^n \subseteq \mathbb{R}^n$, searching for the spots regions/points where the objective function gets its max (min) values
- **Individual swarm members establish a social network** and can profit from the discoveries and previous experience of the other members of the swarm:
  - Each agent iteratively changes its position (i.e., decides how to move) using information from personal past experience and from its social neighborhood

PSO is a first example of *Swarm Intelligence*

Swarm intelligence: Study and design of complex systems that:

- Are potentially made of a ***large number*** of components, a **swarm**
- Each component has **purpose(s)** (as for animals, or artificially designed agents), that implicitly contributes to the "*performance*" of the whole
- Under certain conditions, the system displays forms of ***swarm intelligence*** in terms of generation at the system-level, of effective spatio-temporal patterns and/or optimized decision-making and action-making

Modeling: *study of natural complex systems* with the above characteristics in order to identify the local rules that give raise to complex system-level behaviors and self-organization, make formal models



Engineering: *bottom-up design of artificial systems* that display useful system-level behaviors, possibly, but not necessarily, *taking inspiration* from the natural systems

Mimicking nature:

Bio-mimetic (algorithms, robots)

Bottom-up vs. Top-down design

Ontogenetic and phylogenetic evolution has (necessarily) followed a **bottom-up** approach (grassroots) to "*design*" systems:

- **Instantiation of the basic units** (atoms, cells, organs, organisms, individuals) composing the system and let them **(self-)organize** to generate more complex/organized system-level behaviors, structures, and functions
- **Population + Interaction protocols** are more important than single modules
- System-level structural patterns and behaviors are ***emerging*** properties

From an engineering point of view we can also choose a **top-down** approach:

- Acquisition of comprehensive knowledge about the problem/system, make analysis, decomposition, definition of a possibly optimal strategy
- Amenable to formal analysis, "*predictable*" response

Many different paradigms of SI

Different ways of modeling **communications**, **connection topology**, and **spatial distribution** have given raise to different SI frameworks

- **Point-to-point communication** (one-to-one): two agents get in direct contact (e.g., antennation, trophallaxis, axons and dendrites in neurons)
- **Limited-range information broadcast (**one-to-many): the signal propagates to some limited extent throughout the environment and/or is available for a short time (e.g., fish' use of lateral line to detect water waves, visual detection)
- **Indirect communication:** two individuals interact indirectly when one of them modifies the environment and the other responds to the modified environment, maybe at a later time (e.g., stigmergy, pheromone communication in ants)
- **Physical mobility:** individuals move through the states of the environment, such as the connection topology changes over time (based on communication capability), different environment areas are accessed in parallel
- **Static positioning, state evolution:** connection topology and/or positioning in the environment do not change over time. Local information propagates in multi-hop modality. The internal state of an individual changes over time.

Some SI and SI-related frameworks

- *Stigmergy, Mobility* → **Ant Algorithms** and in particular to **Ant Colony Optimization (ACO)** [Dorigo & Di Caro, 1999], which is based on the shortest path finding abilities of ant colonies
- *Stigmergy* → **Cultural Algorithms** [Reynolds, 1994], population-based algorithms derived from processes of cultural evolution and exchange in societies
- *Limited broadcast, Mobility* → **Particle Swarm Optimization (PSO)** [Kennedy & Eberhart, 2001], related to fish schooling and bird flocking behaviors
- *Point-to-point* → **Hopfield neural networks** [Hopfield, 1982], derived from brain's structure and behavior
- *Point-to-point and neighbor limited broadcast* → **Cellular Automata** [Wolfram, 1984], **Gossip algorithms** [Demers et al., 1987] derived from infection models
- Different combinations of communication and mobility → **Swarm robotics**, Adaptive network routing, Consensus algorithms
- **Genetic algorithms**, **Artificial immune systems**, . . .