LECTURE 15:
SWARM INTELLIGENCE 1 /
PARTICLE SWARM OPTIMIZATION 1

INSTRUCTOR:
GIANNI A. DI CARO
**Limitations of the (classic) CA model**

- **Space is discretized** according to an $n \times m$ regular lattice
- The CA is equivalent to $n \times m$ 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 $\rightarrow$ Continuous

Spatially-fixed cell state $\sim$ FSM

$\rightarrow$ Agent-model:
- internal state
- mobile

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

$\rightarrow$ Relational neighborhood
- Agents form a network
- Logical topology
- Can be dynamic
REFERENCE TASK: GLOBAL FUNCTION OPTIMIZATION

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

Find the **global minimum** of the function
Optimization problems expressed in mathematical form:

\[
\min_x f(x)
\]
subject to \( x \in F \)

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

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

- **Mathematical programming** problem
Basic Properties

Given an optimization problem:

$$\min_x f(x)$$

subject to $x \in F$, $x \in X^n \subseteq \mathbb{R}^n$

- $\min_x f(x)$ is equivalent to $\max_x -f(x)$

- If $F = \emptyset$ the problem has no solution (unfeasible)

- If $F$ is an open set, only the inf (sup) is guaranteed but not min (max)

- The problem is unbounded if $f \to -\infty$
Any constrained optimization problem can be formulated as an unconstrained one by including constraint violations as penalty terms in the objective function.

Sublevel sets (isolines):

\[ \{ \mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = c \} \]
GLOBAL AND LOCAL OPTIMALITY

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

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

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

\[
\begin{align*}
\min \quad & Z_{ILP} = x_2 \\
\text{s.t.} \quad & 2x_1 + x_2 \geq 13 \\
& 5x_1 + 2x_2 \leq 30 \\
& -x_1 + x_2 \geq 5 \\
& x_1, x_2 \in \mathbb{Z}^+ 
\end{align*}
\]
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 \((x, BB(x))\) pairs...

**Black-box / Derivative-free optimization** vs. **White box optimization**
**How do we find minima / maxima?**

- **Using rates of change:** derivatives / gradients / Jacobians / Hessians / ...
- **Sampling / Searching in the** $X^n \subseteq \mathbb{R}^n$ **domain**, 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 constructing a solution**, 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)**


- **Multi-agent black-box optimization inspired by social and roosting behavior of flocking birds:**
  - Each agent (a *particle*) encodes a solution point $x$.
  - 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.
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)
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 rise 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, . . .