An introduction to Metaheuristics

Andrea Roli
aroli@deis.unibo.it

Università degli Studi “G. D’Annunzio” – Chieti
DEIS - Università degli Studi di Bologna

Outline

• Definition of Metaheuristics
• Combinatorial Optimization Problems
• Trajectory Methods
• Population-based Methods
• Hybrid approaches

Metaheuristics

• Approximate algorithms: they do not guarantee to find the optimal solution in bounded time.
• Applied to Combinatorial Optimization Problems and Constraint Satisfaction Problems
• Applied when:
  • Problems have large size
  • The goal is to find a (near-)optimal solution quickly

Metaheuristics

OBJECTIVE: Effectively and efficiently explore the search space

Techniques:

• Use of the search history
• Adaptivity
• General strategies to balance intensification and diversification.

► Sometimes, they are erroneously simply called “local search methods”.
**Two-faced Janus**

Intensification and Diversification are the driving forces of metaheuristic search.

**Intensification**: exploitation of the accumulated search experience (e.g., by concentrating the search in a confined, small search space area)

**Diversification**: exploration 'in the large' of the search space

---

**Other characteristics**

- Metaheuristics are strategies that guide the search process.
- Metaheuristic algorithms are usually non-deterministic.
- The basic concepts of metaheuristics permit an abstract level description.
- Metaheuristics are not problem-specific.
- Metaheuristics may make use of domain-specific knowledge as heuristic controlled by the upper level strategy.

---

**Metaheuristics**

Metaheuristics encompass and combine:

- Constructive methods (e.g., random, heuristic, adaptive, etc.)
- Local search-based methods (e.g., Tabu Search, Simulated Annealing, Iterated Local Search, etc.)
- Population-based methods (e.g., Evolutionary Algorithms, Ant Colony Optimization, Scatter Search, etc.)
A classification

We will classify metaheuristics in two basic classes:

- Trajectory methods
- Population-based methods

Other classifications are possible, based on different key concepts (e.g., the use of memory)

Trajectory methods

- The search process is characterized by a trajectory in the search space
- The search process can be seen as the evolution in (discrete) time of a discrete dynamical system

Examples: Tabu Search, Simulated Annealing, Iterated Local Search, ...

Population-based methods

- Deal in every iteration of the algorithm with a set – a population – of solutions
- The search process can be seen as the evolution in (discrete) time of a set of points in the search space

Examples: Evolutionary Algorithms, Ant Colony Optimization, Scatter Search, ...

Combinatorial Optimization Problems

A **Combinatorial Optimization Problem** $\mathcal{P} = (\mathcal{S}, f)$ can be defined by:

- variables $X = \{x_1, \ldots, x_n\}$;
- variable domains $D_1, \ldots, D_n$;
- constraints among variables;
- **Objective function** $f : D_1 \times \ldots \times D_n \to \mathbb{R}^+$;
- The set of all possible feasible assignments
  $\mathcal{S} = \{s = \{(x_1, v_1), \ldots, (x_n, v_n)\} \mid v_i \in D_i, s \text{ satisfies all the constraints}\}$
Combinatorial Optimization Problems

**Objective:** find a solution $s^* \in S$ with minimum objective function value, i.e., $f(s^*) \leq f(s) \ \forall s \in S$.

Many COPs are $\mathcal{NP}$-hard $\Rightarrow$ no polynomial time algorithm exists (assuming $\mathcal{P} \neq \mathcal{NP}$)

Examples: Traveling Salesman problem (TSP), Quadratic Assignment problem (QAP), Maximum Satisfiability Problem (MAXSAT), Timetabling and Scheduling problems.

---

**Preliminary definitions**

A **neighborhood structure** is a function $\mathcal{N} : S \rightarrow 2^S$ that assigns to every $s \in S$ a set of neighbors $\mathcal{N}(s) \subseteq S$. $\mathcal{N}(s)$ is called the neighborhood of $s$.

A **locally minimal solution** (or **local minimum**) with respect to a neighborhood structure $\mathcal{N}$ is a solution $\hat{s}$ such that $\forall s \in \mathcal{N}(\hat{s}) : f(\hat{s}) \leq f(s)$. We call $\hat{s}$ a strict locally minimal solution if $f(\hat{s}) < f(s) \ \forall s \in \mathcal{N}(\hat{s})$.

---

**Neighborhood: Examples**

For problems defined on binary variables, the neighborhood can be defined on the basis of the Hamming distance ($H_d$) between two assignments. E.g.,

$$\mathcal{N}(s_i) = \{ s_j \in \{0,1\}^n | H_d(s_i, s_j) = 1 \}$$

For example: $\mathcal{N}(000) = \{001, 010, 100\}$

---

**Neighborhood: Examples**

In TSP, the neighborhood can be defined by means of arc exchanges on Hamiltonian tours. E.g.,
Trajectory methods

- Iterative Improvement
- Simulated Annealing
- Tabu Search
- Variable Neighborhood Search
- Guided Local Search
- Iterated Local Search

Iterative Improvement

- Very basic local search
- Each move is only performed if the solution it produces is better than the current solution (also called \textit{hill-climbing}).
- The algorithm stops as soon as it finds a local minimum.
Iterative Improvement

The algorithm

\[ s \leftarrow \text{GenerateInitialSolution()} \]
\[ \text{repeat} \]
\[ s \leftarrow \text{BestOf}(s, N(s)) \]
\[ \text{until no improvement is possible} \]

A pictorial view

Escaping strategies...

Problem: Iterative Improvement stops at Local minima, which can be very “poor”.

⇒ Strategies are required to prevent the search from getting trapped in local minima and to escape from them.
Three basic ideas

1) **Accept up-hill moves**
   
i.e., the search moves toward a solution with a *worse* objective function value

   **Intuition:** climb the hills and go downward in another direction

---

Three basic ideas

2) **Change neighborhood structure during the search**

   **Intuition:** different neighborhoods generate different search space topologies

---

Three basic ideas

3) **Change the objective function so as to “fill-in” local minima**

   **Intuition:** modify the search space with the aim of making more “desirable” not yet explored areas

---

Simulated Annealing

Simulated Annealing exploits the first idea: *accept also up-hill moves*.

- Origins in statistical mechanics (Metropolis algorithm)
- It allows moves resulting in solutions of worse quality than the current solution
- The probability of doing such a move is decreased during the search.

Usually, \( p(\text{accept up-hill moves}^{'}) = \exp\left(-\frac{f(x') - f(x)}{T}\right) \).
A pictorial view

SA: the algorithm

\[ s \leftarrow \text{GenerateInitialSolution()} \]
\[ T \leftarrow T_0 \]
\[ \text{while termination conditions not met do} \]
\[ \quad s' \leftarrow \text{PickAtRandom}(\mathcal{N}(s)) \]
\[ \quad \text{if } f(s') < f(s) \text{ then} \]
\[ \quad \quad s \leftarrow s' \{ s' \text{ replaces } s \} \]
\[ \quad \text{else} \]
\[ \quad \quad \text{Accept } s' \text{ as new solution with probability } p(T, s', s) \]
\[ \quad \text{end if} \]
\[ \quad \text{Update}(T') \]
\[ \text{end while} \]

Cooling schedules

The temperature \( T \) can be varied in different ways:

- Logarithmic: \( T_{k+1} = \frac{T_k}{\log(k+k_0)} \).
  The algorithm is guaranteed to converge to the optimal solution with probability 1. Too slow for applications.

- Geometric: \( T_{k+1} = \alpha T_k \), where \( \alpha \in ]0, 1[ \).

- Non-monotonic: the temperature is decreased (intensifications is favored), then increased again (to increase diversification).

Applications

SA is usually not very effective when used as stand-alone metaheuristic.

References:

Tabu Search

Tabu Search exploits the second idea: change neighborhood structure.

- Explicitly exploits the search history to dynamically change the neighborhood to explore
- Tabu list: keeps track of recent visited solutions or moves and forbids them \( \Rightarrow \) escape from local minima and no cycling
- Many important concepts developed “around” the basic TS version (e.g., general exploration strategies)
Basic TS: the algorithm

\[ s \leftarrow \text{GenerateInitialSolution()} \]
\[ \text{TabuList} \leftarrow \emptyset \]
\[ \textbf{while} \text{ termination conditions not met} \textbf{do} \]
\[ s \leftarrow \text{ChooseBestOf}(s \cup \mathcal{N}(s) \setminus \text{TabuList}) \]
\[ \text{Update}(	ext{TabuList}) \]
\[ \textbf{end while} \]

Tabu Search

Storing a list of solutions is often inefficient, therefore \textit{moves} are stored instead.

BUT: storing moves we could cut good not yet visited solutions

\[ \downarrow \]

we use \textit{ASPIRATION CRITERIA} (e.g., accept a forbidden move toward a solution better than the current one)

TS: the algorithm

\[ s \leftarrow \text{GenerateInitialSolution()} \]
\[ \text{InitializeTabuLists}(T L_1, \ldots , T L_r) \]
\[ k \leftarrow 0 \]
\[ \textbf{while} \text{ termination conditions not met} \textbf{do} \]
\[ \text{AllowedSet}(s, k) \leftarrow \{z \in \mathcal{N}(s) \mid \text{no tabu condition is violated or at least one aspiration condition is satisfied}\} \]
\[ s \leftarrow \text{ChooseBestOf}(s \cup \text{AllowedSet}(s, k)) \]
\[ \text{UpdateTabuListsAndAspirationConditions()} \]
\[ k \leftarrow k + 1 \]
\[ \textbf{end while} \]

Applications

- Among the best performing metaheuristics (when applied with general strategies to balance intensification and diversification)
- Applied to many COPs
- TS approaches dominate the Job Shop Scheduling area

References:
Variable Neighborhood Search

VNS exploits the second idea: *change neighborhood structure*.

- VNS uses different neighborhood structures during search
- A neighborhood $\mathcal{N}_i$ is substituted by neighborhood $\mathcal{N}_j$ as soon as local search can not improve the current best solution.

Applications

- Graph based COPs (e.g., $p$-Median problem, the Steiner tree problem, $k$-Cardinality Tree problem)
- Some variants also very effective

References:

VNS: the algorithm

Select a set of neighborhood structures $\mathcal{N}_k$, $k = 1, \ldots, k_{max}$

$s \leftarrow \text{GenerateInitialSolution}()$

while termination conditions not met do

$k \leftarrow 1$

while $k < k_{max}$ do {Inner Loop}

$s' \leftarrow \text{PickAtRandom}(\mathcal{N}_k(s))$ {Shaking phase}

$s'' \leftarrow \text{LocalSearch}(s')$

if $f(s'') < f(s)$ then

$s \leftarrow s''; k \leftarrow 1$

else

$k \leftarrow k + 1$

end if

end while

end while

Guided Local Search

GLS exploits the third idea: *dynamically change the objective function*.

- Basic principle: help the search to move out gradually from local optima by changing the search landscape
- The objective function is dynamically changed with the aim of making the current local optimum “less desirable”
A pictorial view

A pictorial view

A pictorial view

A pictorial view
**Guided Local Search**

GLS penalizes solutions which contain some defined features (e.g., arcs in a tour, unsatisfied clauses, etc.)

If feature $i$ is present in solution $s$, then $I_i(s) = 1$, otherwise $I_i(s) = 0$

**Guided Local Search**

Each feature $i$ is associated a penalty $p_i$ which weights the importance of the features.

The objective function $f$ is modified so as to take into account the penalties.

$$f'(s) = f(s) + \lambda \sum_{i=1}^{m} p_i \cdot I_i(s)$$

$\lambda$ scales the contribution of the penalties wrt to the original objective function

**GLS: The Algorithm**

1. $s \leftarrow \text{GenerateInitialSolution}()$
2. **while** termination conditions not met **do**
   1. $s \leftarrow \text{LocalSearch}(s, f')$
   2. **for all** selected features $i$ **do**
      1. $p_i \leftarrow p_i + 1$
   **end for**
   3. Update($f', p$)
      {where $p$ is the penalty vector}
3. **end while**
GLS: Applications

Crucial: optimally tune parameters and penalty updating procedure

Successfully applied to the weighted MAXSAT, the VR problem, the TSP and the QAP

References:

Iterated Local Search

- ILS can be seen as a general trajectory method framework
- Three basic blocks:
  - Local Search
  - Perturbation
  - Acceptance criteria

ILS: basic scheme

1. Generate an initial solution
2. Apply local search (e.g., SA, TS, etc.)
3. Perturb (i.e., slightly change) the obtained solution
4. Apply again local search with the perturbed solution as starting solution
5. Decide whether to accept the new solution or not
6. Go to step 3

A pictorial view

- Objective function vs. Solution space
A pictorial view

ILS: the algorithm

\[
\begin{align*}
    s_0 & \leftarrow \text{GenerateInitialSolution()} \\
    s^* & \leftarrow \text{LocalSearch}(s_0) \\
    \text{while} & \text{ termination conditions not met do} \\
    & s' \leftarrow \text{Perturbation}(s^*, \text{history}) \\
    & s^{*'} \leftarrow \text{LocalSearch}(s') \\
    & s^* \leftarrow \text{ApplyAcceptanceCriterion}(s^*, s^{*'}, \text{history}) \\
    \text{end while}
\end{align*}
\]
Design principles

- 'Local search' can be any kind of trajectory method.
- The perturbation should be strong enough to move the starting point to another local minimum basin of attraction, but it should keep some parts of the current solution.
- The acceptance criteria can range from very simple (e.g., accept the new solution if better than the current one) to more complex (e.g., with probabilistic acceptance).

Applications

- TSP, QAP, Single Machine Total Weighted Tardiness (SMTWT) problem, Graph Coloring Problem

References:

Lessons learnt

- The effectiveness of a metaheuristic strongly depends on the dynamical interplay of intensification and diversification.
- General search strategies have to be applied to effectively explore the search space.
- The use of search history characterizes the nowadays most effective algorithms.
- Optimal parameter tuning is crucial and sometimes very difficult to achieve.

Population-based methods

- Evolutionary Algorithms
  - Evolutionary Programming
  - Evolution Strategies
  - Genetic Algorithms
- Ant Colony Optimization

But also (not covered by this introduction): Scatter Search, Population-Based Incremental Learning and Estimation of Distribution Algorithms.
Evolutionary Algorithms

- Inspired by Nature’s capability to evolve living beings well adapted to their environment.
- Computational models of evolutionary processes.

They include:
- Evolutionary Programming
- Evolution Strategies
- Genetic Algorithms

Basic principle: moving a population of solutions toward good regions of the search space.
Evolutionary Algorithms

Basic principle: moving a population of solutions toward good regions of the search space.

The Evolutionary Cycle

<table>
<thead>
<tr>
<th>Selection</th>
<th>Parents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombination</td>
<td></td>
</tr>
<tr>
<td>Mutation</td>
<td></td>
</tr>
<tr>
<td>Replacement</td>
<td></td>
</tr>
<tr>
<td>Offspring</td>
<td></td>
</tr>
</tbody>
</table>

EA: the algorithm

\[
P \leftarrow \text{GenerateInitialPopulation()}
\]
\[
\text{Evaluate}(P)
\]
\[
\textbf{while} \ \text{termination conditions not met} \ \textbf{do}
\]
\[
P' \leftarrow \text{Recombine}(P)
\]
\[
P'' \leftarrow \text{Mutate}(P')
\]
\[
\text{Evaluate}(P'')
\]
\[
P \leftarrow \text{Select}(P'' \cup P)
\]
\[
\textbf{end while}
\]

The Seven Features

1. Description of the individuals
2. Evolution process
3. Neighborhood structure
4. Information sources
5. Infeasibility
6. Intensification strategy
7. Diversification strategy
Description of the individuals

Solutions can be represented in many ways:

- bit-strings
- integer/real arrays
- tree structures or more complex data structures

▶ the representation is crucial for the success of the algorithm

Evolution process

The evolution process determines which individuals will enter the population at each iteration.

- Generational replacement: the offspring entirely replaces the old population
- Steady state: some new individuals are inserted in the old population
- The population size can be constant or varying

Neighborhood structure

The neighborhood function $\mathcal{N}_C : I \rightarrow 2^I$ defines, for every individual $i \in I$, the set of individuals $\mathcal{N}_C(i) \subseteq I$ which can be recombined with it.

- Unstructured population: every individual can be recombined with any other one (e.g., Simple Genetic Algorithm)
- Structured population: otherwise (e.g., Parallel Genetic Algorithm with “islands”)

Information sources

Several kinds of recombination are possible:

- two-parent crossover
- multi-parent crossover
- population statistics-based recombination operators
Infeasibility

The result of a recombination could be an individual violating some constraints.

Three possible ways of dealing with infeasibility:

- **Reject**: discard infeasible solutions
- **Penalize**: decrease the fitness of individuals violating constraints
- **Repair**: apply some operators to change the solution trying to obtain a feasible one

Intensification strategy

It is possible to apply operators or algorithms to improve the fitness of single individuals.

For example:

- Before the replacement, apply local search to every individual of the population (→ memetic algorithms).
- Apply mutation operators based on local improvements (e.g., some steps of local search).

Diversification strategy

To avoid premature convergence of the search, diversification techniques are introduced.

For example:

- Random mutation (most often adopted)
- Introduce into the population new individuals 'coming' from not yet explored areas of the search space

Applications

- Applied to nearly any COPs and optimization problems
- Particularly effective in robotics applications

References:

Ant Colony Optimization

Population-based metaheuristic inspired by the foraging behavior of ants, which enables them to find the shortest path between the nest and a food source.

- While walking ants deposit a substance called *pheromone* on the ground.
- When they decide about a direction to go, they choose with higher probability paths that are marked by stronger pheromone concentrations.
- This basic behavior is the basis for a cooperative interaction which leads to the emergence of shortest paths.

ACO construction graph

\[ G = (\mathcal{C}, \mathcal{L}) \]

- vertices are the solution components \( \mathcal{C} \)
- \( \mathcal{L} \) are the connections
- states are paths in \( G \).

Solutions are states, i.e., encoded as paths on \( G \)

Constraints are also provided in order to construct feasible solutions

Ant Colony Optimization

ACO algorithms are based on a parametrized probabilistic model – the *pheromone model* – that is used to model the chemical pheromone trails.

Artificial ants incrementally construct solutions by adding opportune defined solution components to a partial solution under consideration.

Artificial ants perform randomized walks on the construction graph: a completely connected graph \( G = (\mathcal{C}, \mathcal{L}) \).

Example

One possible TSP model for ACO:

- nodes of \( G \) (the components) are the cities to be visited;
- states are partial or complete paths in the graph;
- a solution is an Hamiltonian tour in the graph;
- constraints are used to avoid cycles (an ant can not visit a city more than once).
**Sources of information**

- Connections, components (or both) can have associated **pheromone** trail and **heuristic** value.

- Pheromone trail takes the place of natural pheromone and encodes a long-term memory about the whole ants’ search process.

- Heuristic represents a priori information about the problem or dynamic heuristic information (in the same way as static and dynamic heuristics are used in constructive algorithms).
A pictorial view
**ACO: the algorithm**

```plaintext
while termination conditions not met do
    ScheduleActivities
    AntBasedSolutionConstruction()
    PheromoneUpdate()
    DaemonActions() {optional}
end ScheduleActivities
end while
```

**Solution construction**

- Ants move by applying a *stochastic local decision policy* that makes use of the pheromone values and the heuristic values on components of the construction graph.
- While moving, the ant keeps in memory the partial solution it has built in terms of the path it was walking on the construction graph.
Pheromone Update

- When adding a component $c_j$ to the current partial solution, an ant can update the pheromone trail (*online step-by-step pheromone update*).
- Once an ant has built a solution, it can retrace the same path backward and update the pheromone trails of the used components according to the quality of the solution it has built (*online delayed pheromone update*).
- *Pheromone evaporation* is always applied → the pheromone trail intensity on the components decreases over time.

Daemon Actions

- Can be used to implement centralized actions which cannot be performed by single ants. E.g.,
  - local search procedure applied to the solutions built by the ants
  - collection of global information used to decide whether to deposit additional pheromone to bias the search process from an non-local perspective

Variants

- Ant Colony System
- $\mathcal{MAX}$-$\mathcal{MIN}$ Ant System – among the most effective ACO implementations
- *Hyper-Cube Framework*: generalizes pheromone-based construction mechanisms

Applications

- Routing in communication networks, Sequential Ordering Problem, Resource Constraint Project Scheduling
  - ACO is effective when combined with local search and/or tree-search strategies.

References:

## Hybrid metaheuristics

- Component exchange among metaheuristics
- Cooperative search
- Integrating metaheuristics and systematic methods

## Component exchange

- Integrate trajectory methods into population-based methods (e.g., memetic algorithms)
- Apply general I&D strategies of a metaheuristic into another one (e.g., TS restart strategy applied into ILS)

> The combination of population-based methods with trajectory methods produces hybrid algorithms which are often more efficient than single methods.

## Cooperative search

- *Loose* form of integration
- The search is performed by different algorithms (possibly running in parallel)
- The algorithms *exchange information* during the search process
- Crucial:
  - kind of information exchanged
  - implementation

## Metaheuristics and systematic methods

1. Metaheuristics are applied before systematic methods, providing a valuable input, or vice versa.

2. Metaheuristics use CP and/or tree search to efficiently explore the neighborhood.
Metaheuristics and systematic methods

3. A “tree search”-based algorithm applies a metaheuristic in order to improve a solution (i.e., a leaf of the tree) or a partial solution (i.e., an inner node). Metaheuristic concepts can also be used to obtain incomplete but efficient tree exploration strategies.


Overview references


Internet resources

- www.metaheuristics.net
- tew.ruca.ua.ac.be/eume/