So far, we have considered methods that systematically explore the full search space, possibly using principle pruning (A* etc.).

The current best such algorithms (IDA* / SMA*) can handle search spaces of up to $10^{100}$ states.

What if we have 10,000 or 100,000 variables / search spaces of up to $10^{30,000}$ states?

A completely different kind of method is called for:

*Local Search Methods* or

*Iterative Improvement Methods*

---

**Local Search Methods**

Applicable when we’re interested in the Goal State — not in how to get there.

E.g. N-Queens, VLSI layout, or map coloring.
Local Search Methods

Key idea (suprisingly simply):

a) Select (random) initial state
   (initial guess at solution)

b) Make local modification to try
   to improve current state.

c) repeat b) till goal state found
   or can’t improve from current state
   (or out of time).
Example

Map coloring:
   a) start with random coloring of map sections
   b) change the color of one of the sections to reduce conflicts
   c) repeat b)

Slide CS472 – Local Search 5

Slide CS472 – Local Search 6
function HILL-CLIMBING(problem) returns a solution state
inputs: problem, a problem
static: current, a node
        next, a node

current ← MAKE-NODE(INITIAL-STATE[problem])
loop do
    next ← a highest-valued successor of current
    if VALUE[next] < VALUE[current] then return current
    current ← next
end

Hill Climbing Pathologies

- Local maximum
- Plateau
- Ridge
Improvements to Basic Local Search

Issue: How to move more quickly to successively higher plateaus and avoid getting “stuck” / local minima.

Idea: Introduce uphill moves (“noise”) to escape from long plateaus (or true local minima).

Strategies:

- Multiple runs from randomly generated initial states
- Random-restart hill-climbing
- Tabu search
- Simulated Annealing

Slide CS472 – Local Search 9

- Genetic Algorithms

Slide CS472 – Local Search 10
Variations on Hill-Climbing

1. **random restarts**: simply restart at a new random state after a pre-defined number of local steps.

2. **tabu**: prevent returning quickly to same state.
   Implementation: Keep fixed length queue (“tabu list”): add most recent step to queue; drop “oldest” step.
   Never make step that’s currently on the tabu list.

---

**Slide CS472 – Local Search 11**

---

**Simulated Annealing**

**Idea:**
Use conventional hill-climbing techniques, but occasionally take a step in a direction other than that in which the rate of change is maximal.

As time passes, the probability that a down-hill step is taken is gradually reduced and the size of any down-hill step taken is decreased.

Kirkpatrick *et al.* 1982; Metropolis *et al.* 1953.

---

**Slide CS472 – Local Search 12**
SA Algorithm

\(current, next\): nodes/states

\(T\): “temperature” controlling probability of downward steps

\(schedule\): mapping from time to “temperature”

\(h\): heuristic evaluation function

---

\(current \leftarrow \text{initial state}\)

for \(t \leftarrow 1\) to inf do

\(T \leftarrow schedule[t]\)

if \(T = 0\) then return \(current\)

\(next \leftarrow \text{randomly selected successor of } current\)

\(\Delta E \leftarrow h(next) - h(current)\)

if \(\Delta E > 0\) then \(current \leftarrow next\)

else \(current \leftarrow next\) only with probability \(e^{\Delta E/T}\)
Genetic Algorithms

- Approach mimics *evolution*.
  (See Section 20.8 R&N.)
- Usually presented using rich new vocabulary:
  - fitness function, population, individuals, genes, crossover, mutations, etc.
- Still, can be viewed quite directly in terms of standard *local search*.

---

Features of evolution

- High degree of parallelism
- New individuals ("next state / neighboring states"): derived from "parents" ("crossover operation") genetic mutations
- Selection of next generation:
  Survival of the fittest.
Genetic Algorithms

Inspired by biological processes that produce genetic change in populations of individuals.

Genetic algorithms (GAs) are local search procedures that usually include three basic elements:

1. A Darwinian notion of fitness: the most fit individuals have the best chance of survival and reproduction.
2. Mating operators: individuals contribute their genetic material to their children.
3. Mutation: individuals are subject to random changes in their genetic material.

General Idea

- Maintain a population of individuals (states / strings / candidate solutions)
- Each individual is evaluated using a fitness function. The fitness scores force individuals to compete for the privilege of survival and reproduction.
- Generate a sequence of generations:
  - From the current generation, select pairs of individuals (based on fitness) to generate new individuals, using crossover.
• Introduce some noise through random \textbf{mutations}.

• Hope that average and maximum fitness (i.e. value to be optimized) increases over time.

---

\textbf{Genetic algorithms as search}

• Genetic algorithms are local heuristic search algorithms.

• Especially good for problems that have large and poorly understood search spaces.

• Genetic algorithms use a randomized parallel beam search to explore the state space.

• You must be able to define a good fitness function, and of course, a good state representation.
Binary string representations

- Individuals are usually represented using bit strings.
- Individuals represented can be arbitrarily complex.
- E.g. each component of the state description is allocated a specific portion of the string, which encodes the values that are acceptable.
- Bit string representation allows crossover operation to change multiple values in the state description. Crossover and mutation can also produce previously unseen values.

Example

World championship chocolate chip cookie recipe.

<table>
<thead>
<tr>
<th>flour</th>
<th>sugar</th>
<th>salt</th>
<th>chips</th>
<th>vanilla</th>
<th>fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4.5</td>
<td>3</td>
<td>1</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2.2</td>
<td>2.5</td>
<td>2.5</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4.1</td>
<td>2.5</td>
<td>1.5</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>1.5</td>
<td>2</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1.5</td>
<td>1.5</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

generation 1

Slide CS472 – Local Search 21

Slide CS472 – Local Search 22
\[ \text{GA}(\text{Fitness}, \text{Fitness\_threshold}, p, r, m) \]

- \( P \leftarrow \) randomly generate \( p \) individuals
- For each \( i \) in \( P \), compute \( \text{Fitness}(i) \)
- While \([\max_i \text{Fitness}(i)] < \text{Fitness\_threshold}\)
  1. Probabilistically select \((1 - r)p\) members of \( P \) to add to \( P_s \).
  2. Probabilistically choose \( \frac{r p}{2} \) pairs of individuals from \( P \). For each pair, \( \langle i_1, i_2 \rangle \), apply \text{crossover} and add the offspring to \( P_s \).
  3. \text{Mutate} \( m \cdot p \) random members of \( P_s \).
  4. \( P \leftarrow P_s \)
  5. For each \( i \) in \( P \), compute \( \text{Fitness}(i) \)

\[ \text{Slide CS472 – Local Search 23} \]

- Return the individual in \( P \) with the highest fitness.

\[ \text{Slide CS472 – Local Search 24} \]
Selecting Most Fit Individuals

Individuals are chosen probabilistically for survival and crossover based on **fitness proportionate selection**:

\[
\Pr(i) = \frac{\text{Fitness}(i)}{\sum_{j=1}^{p} \text{Fitness}(i_j)}
\]

Slide CS472 – Local Search 25

Other selection methods include:

- **Tournament Selection**: 2 individuals selected at random. With probability \( p \), the most fit is selected. With probability \( 1 - p \), the less fit is selected.

- **Rank Selection**: The individuals are sorted by fitness and the probability of selecting an individual is proportional to its rank in the list.

Slide CS472 – Local Search 26
## Crossover Operators

Single-point crossover:

- **Parent A:** 1 0 0 1 0 1 1 1 0 1
- **Parent B:** 0 1 0 1 1 1 0 1 1 0

<table>
<thead>
<tr>
<th>Child AB</th>
<th>Parent A</th>
<th>Parent B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 1 0 1 0 1 1 0</td>
<td>1 0 0 1 0 1 1 1 0 1</td>
<td>0 1 0 1 1 1 1 0 1</td>
</tr>
</tbody>
</table>

Two-point crossover:

- **Parent A:** 1 0 0 1 0 1 1 1 0 1
- **Parent B:** 0 1 0 1 1 1 0 1 1 0

<table>
<thead>
<tr>
<th>Child AB</th>
<th>Parent A</th>
<th>Parent B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 1 1 1 0 1 0 1</td>
<td>1 0 0 1 0 1 1 1 1 0</td>
<td>0 1 0 1 0 1 1 1 1 0</td>
</tr>
</tbody>
</table>
Uniform Crossover

Uniform crossover:

\[ \text{Parent } A: \quad 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1 \]
\[ \text{Parent } B: \quad 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1 \ 0 \]

\[ \text{Child } AB: \quad 1 \ 1 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \]
\[ \text{Child } BA: \quad 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 1 \ 0 \]

Mutation

Mutation: randomly toggle one bit

\[ \text{Individual } A: \quad 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1 \]
\[ \text{Individual } A': \quad 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1 \]
Mutation

- The **mutation** operator introduces random variations, allowing solutions to jump to different parts of the search space.
- What happens if the mutation rate is too low?
- What happens if the mutation rate is too high?
- A common strategy is to use a high mutation rate when search begins but to decrease the mutation rate as the search progresses.

Slide CS472 – Local Search 31

Deriving illegal structures

Consider the traveling salesman problem, where an individual represents a potential solution. The standard crossover operator can produce illegal children:

<table>
<thead>
<tr>
<th>Parent A:</th>
<th>ITH</th>
<th>Pitt</th>
<th>Chicago</th>
<th>Denver</th>
<th>Boise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent B:</td>
<td>Boise</td>
<td>Chicago</td>
<td>ITH</td>
<td>Phila</td>
<td>Pitt</td>
</tr>
<tr>
<td>Child AB:</td>
<td>ITH</td>
<td>Pitt</td>
<td>Chicago</td>
<td>Phila</td>
<td>Pitt</td>
</tr>
<tr>
<td>Child BA:</td>
<td>Boise</td>
<td>Chicago</td>
<td>ITH</td>
<td>Denver</td>
<td>Boise</td>
</tr>
</tbody>
</table>

Slide CS472 – Local Search 32
Two solutions:

1. define special genetic operators that only produce syntactically and semantically legal solutions.
2. ensure that the fitness function returns extremely low fitness values to illegal solutions.

Applications: Parameter Optimization

- Parameter optimization problems are well-suited for GAs. Each individual represents a set of parameter values and the GA tries to find the set of parameter values that achieves the best performance.
- The crossover operator creates new combinations of parameter values and, using a binary representation, both the crossover and mutation operators can produce new values.
- Many learning systems can be recast as parameter optimization problems. For example, most neural networks use a fixed architecture so learning consists
entirely of adjusting weights and thresholds.

Crossover with Variable-Length Bitstrings

Start with

\[
\begin{array}{ccc}
  a_1 & a_2 & c \\
  a_1 & a_2 & c \\
  i_1 : & 10 & 011 & 11 & 10 & 0 \\
  i_2 : & 01 & 110 & 10 & 01 & 0 \\
\end{array}
\]

1. choose crossover points for \( i_1 \), e.g., after bits 1, 8

2. now restrict points in \( i_2 \) to those that produce bitstrings with well-defined semantics, e.g., \( \langle 1, 3 \rangle, \langle 1, 8 \rangle, \langle 6, 8 \rangle \).
if we choose \( \langle 1, 3 \rangle \), result is

\[
\begin{array}{cccc}
a_1 & a_2 & c \\
\hline
i_3 & 11 & 10 & 0 \\
a_1 & a_2 & c & a_1 & a_2 & c & a_1 & a_2 & c \\
i_4 & 00 & 01 & 1 & 11 & 11 & 0 & 10 & 01 & 0
\end{array}
\]

Genetic Programming

In **Genetic Programming**, programs are evolved instead of bit strings. Programs are often represented by trees. For example:

\[
\sin(x) + \sqrt{x^2 + y}
\]
Remarks

- In practice, several 100 to 1000’s of strings. Value of crossover difficult to determine (so far).
- **Crowding** can occur when an individual that is much more fit than others reproduces like crazy, which reduces diversity in the population.
- In general, GA’s are highly sensitive to the representation.
- Given enough compute time, it’s the best search algorithm in certain domains.
Perspective
Jury still out on Genetic Algorithms in general, but nature suggests it can be a highly powerful mechanism for evolving highly complex systems (e.g., humans).
Formal properties far from understood.

Local Search — Summary
Surprisingly efficient search method.
Wide range of applications.
   any type of optimization / search task
Formal properties elusive.
Intuitive explanation:
   Search spaces (e.g., 10^{1000}) are often too too large for systematic search anyway ...
Area will most likely continue to thrive.
Often best available with lack of global info.