Introduction to Evolutionary Multiobjective Optimization (EMO)

Tsung-Che Chiang
Department of Computer Science and Information Engineering
National Taiwan Normal University

2009.11.18

Outline

- Multiobjective optimization (MO)
- Evolutionary algorithms (EA)
- Three styles of MOEAs
- Research directions
- Summary
Multiobjective Optimization

- **Significance**
  - One of the four fastest growing areas in the field of Computational Intelligence in WCCI06
  - MOEA-dedicated conferences and special sessions
    - International Conference on Evolutionary Multi-criterion Optimization (EMO)
    - Special sessions in CEC & GECCO
  - Special issues for EMO in flagship journals

---

Multiobjective Optimization

- Multiobjective optimization problem (MOOP)

\[
\begin{align*}
\text{Minimize/Maximize } & f_m(x), & m = 1, 2, \ldots, M; \\
\text{subject to } & g_j(x) \geq 0, & j = 1, 2, \ldots, J; \\
& h_k(x) = 0, & k = 1, 2, \ldots, K; \\
& x_i^{(L)} \leq x_i \leq x_i^{(U)}, & i = 1, 2, \ldots, n.
\end{align*}
\]
Multiobjective Optimization

- An example:
  Finding a way to move from Taipei to Kaohsiung

```
又要馬兒跑，又要馬兒不吃草
魚與熊掌難以兼得
```

Multiobjective Optimization

- The difficulty of MO: the concerned objectives are usually conflicting.

```
又要馬兒跑，又要馬兒不吃草
魚與熊掌難以兼得
```
Multiobjective Optimization

- **Methods** to solve MOOP
  - A priori
  - A posteriori
  - Interactive

Multiobjective Optimization

- A priori methods vs. a posteriori methods

[Diagram showing a priori and a posteriori methods with aggregation function, optimization order, single solution, and Pareto optimal set.]
Multiobjective Optimization

- A priori method - Aggregation

\[
\text{minimize } \sum_{i=1}^{k} w_i f_i(x)
\]

\[w = (0, 1)\]

\[w = (1, 0)\]

Multiobjective Optimization

- A priori method - Lexicographical ordering

\[f_2 \rightarrow f_1\]

\[f_1 \rightarrow f_2\]

\[f_2 \rightarrow f_1\]
Multiobjective Optimization

- A posteriori method

\[ x_i \text{ dominates } x_j \iff \]

1. \( x_i \) is not worse than \( x_j \) for all objectives, and
2. \( x_i \) is better than \( x_j \) for at least one objective.

![Graph showing Pareto front and dominance]

Assume minimization of both objectives

Multiobjective Optimization

- Disadvantages of a priori methods
  - It is difficult to give the preference information in advance.
  - The decision maker needs to adjust the preference to obtain alternative solutions. It is hard and laboring.
  - The condition becomes harder when there are multiple decision makers.
Multiobjective Optimization

- Goal of a posteriori methods

- Proximity

- Distribution

Performance metrics (single focus)

- Number of fronts
- Generational distance (GD)
- Spacing (SP)

\[ GD(P^*, P) = \sum_{x \in F} d(x, P^*) \]

\[ SP = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (d - d_i)^2} \]

\[ d_i = \min_j \left( \sum_{m=1}^{n} |f_m^i - f_m^j| \right) \]
Multiobjective Optimization

- **Performance metrics** (multi-focus)
  - Inverted Generational Distance (IGD)
  - Hypervolume
  - Epsilon indicator

\[
IGD(P^*, P) = \frac{\sum_{v \in P} d(v, P)}{|P^*|}
\]

\[
I_e(P, P^*) = \max_{u \in P} \min_{v \in P^*} \frac{f(u)}{f(v)}
\]

- **Popular benchmark problems**
  - Continuous functions
  - Knapsack
  - Permutation flow shop scheduling
Evolutionary Algorithms

- EAs are algorithms imitating the natural evolutionary process.

- EAs are
  - iterative (not constructive)
  - approximate (not always optimal)
  - nondeterministic
  - not problem-specific

**Flow of EA**

```
```

next generation

Stop? Y N

Final Population
Evolutionary Algorithms

- EA for optimization (an clustering example)

1. Initial Population
2. Evaluation
3. Mating selection
4. Reproduction
5. Evaluation
6. Environmental selection

Stop?

next generation

Final Population

"Introduction to EMO," talk@ee.ntu, 2009.11.18
Three Types of MOEAs

In the following, three types of MOEAs will be introduced:
- Pareto dominance-based
- Performance metric-based
- Aggregation-based

Pareto dominance-based MOEAs

NSGA-II (Deb et al., 2002)
- Pareto ranking

Crowding distance
- \( \text{dist}(\star) > \text{dist}(\bullet) \) (\( \bullet \) is better)
Pareto dominance-based MOEAs

- **NSGA-II** (Deb et al., 2002)

  
  Non-dominated sorting

  \[ P_t \]

  \[ Q_t \]

  \[ R_t \]

  Crowding distance sorting

  \[ F_1 \]

  \[ F_2 \]

  \[ P_{t+1} \]

  Rejected


Pareto dominance-based MOEAs

- **SPEA2** (Zitzler et al., 2001)

  - Raw fitness assignment \( R(i) \)
Pareto dominance-based MOEAs

- **SPEA2** (Zitzler et al., 2001)
  - Density estimation
    \[ d(i) = \text{the } k^{\text{th}} \text{ smallest distance from } i \text{ to the others} \]
    \[ D(i) = 1/(d(i)+2) \]
  - Fitness
    \[ F(i) = R(i) + D(i) \]
    (The smaller the better)

Performance metric-based MOEAs

- **Key Idea**
  - Use the concerned performance metric directly in fitness assignment.
  - The fitness of a solution depends on its contribution to the performance metric.
Performance metric-based MOEAs

- **IBEA** (Zitzler and Kunzli, 2004)
  - Contribution estimation (of $B$)

  $$I_{HD}(A, B) = \begin{cases} I_H(B) - I_H(A) & \text{if } \forall x^2 \in B \exists x^1 \in A : x^1 > x^2 \\ I_H(A + B) - I_H(A) & \text{else} \end{cases}$$

  ![Diagram](image)

"Introduction to EMO," talk@ee.ntu, 2009.11.18

---

Performance metric-based MOEAs

- **IBEA** (Zitzler and Kunzli, 2004)
  - Fitness assignment (the larger the better)

  $$F(x^1) = \sum_{x^2 \in P \setminus \{x^1\}} -e^{-I(H(x^2), x^1)^2/c}$$

  ![Diagram](image)

"Introduction to EMO," talk@ee.ntu, 2009.11.18
Performance metric-based MOEAs

- **IBEA** (Zitzler and Kunzli, 2004)
  - After producing offspring by mating selection, crossover, and mutation, environmental selection iterates the following three steps until the population size is reached:
    1. Choose an individual $x^*$ with the smallest fitness.
    2. Remove $x^*$ from the population.
    3. Update the fitness values of the remaining individuals.

- **SMS-EMOA** (Beume et al., 2007)
  - Main algorithm

```
Algorithm 1. SMS-EMOA
1: $P_0 \leftarrow \text{init}()$
2: $t \leftarrow 0$
3: repeat
4: $q_{t+1} \leftarrow \text{generate}(P_t)$
5: $P_{t+1} \leftarrow \text{Reduce}(P_t \cup \{q_{t+1}\})$
6: $t \leftarrow t + 1$
7: until termination condition fulfilled
```
Performance metric-based MOEAs

- **SMS-EMOA** (Beume et al., 2007)
  - Key function

  Algorithm 3. Reduced \( Q \)

  1. \( \{ P_1, \ldots, P_e \} \leftarrow \text{non-dominated-sort}(Q) \)
  2. if \( e > 1 \) then
  3. \( r \leftarrow \arg \max_{r \in P_1} [d(s, Q)] \)
  4. else
  5. \( r \leftarrow \text{argmin}_{r \in P_1} [\delta_{P}(s, P_1)] \)
  6. end if
  7. return \( (Q \setminus r) \)

  - Number of dominating solutions
  \[ d(s, P(t)) := |\{ y \in P(t) | y < s \}| \]

  - Contribution to the hypervolume
  \[ \Delta_{P}(s, P_0) := \mathcal{H}(P_0) - \mathcal{H}(P_0 \setminus \{ s \}). \]

Performance metric-based MOEAs

- **SMS-EMOA** (Beume et al., 2007)
  - Rationale

  y9 is more interesting than y8.

  \[ \Delta_{P}(s, P_0) \] prefers y8

  \[ d(s, P(t)) \] prefers y9
Performance metric-based MOEAs

- **SMS-EMOA** (Beume et al., 2007)
  - Characteristic

![Graph showing performance metric-based MOEAs](image)

Aggregation-based MOEAs

- **I-MOGLS** (Ishibuchi et al., 1998, 2003)
  - **GLS** = Genetic algorithm + Local Search
  - Fitness assignment: linear weighted sum $\sum w_i f_i(x)$
  - Mating selection
    - random weight vector
    - tournament

![Graph showing aggregation-based MOEAs](image)
Aggregation-based MOEAs

- **I-MOGLS** (Ishibuchi et al., 1998, 2003)
  - Issues brought by LS
    - Who does LS: random weight vector + tournament
    - Which direction: linear weighted sum
    - How to move: first improvement
    - When to stop: no better solutions in the neighborhood
    - Who survives: every terminating solution

- **TPSPGA** (Chang et al., 2005)
  - First phase
Aggregation-based MOEAs

- **TPSPGA** (Chang et al., 2005)
  - Second phase

```
Sub-population
Population
```

- **MOEA/D** (Zhang and Li, 2007, 2009)
  - “D” for problem Decomposition
  - Weight vector-based aggregation function
    - Linear weighted sum
    - Tchebycheff
    - Boundary intersection
  - Every weight vector defines a subproblem, and every subproblem keeps the best solution.
    - a large number of weight vector
    - a very small sub-population (In fact, the size is 1.)
Aggregation-based MOEAs

- MOEA/D (Zhang and Li, 2007, 2009)
  - Neighborhood
    - Mating restriction
    - Diversity control

Research Directions

- Algorithm design for general MOEA components:
  - Fitness assignment mechanism
    - Aggregation? Dominance?
    - Proximity vs. Diversity
  - Selection
    - Mating restriction & subpopulation
    - Archiving
  - Hybridization
  - Parallelization
Research Directions

- Algorithm design for specific problem natures:
  - Constrained
  - Dynamic
  - Uncertain
  - Expensive evaluation
  - Many objectives (say more than 10)

Research Directions

- Generation of benchmark problems
  - ZDT
  - DTLZ
  - Li & Zhang

"Introduction to EMO," talk@ee.ntu, 2009.11.18
Research Directions

- Generation of benchmark problems
  - ZDT
  - DTLZ
  - Li & Zhang

Research Directions

- Development of performance metrics
  - Zitzler et al. (2003)

<table>
<thead>
<tr>
<th>ref</th>
<th>name / reference</th>
<th>( I_x ) option / section &amp; B</th>
<th>( I_x ) additive option / section &amp; B</th>
<th>( I_x ) coverage / [7]</th>
<th>( I_y ) hypervolume indicator / [21]</th>
<th>( I_y ) utility function indicator / [8]</th>
<th>( I_y ) utility function indicator / [18]</th>
<th>( I_y ) area of interaction / [16]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I_x )</td>
<td>( I_x(A, B) ) ( &lt; 0 )</td>
<td>( I_x(A, B) ) ( \leq 1 )</td>
<td>( I_x(A, B) ) ( &gt; 1 )</td>
<td>( I_x(A, B) = 1 )</td>
<td>( I_x(A, B) = 0 )</td>
<td>( I_x(A, B) ) ( &lt; 0 )</td>
<td>( I_x(A, B) ) ( \leq 0 )</td>
<td>( I_x(A, B) ) ( &gt; 0 )</td>
</tr>
<tr>
<td>( I_x )</td>
<td>( I_x(A, B) ) ( \leq 0 )</td>
<td>( I_x(A, B) ) ( &lt; 1 )</td>
<td>( I_x(A, B) ) ( \geq 1 )</td>
<td>( I_x(A, B) ) ( = 1 )</td>
<td>( I_x(A, B) ) ( = 0 )</td>
<td>( I_x(A, B) ) ( &gt; 0 )</td>
<td>( I_x(A, B) ) ( \geq 0 )</td>
<td>( I_x(A, B) ) ( &lt; 0 )</td>
</tr>
<tr>
<td>( I_y )</td>
<td>( I_y(A, B) ) ( \leq 1 )</td>
<td>( I_y(A, B) ) ( &gt; 1 )</td>
<td>( I_y(A, B) ) ( = 1 )</td>
<td>( I_y(A, B) ) ( = 0 )</td>
<td>( I_y(A, B) ) ( &lt; 0 )</td>
<td>( I_y(A, B) ) ( \leq 0 )</td>
<td>( I_y(A, B) ) ( &gt; 0 )</td>
<td>( I_y(A, B) ) ( \geq 0 )</td>
</tr>
<tr>
<td>( I_y )</td>
<td>( I_y(A, B) ) ( &gt; 0 )</td>
<td>( I_y(A, B) ) ( \leq 0 )</td>
<td>( I_y(A, B) ) ( &lt; 1 )</td>
<td>( I_y(A, B) ) ( \geq 1 )</td>
<td>( I_y(A, B) ) ( = 0 )</td>
<td>( I_y(A, B) ) ( &gt; 0 )</td>
<td>( I_y(A, B) ) ( \geq 0 )</td>
<td>( I_y(A, B) ) ( &lt; 0 )</td>
</tr>
</tbody>
</table>
Research Directions

- Development of performance metrics
  - Zitzler et al. (2003)

<table>
<thead>
<tr>
<th>relation</th>
<th>objective vectors</th>
<th>approximation sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>strictly dominates</td>
<td>$z^1 &gt; z^2$</td>
<td>$z^1$ is better than $z^2$ in all objectives</td>
</tr>
<tr>
<td>dominates</td>
<td>$z^1 \geq z^2$</td>
<td>$z^1$ is not worse than $z^2$ in all objectives and better in at least one objective</td>
</tr>
<tr>
<td>better</td>
<td>$z^1 \geq z^2$</td>
<td>$z^1$ is not worse than $z^2$ in all objectives</td>
</tr>
<tr>
<td>weakly dominates</td>
<td>$z^1 \geq z^2$</td>
<td>$z^1$ is not worse than $z^2$ in all objectives</td>
</tr>
<tr>
<td>incomparable</td>
<td>$z^1 \parallel z^2$</td>
<td>neither $z^1$ weakly dominates $z^2$ nor $z^2$ weakly dominates $z^1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$A \parallel B$ every $z^1 \in B$ is weakly dominated by at least one $z^1 \in A$</td>
</tr>
</tbody>
</table>
Research Directions

- **Applications**
  - Multiobjectivization
    - Constrained optimization
    - Rule mining
    - Clustering
  - Innovization
    - Knowledge discovery of the Pareto optimal solutions (what makes a solution optimal?)

- **Decision making**
  - Generic consideration
    - Knee points
    - Pareto-optimal points having multiplicity
    - Pareto-optimal solutions which do not lie close to variable boundaries
  - Subjective consideration
    - Preference information
    - Interaction
  - Visualization
Research Directions

- **Software development**
  - Optimization methods
  - GUI
    - Visualization, post-processing, statistical charts
    - Decision support
  - Meta-modeling and model validation
  - Framework & libraries
    - ParadisEO (http://paradiseo.gforge.inria.fr/)
    - PISA (http://www.tik.ethz.ch/~sop/pisa/)
  - Parallel implementation

```
template<class EOT> 
class eoPop: public std::vector<EOT>, public eoObject, public eoPersistent 
{
  public:
    using std::vector<EOT>::size;
    using std::vector<EOT>::resize;
    using std::vector<EOT>::operator[];
    using std::vector<EOT>::begin;
    using std::vector<EOT>::end;
    
typedef typename EOT::Fitness Fitness;
```

"Introduction to EMO," talk@ee.ntu, 2009.11.18
Research Directions

- Software development
  - PISA library

- Research Directions
  - Software development
    - Parallelization
      - Algorithms
      - Components
        - evaluations
        - local search
      - Evaluation
        - multiple runs
        - objectives
        - solvers

*Introduction to EMO,* talk@ee.ntu, 2009.11.18
Summary

- In this talk, we
  - introduce the multiobjective optimization problem and the evolutionary algorithm
  - describe three major types of MOEAs
  - indicate the potential research directions
- This field is still growing rapidly, and there are many research opportunities.