Parameter Selection

Winter 2024

Control parameters

- Control parameters (policy parameters; tuning parameters) are configurable components of an algorithm
- They can be changed (fine-tuned) to seek better performance
- Also known as Algorithm Configuration
- Population size and number of iterations are not considered here

Specific parameters

- GA: crossover rate, mutation rate
- PSO: inertia weight, acceleration constants, velocity limit
- SA: cooling schedule, N_{iter}
- TS: tabu tenure, aspiration criterion, frequency penalty
- ACO: weighting parameters, Q in heuristic

Parameter setting

- Parameters may be expressed as continuous value within a specified range
  - Crossover rate between 0.9 and 0.7
  - Acceleration constants between 1 and 4
- Or as discrete levels
  - Mutation rate is picked from [0.01   0.1   0.2]
  - Inertia weight is picked from [1.5~0.8   1.2~0.6   1.0~0.4]

Criteria

- Comparing different settings of parameters based on the criteria we discussed previously
  - Solution quality
  - Computation time
  - Consistency

Analysis

- Statistics would be the means to compare different setting parameters
- Please refer to Performance Evaluation
- Parametric tests
  - Large sample size
  - Small sample size
- Non-parametric tests
  - Paired data   Wilcoxon signed rank
  - Unpaired data   Mann-Whitney

Classification

- Off-line tuning
  - Values of parameters are determined preliminarily and fixed during the progress of metaheuristics
- On-line tuning
  - Parameters are controlled and updated dynamically or adaptively during execution

Parameter tuning methods

- Trial and error
- Design of Experiment
- Taguchi method
- Hybrid with other metaheuristics
- Learn from moves
- Consider parameters as extra decision variables
- Racing

Trial and Error

- Most naive way to select parameters
- Adopt the setting suggested in the literature
- Sometimes based on experience or by prior knowledge on the problem
  - Using smaller acceleration constants for smaller search space
  - Larger mutation rate when the problem has multiple local optima
- Disadvantage
  - Labor-intensive; informal and less convincing

Design of Experiment: Factorial experiment

- Full factorial experiment (grid search) is an experiment whose design consists of two or more factors, each with discrete levels
- Experiment may take on all possible combinations of the levels across all such factors; guarantee best result
- Disadvantage:
  - m^n experiments are needed for n parameters, each with m levels
  - Very laborious when the numbers of parameters and associated levels are large

Design of Experiment: Fractional factorial experiment

- To reduce the number of experiments to a practical level, only a small set from all the possibilities is selected
- Use Taguchi orthogonal arrays to lay out the experiment
  - http://www.york.ac.uk/depts/maths/tables/orthogonal.htm

Example for Taguchi method (1)

1. For an maximization problem, there are 4 control parameters, and each one has 3 levels:
   low, medium, high
2. Choose the appropriate orthogonal array L9
   - Run 9 experiments
   - The first experiment picks level 1 for all the parameters
   - The second experiment picks level 1 for the first parameter and level 2 for the others
   - Continue for other experiments

Experiment	P1	P2	P3	P4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Example for Taguchi method (2)

3. Collect experiment results
4. For each parameter, compare the average cross the related experiments
5. Choose the best level for the parameter, e.g., Level 2 for Parameter 1

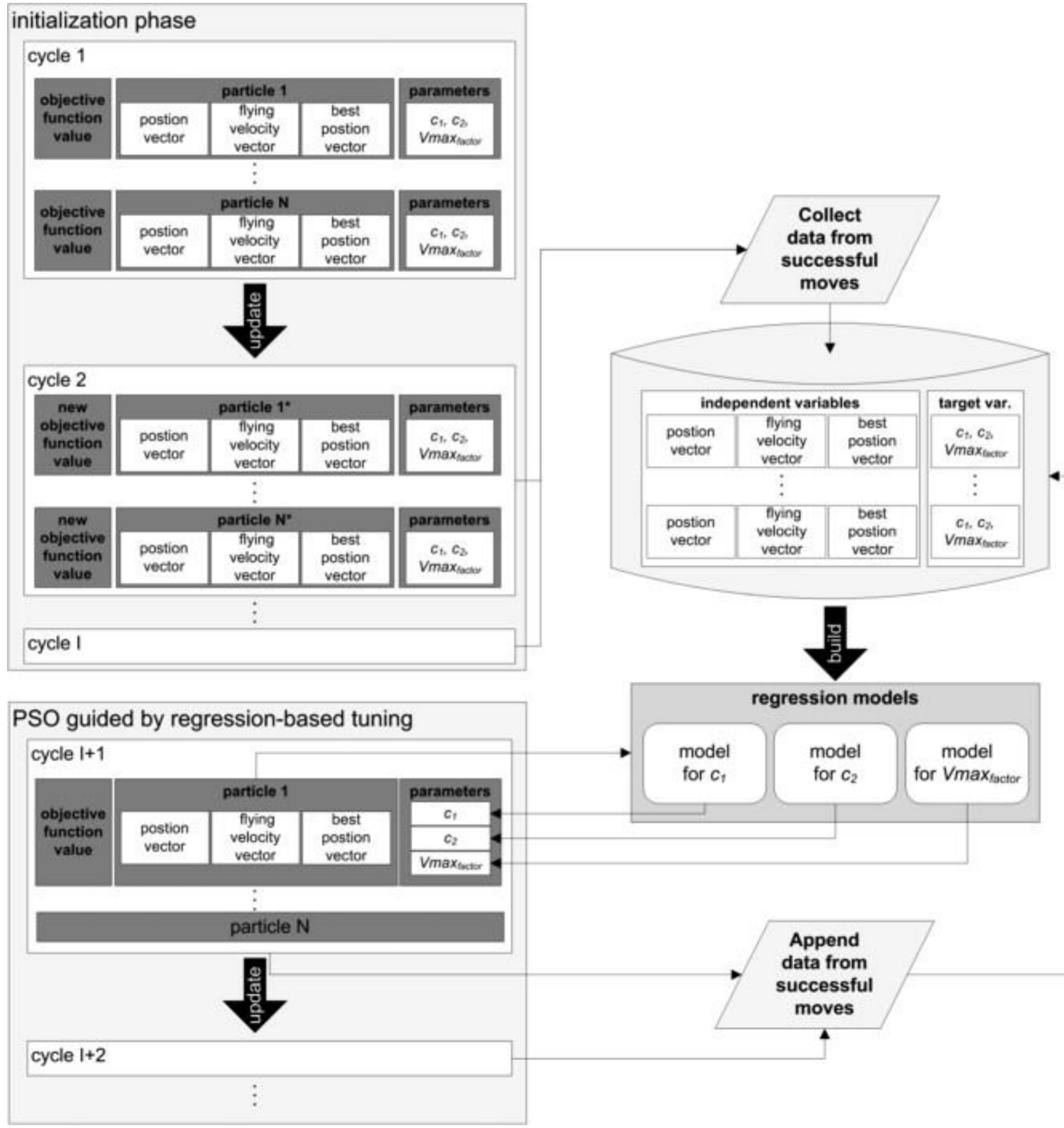
Experiment	P1	P2	P3	P4	Results	
1	1	1	1	1	100	
2	1	2	2	2	60	(100+60+80)/3=80
3	1	3	3	3	80	
4	2	1	2	3	70	
5	2	2	3	1	50	**(70+50+150)/3=90**
6	2	3	1	2	150	
7	3	1	3	2	90	
8	3	2	1	3	80	(90+80+40)/3=70
9	3	3	2	1	40	

Hybrid with other metaheuristics

- Nested loop: Use one metaheuristic to find the optimal parameters for another metaheuristic
- PSO is good at continuous parameter setting whereas GA is good at discrete one
- Disadvantage:
  - Nested loop may often be time-consuming
  - Methodology loop: A is controlled by B, which is controlled by C…

Learn from moves

- Use data mining or regression to find suitable parameters
- Goal: find the set of parameters leading to good performance



Lessmann, Caserta, Arango (2011). "Tuning metaheuristics: A data mining based approach for particle swarm optimization." Expert Systems with Applications, 38(10), 12826-12838.

Optimize parameters as decision variables

- Include parameters as a part of the optimization target; optimize them along with original decision variables
- Disadvantage:
  - Increase the dimension of search space; often influence the convergence speed
  - Parameters have to be time-invariant: do not change over time

Racing

- Initially perform only a few tests for each vector of parameters
- Separate the ones that are clearly good
- Iteratively increase the number of tests for those vectors only that are not significantly worse or better than the good ones
- Iterated Racing: http://iridia.ulb.ac.be/IridiaTrSeries/link/IridiaTr2011-004.pdf

Application and library

- ParamILS (Linux)

http://www.cs.ubc.ca/labs/beta/Projects/ParamILS/

- Sequential Model-based Algorithm Configuration (Python)

https://www.automl.org/automated-algorithm-design/algorithm-configuration/smac/

- Sequential Parameter Optimization (R)

https://martinzaefferer.de/?page_id=92

- AClib (Python)

http://www.aclib.net/

Remarks

- Control parameters may often influence the result
- There is a tradeoff between computation time and optimization performance
  - Full factorial experiments guarantee best performance but require lengthy computations
  - Other methods take much less time but the performance may be questionable