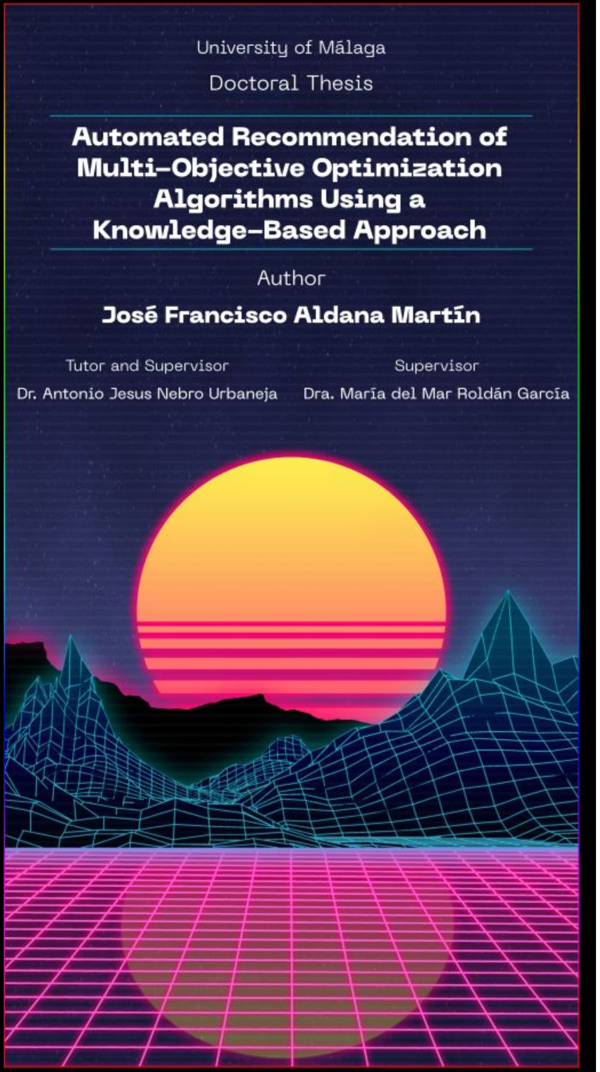


Automated Recommendation of Multi-Objective Optimization Algorithms Using a Knowledge-Based Approach

Doctoral thesis

José Francisco Aldana Martín

Tutor and Supervisor:
Antonio Jesús Nebro Urbaneja
Supervisor:
María del Mar Roldán García

September, 2024

UNIVERSIDAD DE MÁLAGA

khaos RESEARCH

SOFTWARE
UNIVERSIDAD DE MÁLAGA

Outline

1. Introduction
    a. Context and motivation
    b. Objectives
    c. Theoretical foundations
2. PhD contributions
    a. A Semantic Approach to Standardizing Multi-Objective Optimization
    b. Similarity Between Multi-Objective Problems Via Landscape Analysis
    c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
    d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
    e. Algorithmic Recommendations Based on Semantic Knowledge
3. Conclusions and future work
    a. Conclusions
    b. Future work

1) Introduction

a. Context and motivation
b. Objectives
c. Theoretical foundations

1.a) Introduction: Context and motivation

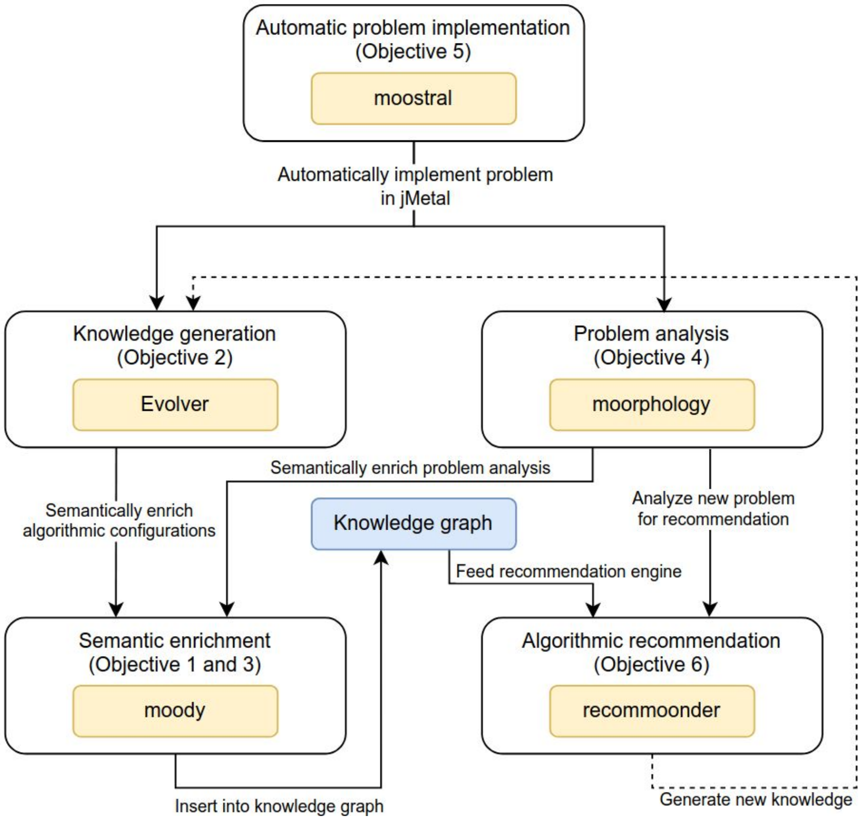
- Multi-objective optimization problems are found in a many different domains
  - Engineering, biology, medicine, economy, …
- However, domain expert that encounter this problem do not often have the expertise in optimization to choose the best algorithm/configuration and resort to default configurations
  - Default NSGA-II from the year 2000
- The automatic configuration of metaheuristics is an active line of research
  - However, is a computationally intensive process that may not be suitable for domain experts
- This PhD proposes an alternative approach based on the use of existing knowledge to provide end-users recommendations to solve their problems.

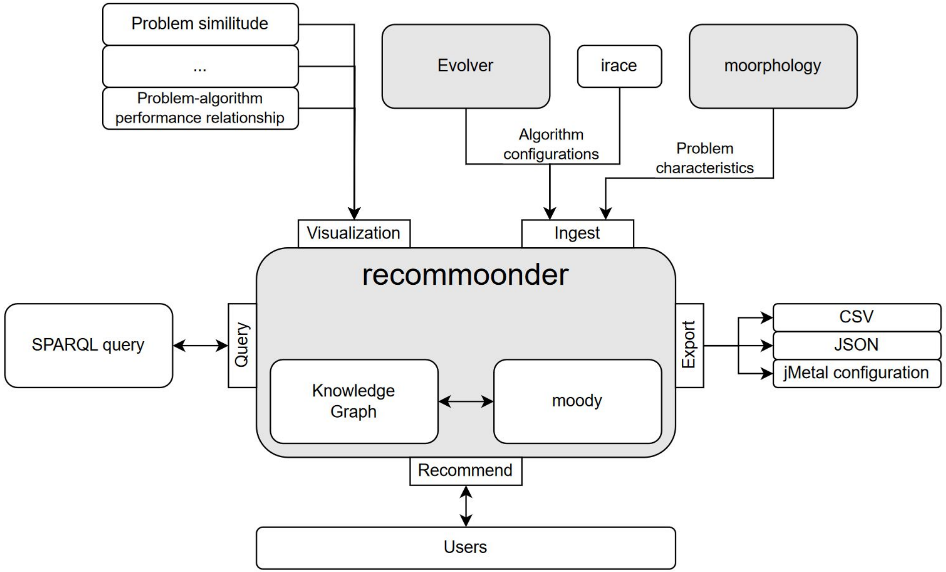
1.a) Introduction: Context and motivation

The main hypothesis of this PhD is that:

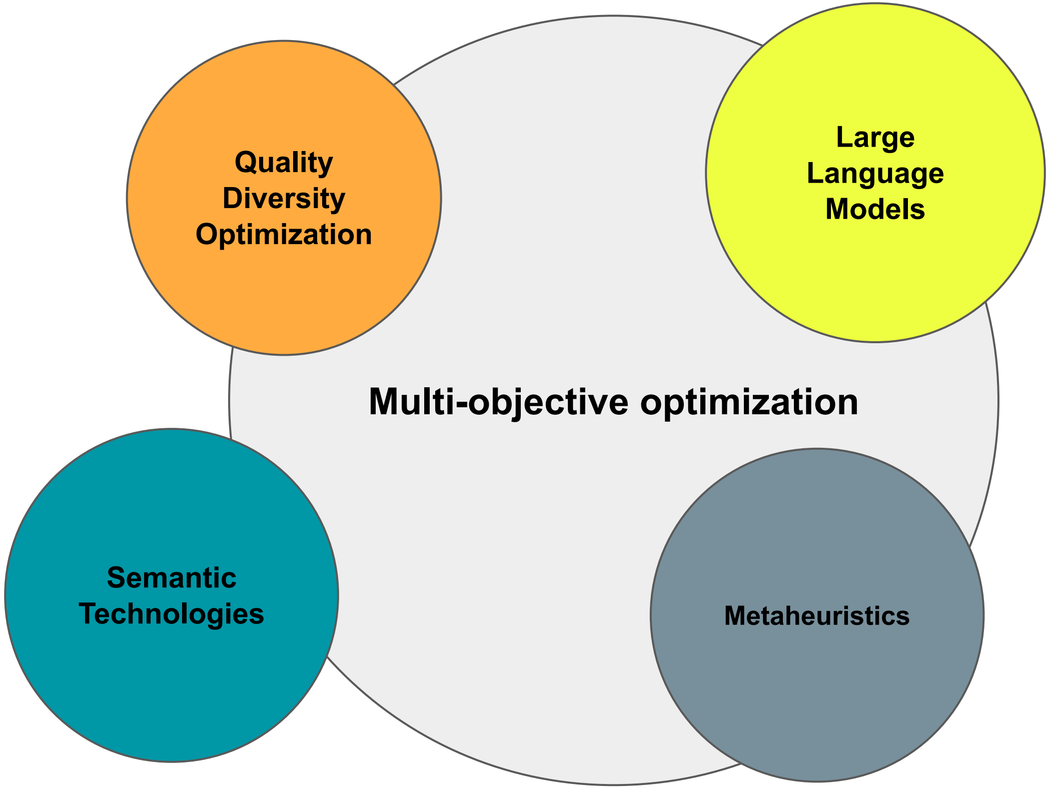
“Given previous knowledge on the relationship between a specific algorithmic configuration and the quality of the result of said algorithm solving a problem and given a similitude metric between two problems, it is possible to provide recommendations to non-expert users to choose an algorithmic configuration to efficiently solve a specific problem.”

1.b) Introduction: Objectives



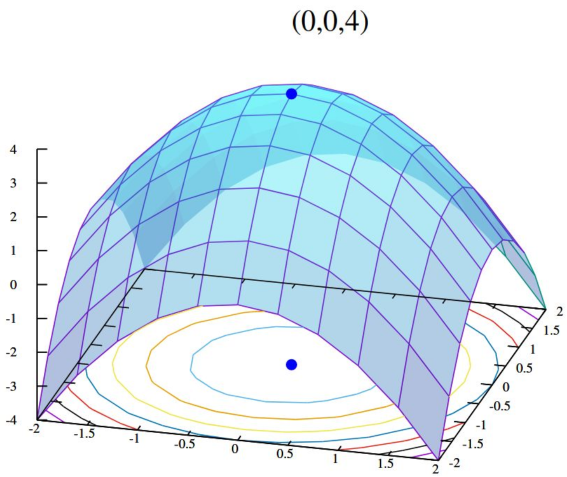


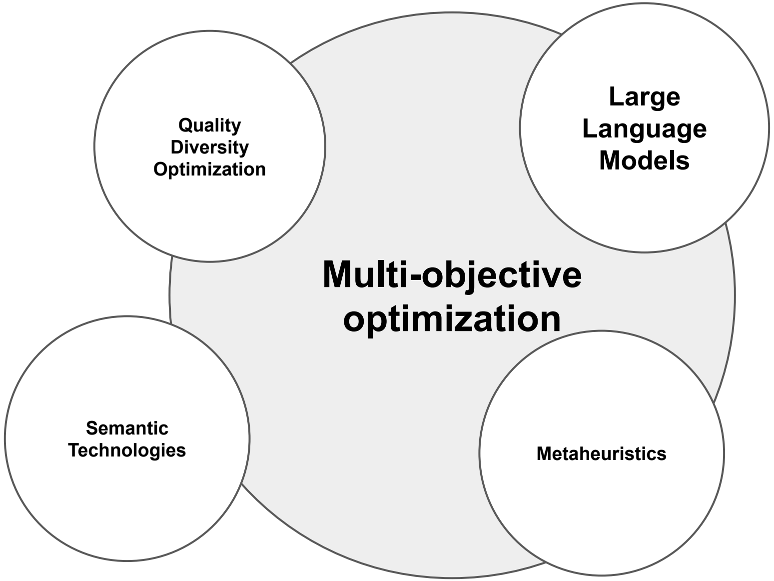
1.c) Introduction: Theoretical foundations



1.c) Introduction: Theoretical foundations

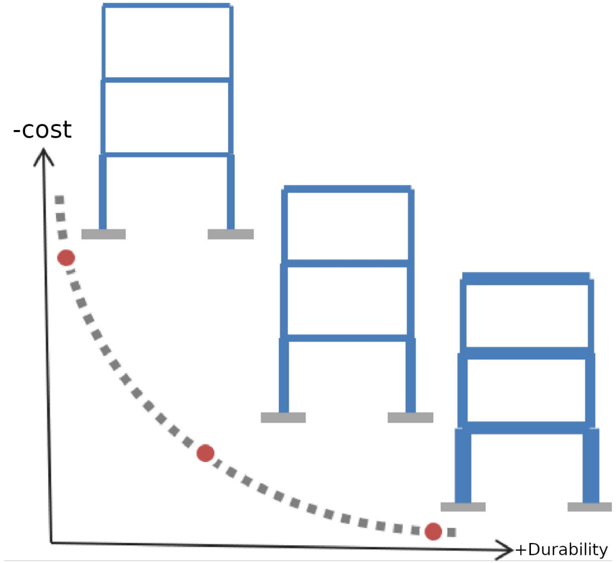
Optimization problem
- Definition
- Multi-objective optimization problem
- Quality Indicator

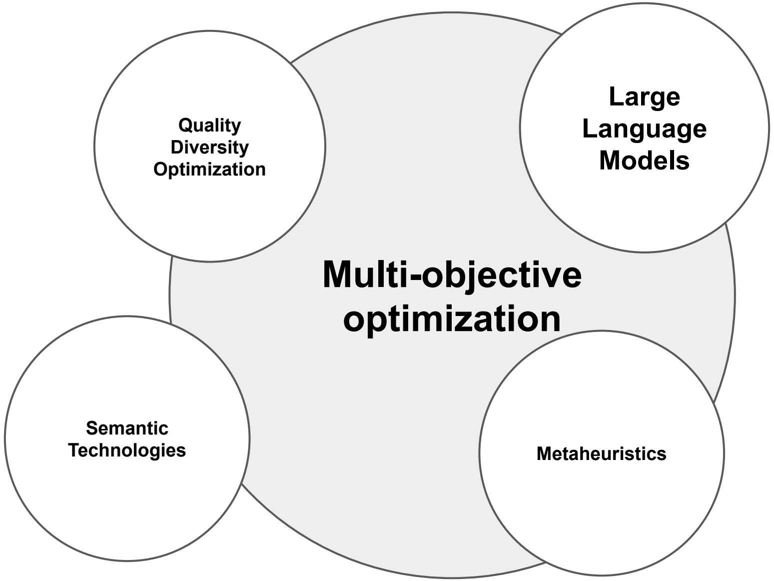




1.c) Introduction: Theoretical foundations

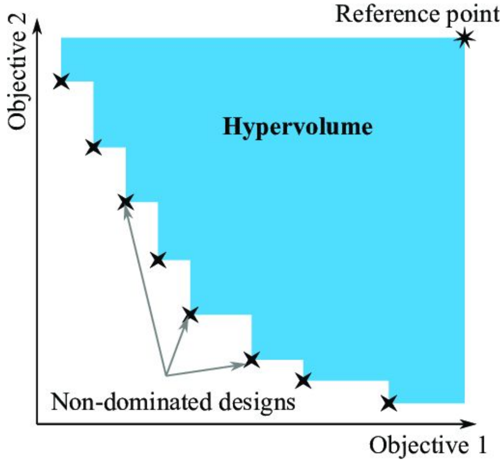
Optimization problem
- Definition
- Multi-objective optimization problem
- Quality Indicator

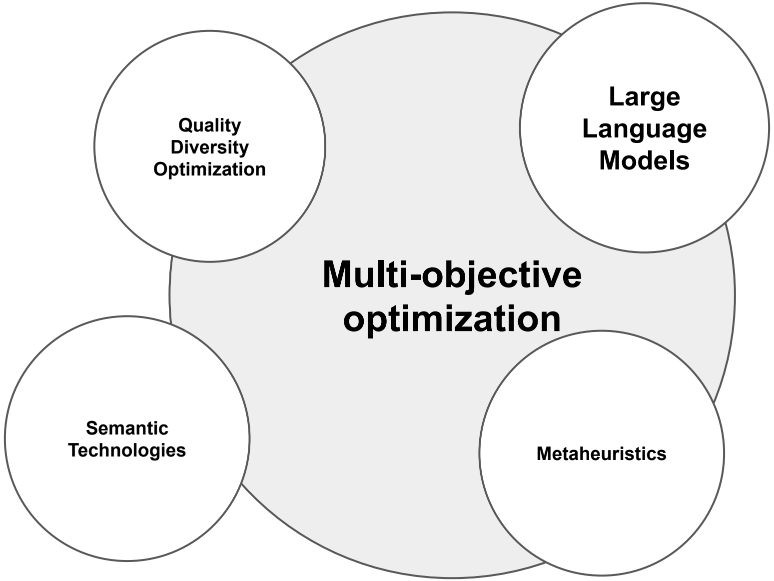




1.c) Introduction: Theoretical foundations

Optimization problem
- Definition
- Multi-objective optimization problem
- Quality Indicator



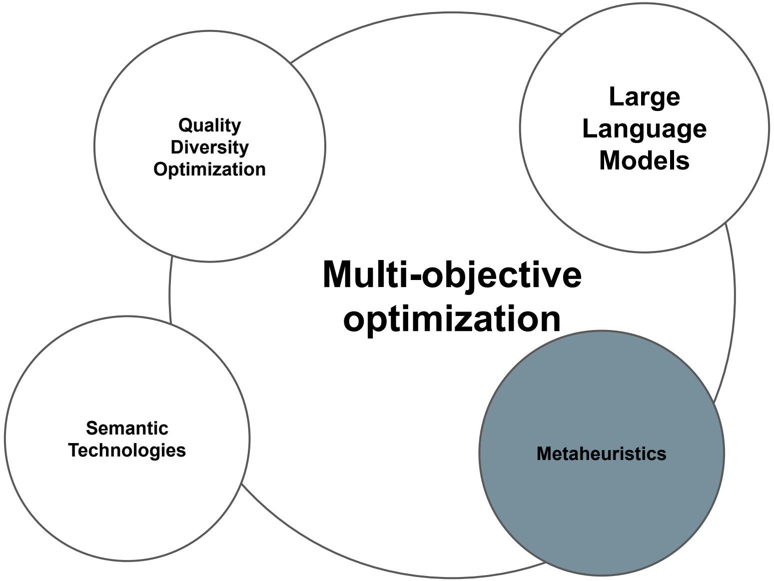


P. -D. Pfister, C. Tang and Y. Fang, "A Multi-Objective Finite-Element Method Optimization That Reduces Computation Resources Through Subdomain Model Assistance, for Surface-Mounted Permanent-Magnet Machines Used in Motion Systems," in IEEE Access, vol. 11, pp. 8609-8621, 2023, https://doi.org/10.1109/ACCESS.2023.3239214

1.c) Introduction: Theoretical foundations

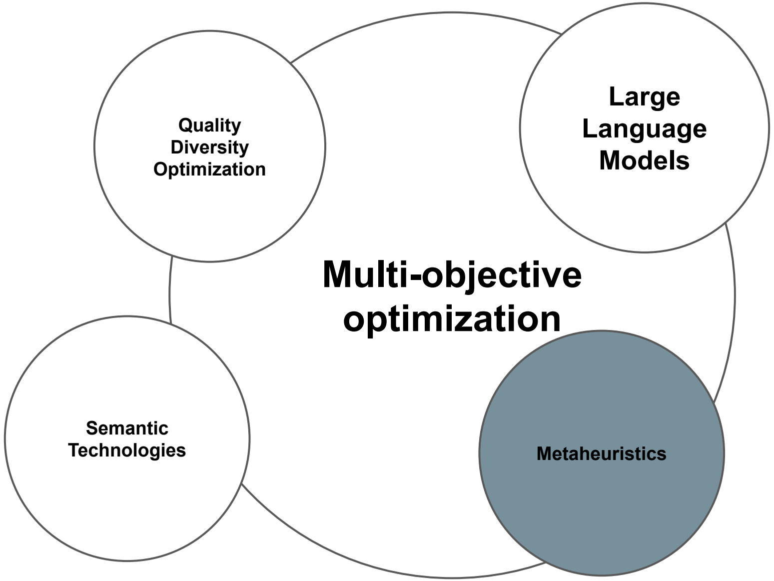
Metaheuristics
- Definition
- Evolutionary algorithm
- Optimization framework

A metaheuristic can be defined as a high level strategy to guide a set of underlying heuristics by combining different concepts for the exploration and exploitation of the search space in order to find a balance between diversification and intensification



1.c) Introduction: Theoretical foundations

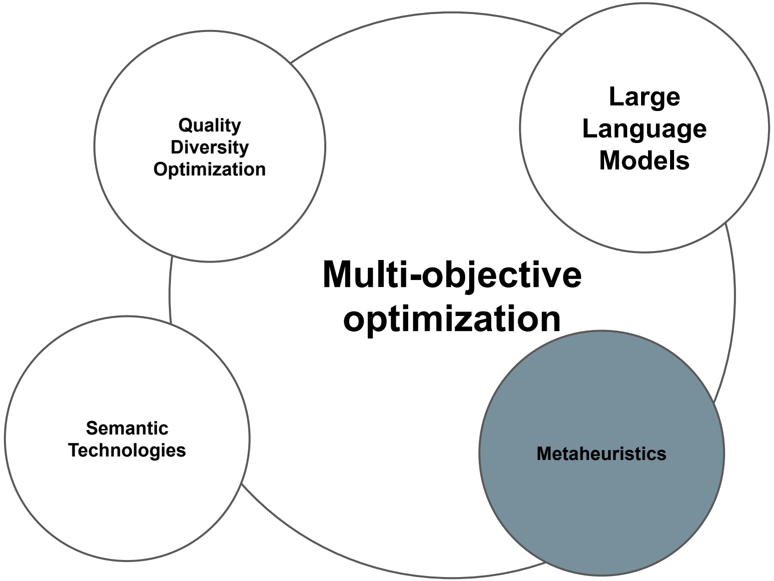
Metaheuristics
- Definition
- Evolutionary algorithm
- Optimization framework



$P(0) \leftarrow$ GenerateInitialPopulation()
$t \leftarrow 0$
Evaluation($P(0)$)
while not TerminationCriterionIsMet() **do**
  $P'(t) \leftarrow$ Selection($P(t)$)
  $Q(t) \leftarrow$ Variation($P'(t)$)
  Evaluation($Q(t)$)
  $P(t+1) \leftarrow$ Replacement($P(t)$, $Q(t)$)
  $t \leftarrow t+1$
end while

1.c) Introduction: Theoretical foundations

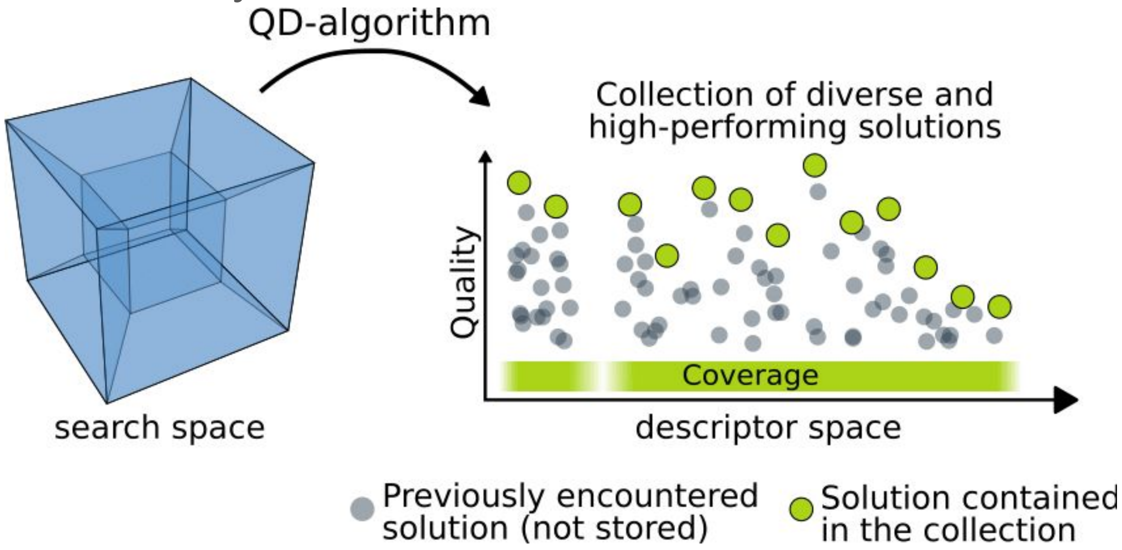
Metaheuristics
- Definition
- Evolutionary algorithm
- Optimization framework

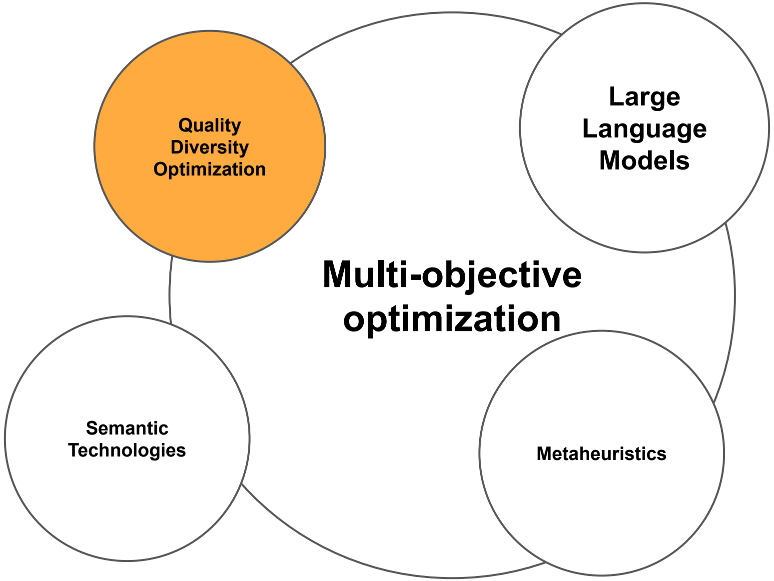


jMetal

1.c) Introduction: Theoretical foundations

Quality-Diversity Optimization
- Behavioral search space
- Diversity archive

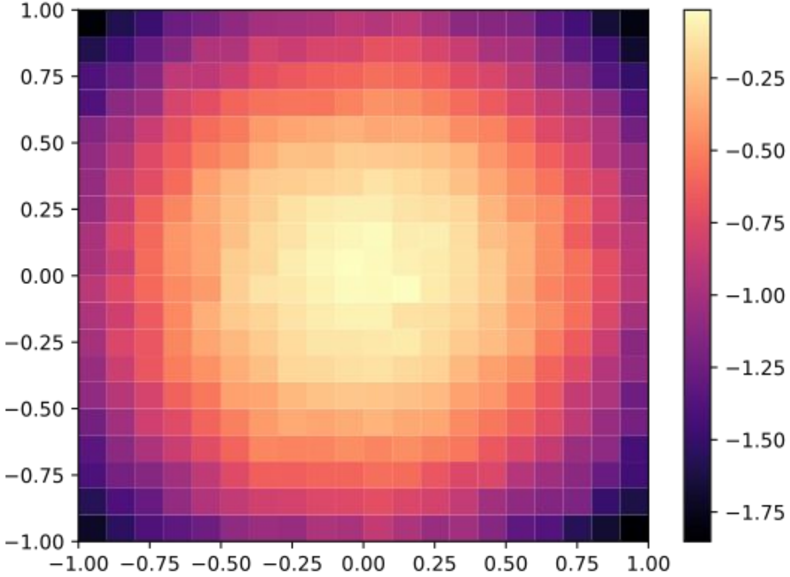


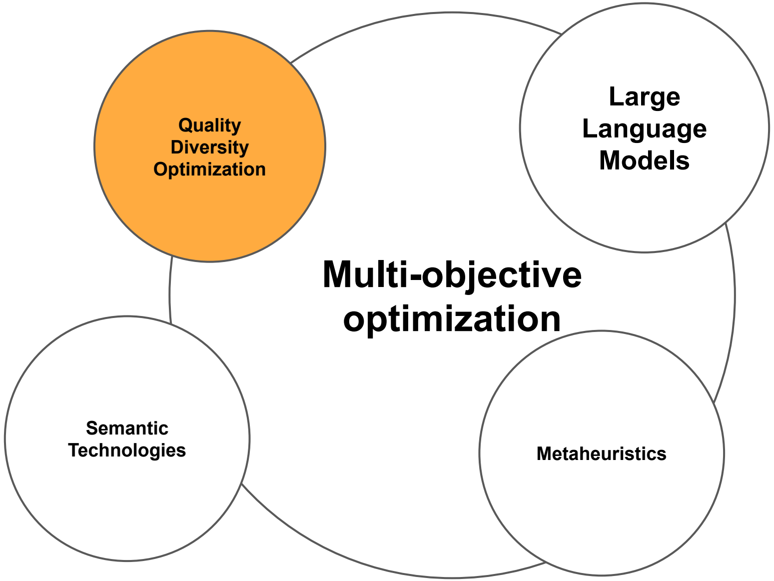


A. Cully and Y. Demiris, "Quality and Diversity Optimization: A Unifying Modular Framework," in IEEE Transactions on Evolutionary Computation, vol. 22, no. 2, pp. 245-259, April 2018, https://doi.org/10.1109/TEVC.2017.2704781

1.c) Introduction: Theoretical foundations

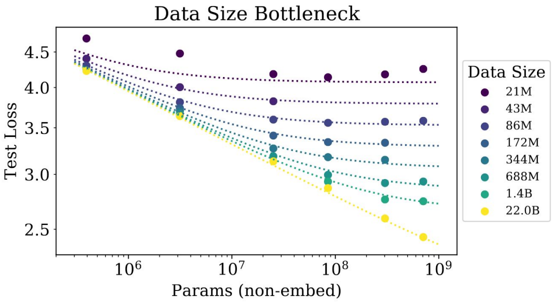
Quality-Diversity Optimization
- Behavioral search space
- Diversity archive

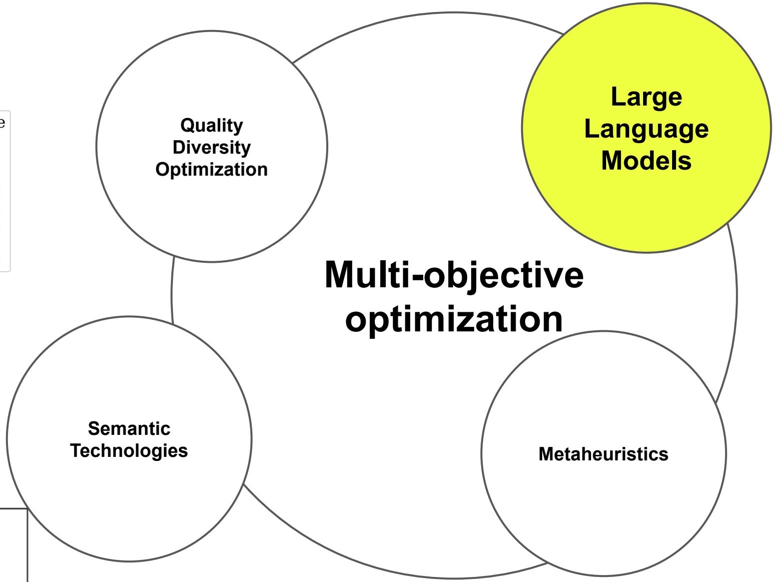




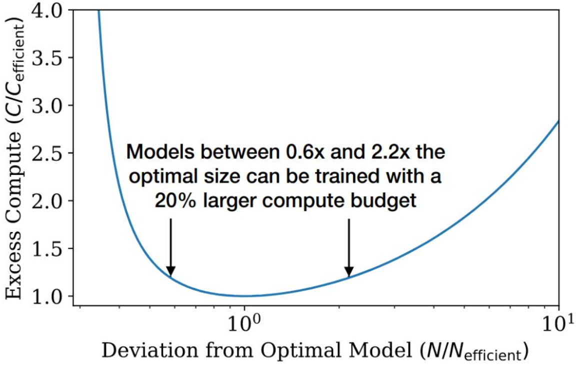
1.c) Introduction: Theoretical foundations

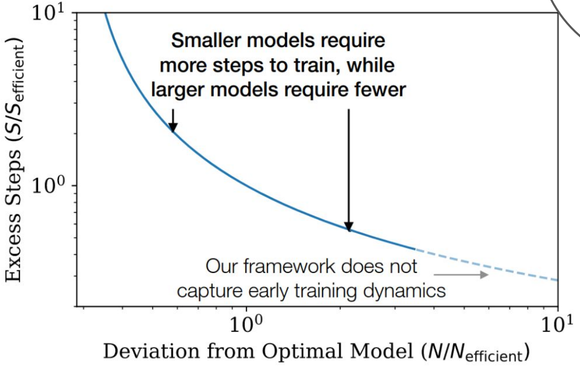
Large Language Model
- Scaling laws
- Neural architecture
- Pre-training
- Adaptation





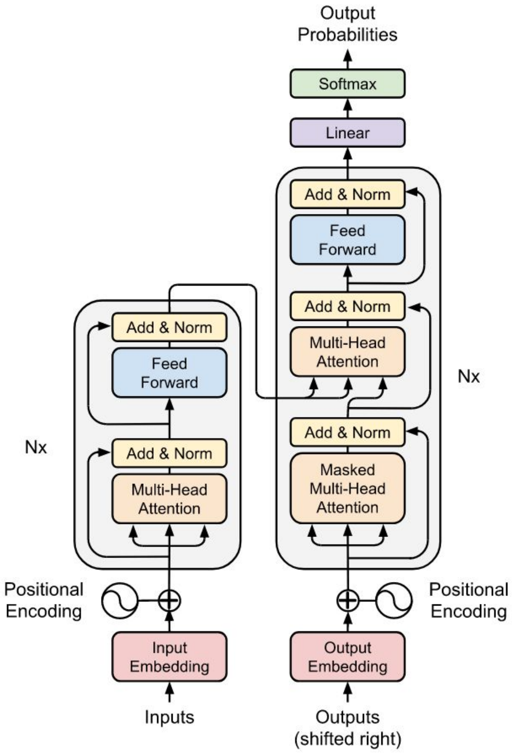
Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). Scaling Laws for Neural Language Models. https://doi.org/10.48550/arXiv.2001.08361

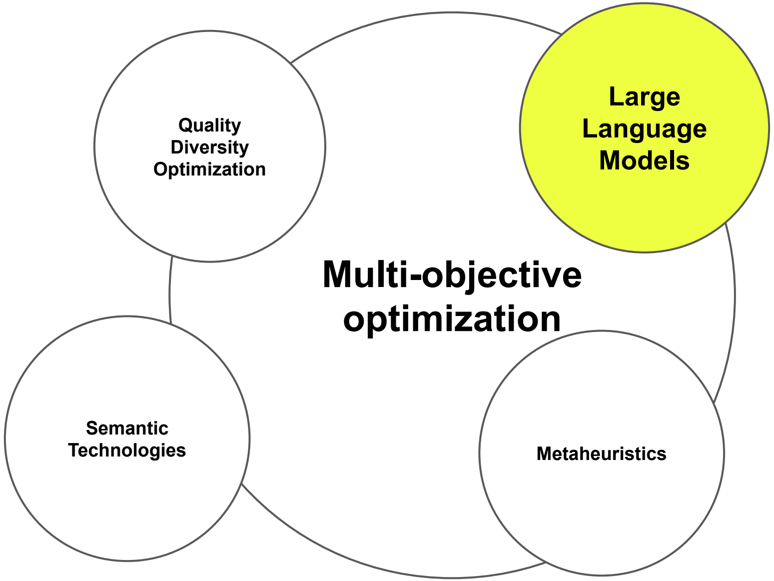




1.c) Introduction: Theoretical foundations

Large Language Model
- Scaling laws
- Neural architecture
- Pre-training
- Adaptation

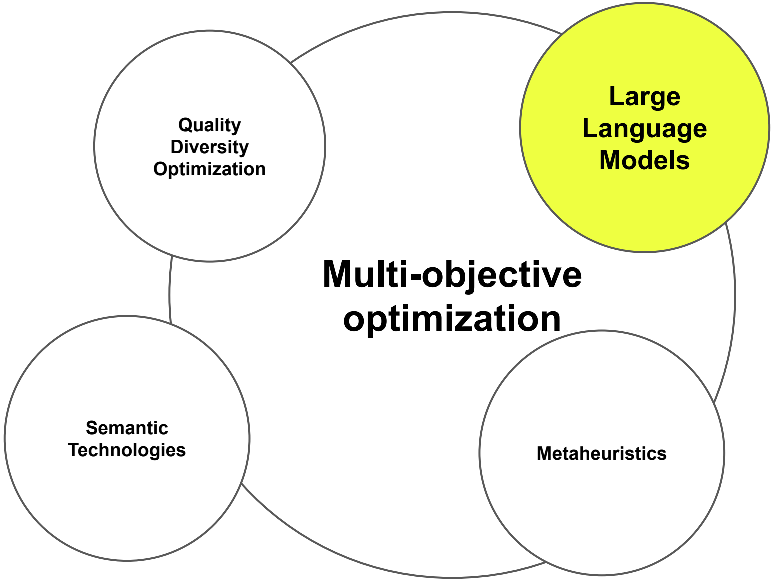




Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. https://doi.org/10.48550/arXiv.1706.03762

1.c) Introduction: Theoretical foundations

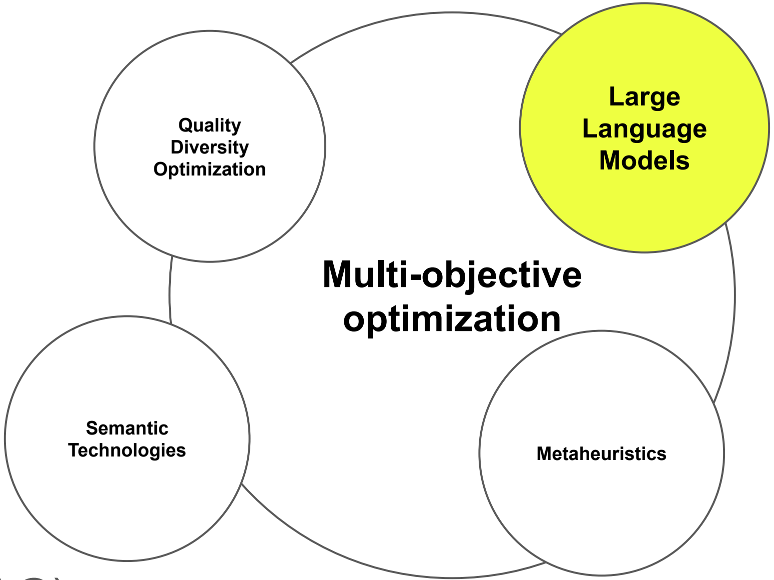
Large Language Model
- Scaling laws
- Neural architecture
- Pre-training
- Adaptation



1.c) Introduction: Theoretical foundations

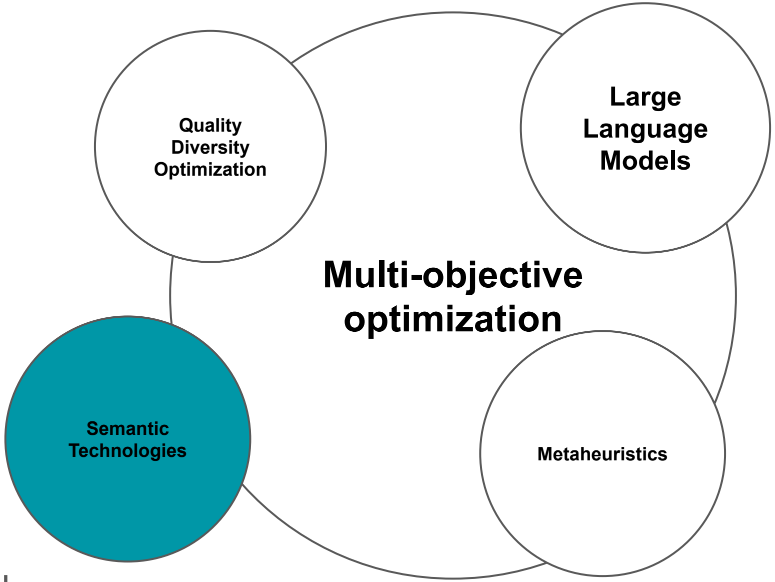
Large Language Model
- Scaling laws
- Neural architecture
- Pre-training
- Adaptation

Impact/Cost/Complexity	
High	Pre-training
	Fine-tuning Retrieval Augmented Generation (RAG)
Low	Prompt engineering One- and few- shot learning



1.c) Introduction: Theoretical foundations

Semantic Technologies
- Ontology (OWL)
- RDF
- SPARQL
- Semantic Reasoning (SWRL)
- Knowledge Graph

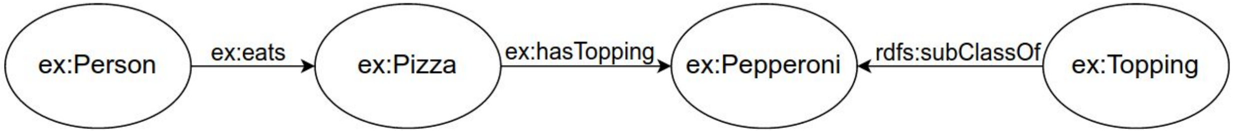


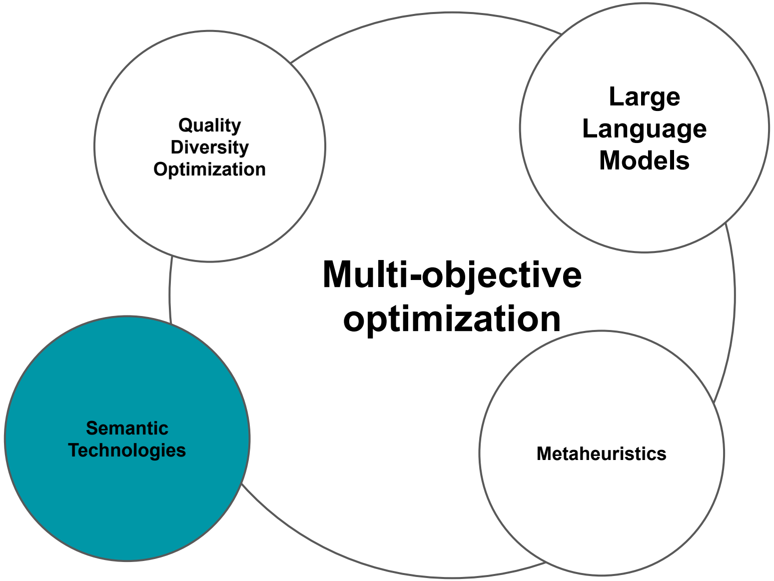
Ontology: Formal, logic-based approach for defining concepts and establishing common vocabulary.

1.c) Introduction: Theoretical foundations

Semantic Technologies
- Ontology (OWL)
- RDF
- SPARQL
- Semantic Reasoning (SWRL)
- Knowledge Graph

```
PREFIX ex: <http://example.org/>
SELECT DISTINCT ?person
WHERE {
    ?person ex:eats ?pizza .
    ?pizza ex:hasTopping ex:Pepperoni .
}
```

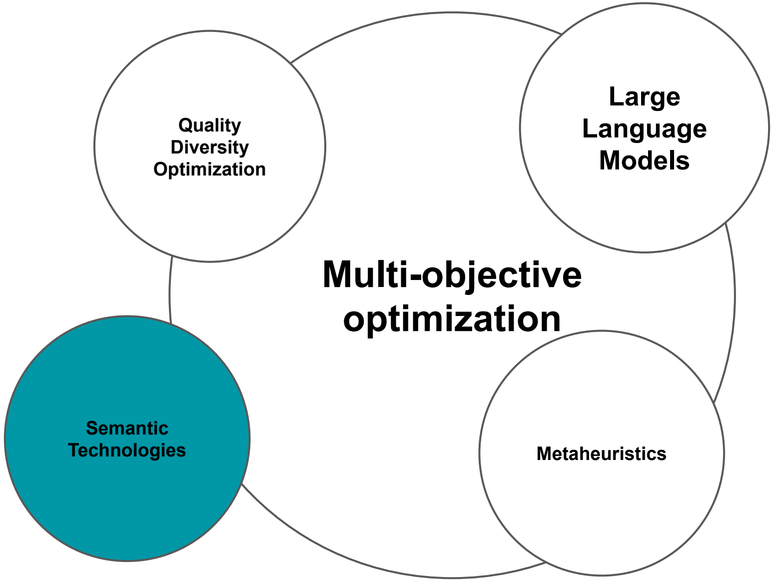


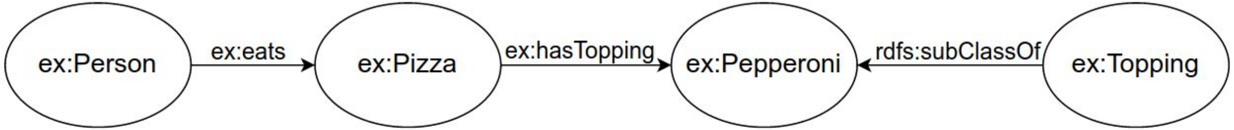


1.c) Introduction: Theoretical foundations

Semantic Technologies
- Ontology (OWL)
- RDF
- SPARQL
- Semantic Reasoning (SWRL)
- Knowledge Graph

```
ex:Pepperoni(?topping)
^   ex:hasTopping(?x, ?topping)
-> ex:Pizza(?x)
```

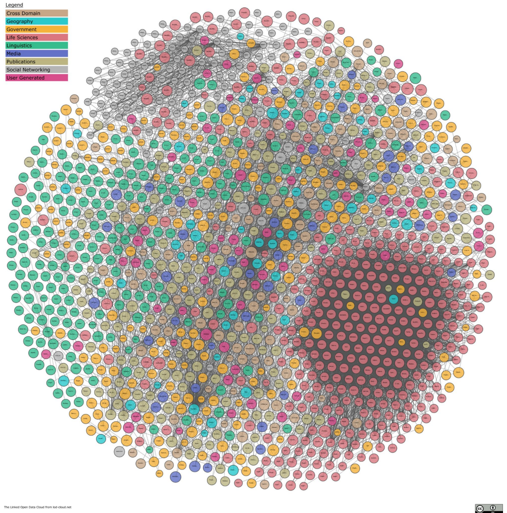


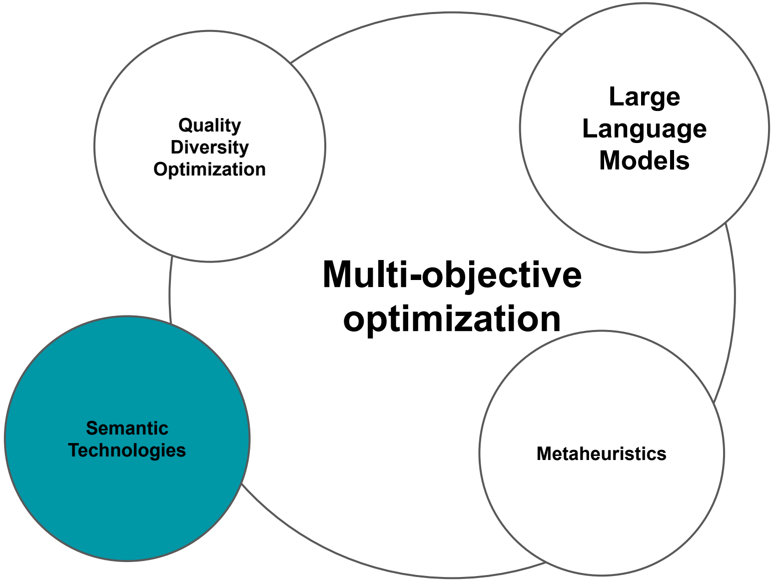


1.c) Introduction: Theoretical foundations

Semantic Technologies
- Ontology (OWL)
- RDF
- SPARQL
- Semantic Reasoning (SWRL)
- Knowledge Graph

The linked Open Data Cloud from
https://lod-cloud.net/





2) PhD contributions

a. A Semantic Approach to Standardizing Multi-Objective Optimization
b. Similarity Between Multi-Objective Problems Via Landscape Analysis
c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
e. Algorithmic Recommendations Based on Semantic Knowledge

Information Sciences 661 (2024) 120184

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

MOODY: An ontology-driven framework for standardizing multi-objective evolutionary algorithms

José F. Aldana-Martín *, María del Mar Roldán-García, Antonio J. Nebro, José F. Aldana-Montes

Dept. de Lenguajes y Ciencias de la Computación, ITIS Software, University of Málaga, ETSI Informática, Campus de Teatinos, Málaga, 29071, Spain

Software tool

moody -

multi-objective optimization ontology

2.a) Motivation

- Lack of Standardization: No unified approach for designing multi-objective optimization algorithms.
  - Researchers often create their own components, leading to inconsistencies.
  - Difficult to compare results due to the absence of a unified framework.
- Computational Demands: Auto-configuration requires significant computational resources, knowledge should be re-used.
  - Thousands of configurations need to be generated and evaluated.
  - Identifying existing configurations for similar problems can save time but methods are unclear.

2.a) Contributions

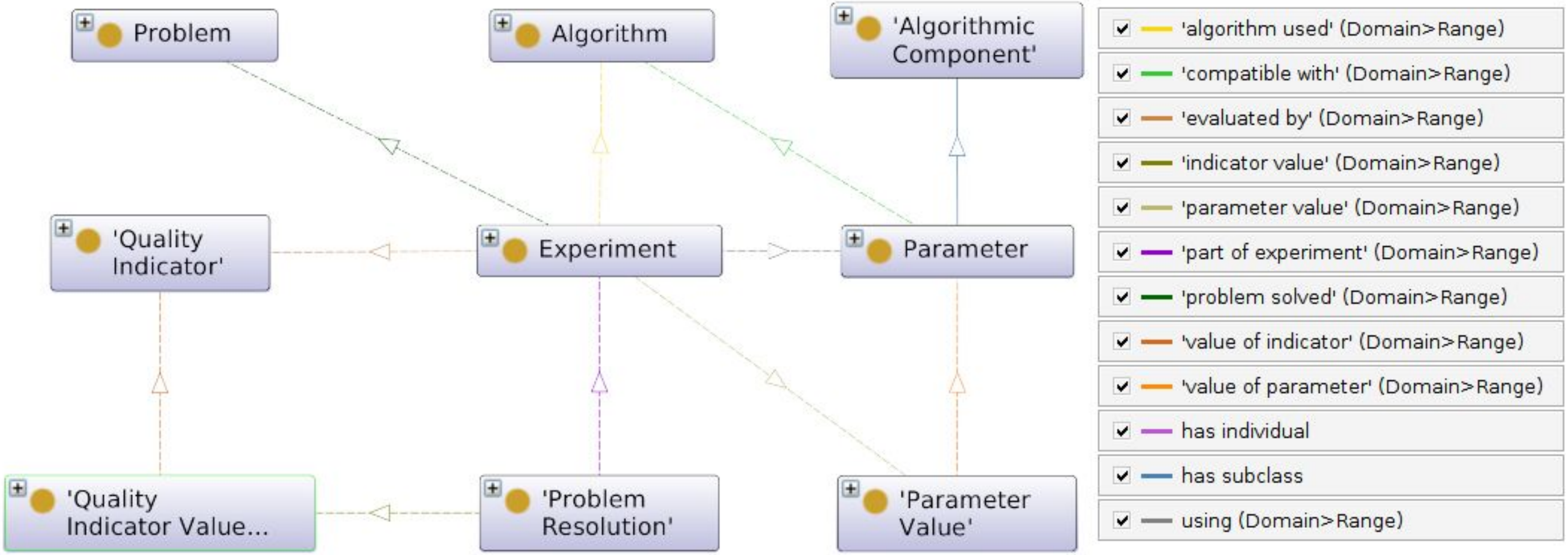
- Ontology Development - moody:
  - Formalizes aspects of multi-objective evolutionary algorithms and their parameters.
  - Integrates configurations into a knowledge graph for cross-compatibility.
  - Follows the FAIR principles. Available at permanent URL: https://w3id.org/moody
  - Enables advanced reasoning and recommendations for superior algorithm configurations.
- Knowledge Graph Creation:
  - Populated with algorithm configurations and optimization experiments.
  - Annotated semantically and formatted in RDF.
- Validated by 4 use cases
  - Enhancing auto-configuration tools
  - Integration of algorithmic configuration from diverse sources
  - SPARQL queries to extract valuable insight
  - Re-using of knowledge between different configuration frameworks

2.a) Semantic approach

- Implementation:
  - moody is implemented as an OWL 2 ontology.
  - Following the Ontology Development 101 methodology in 7 steps:
    1. Determine the domain and scope of the ontology.
    2. Consider reusing existing ontologies.
    3. Enumerate important terms in the ontology.
    4. Define classes and the class hierarchy.
    5. Define the properties of classes and slots.
    6. Define the facets of the slots.
    7. Create instances.

2.a) Semantic approach

Main concepts



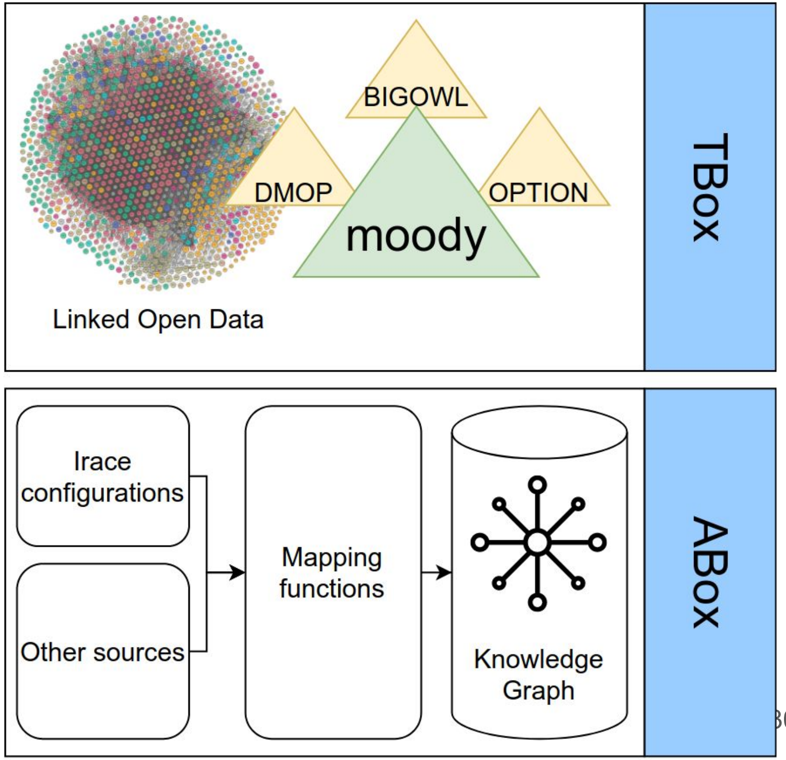
2.a) Domain and ranges of properties

Parameter	range
Aggregative function	{"PenaltyBoundaryIntersection", "Tschebycheff", "WeightedSum"}
Algorithm result	{"externalArchive", "population"}
BLX alpha crossover alpha value	xsd:double[>= "0.0"^^xsd:double, <= "1.0"^^xsd:double]
Create initial solutions	{"latinHypercubeSampling", "random", "scatterSearch"}
Crossover probability	xsd:double[>= "0.0"^^xsd:double, <= "1.0"^^xsd:double]
Crossover repair strategy	{"bounds", "random", "round"}
Crossover	{"BLX_ALPHA", "SBX"}
Maximum number of evaluations	xsd:integer
Maximum number of replaced solutions	xsd:integer
Mutation probability	xsd:double[>= "0.0"^^xsd:double, <= "1.0"^^xsd:double]
Mutation repair strategy	{"bounds", "random", "round"}
Mutation	{"polynomial", "uniform"}
Neighborhood selection probability	xsd:double[>= "0.0"^^xsd:double, <= "1.0"^^xsd:double]
Neighborhood size	xsd:integer
Offspring population size	xsd:integer
Polynomial mutation distribution index	xsd:double[>= "5.0"^^xsd:double, <= "400.0"^^xsd:double]
Population size	xsd:integer
Population size with archive	xsd:integer
SBX crossover distribution index	xsd:double[>= "5.0"^^xsd:double, <= "400.0"^^xsd:double]
Selection tournament size	xsd:integer[>=2, <=10]
Selection	{"random", "tournament"}
Uniform mutation perturbation	xsd:double[>= "0.0"^^xsd:double, <= "1.0"^^xsd:double]

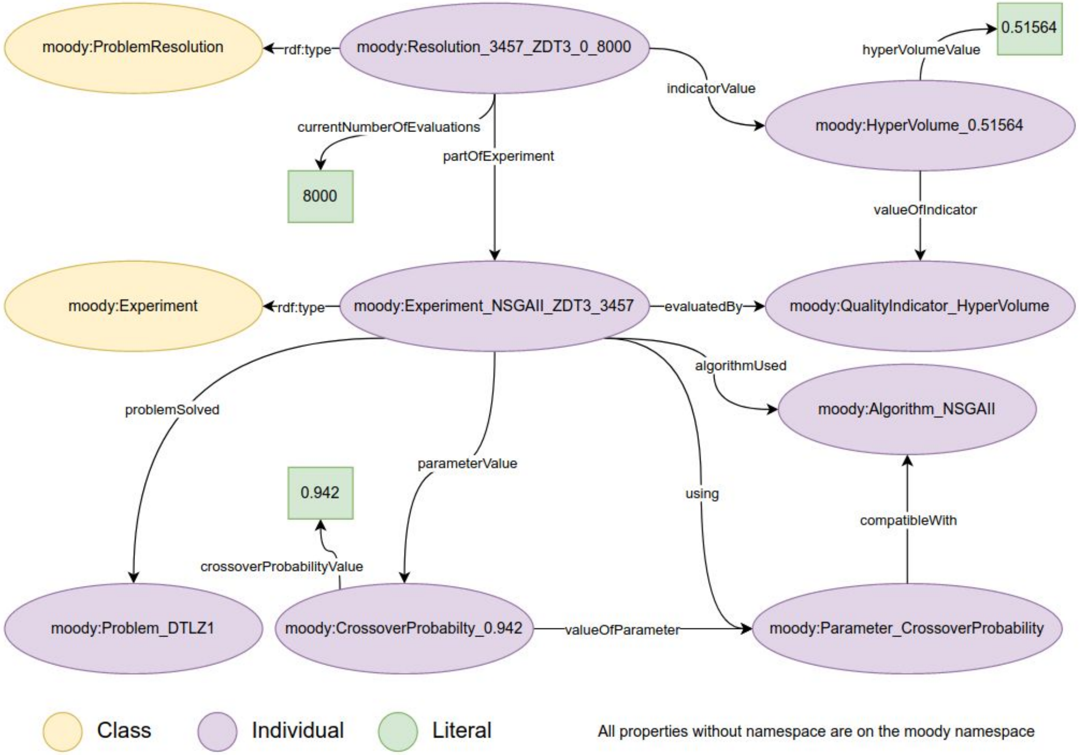
2.a) Mapping to RDF

Data from different sources are integrated using mapping functions:

Possible sources are optimization frameworks and auto-configuration tools.



2.a) Mapping to RDF



2.a) Use cases

Automatic validation of configurations via semantic reasoning

```
moody:Experiment(?x)
^ moody:SbxCrossoverDistributionIndex(?p)
^ moody:sbxCrossoverDistributionIndexValue(?pv, ?cv)
^ moody:parameterValue(?x, ?pv)
^ moody:valueOfParameter(?pv, ?p)
^ moody:Crossover(?p2)
^ moody:crossoverValue(?pv2, ?cv2)
^ moody:parameterValue(?x, ?pv2)
^ moody:valueOfParameter(?pv2, ?p2)
^ swrlb:equal(?cv2, "BLX_ALPHA")
-> moody:InvalidExperiment(?x)
```

2.a) Use cases

Automatic validation of configurations via semantic reasoning

```
moody:Experiment(?x)
^ moody:SbxCrossoverDistributionIndex(?p)
^ moody:sbxCrossoverDistributionIndexValue(?pv, ?cv)
^ moody:parameterValue(?x, ?pv)
^ moody:valueOfParameter(?pv, ?p)
^ moody:Crossover(?p2)
^ moody:crossoverValue(?pv2, ?cv2)
^ moody:parameterValue(?x, ?pv2)
^ moody:valueOfParameter(?pv2, ?p2)
^ swrlb:equal(?cv2, "BLX_ALPHA")
-> moody:InvalidExperiment(?x)
```

Non compatible components

2.a) Use cases

Re-using of knowledge between different configuration frameworks

Use configurations found in jMetal for algorithms implemented in pagmo

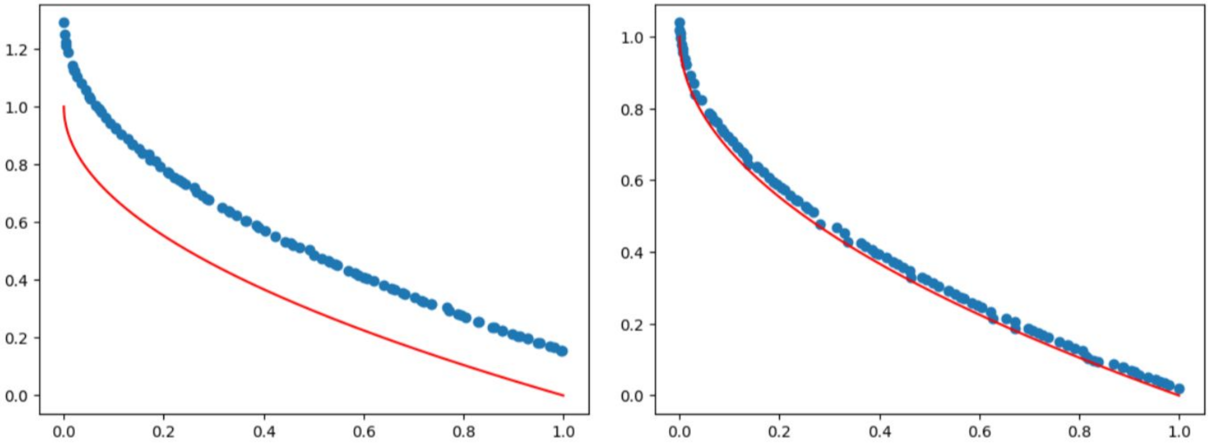


Figure 3.5: In red, reference front for the ZDT problem. In blue, Front for the ZDT4 problem obtained by NSGA-II with standard settings (left), and front obtained by exporting a configuration from *moody* (right). The target framework is pagmo.

2) PhD contributions

a. A Semantic Approach to Standardizing Multi-Objective Optimization
b. Similarity Between Multi-Objective Problems Via Landscape Analysis
c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
e. Algorithmic Recommendations Based on Semantic Knowledge

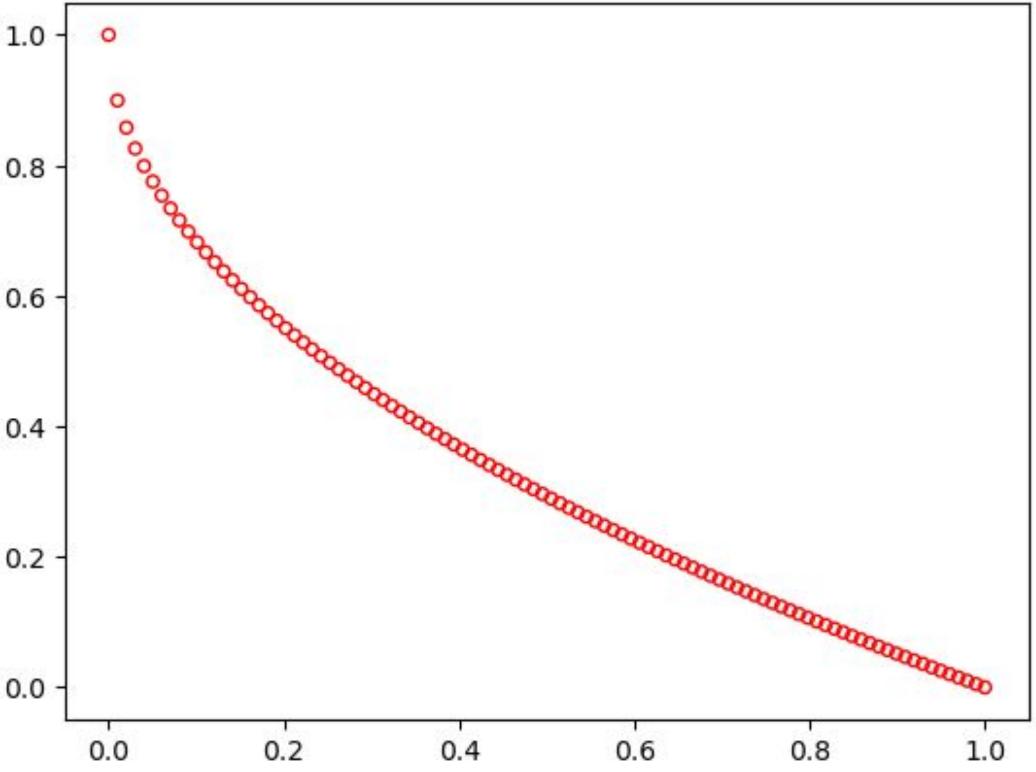
Software tool

moorphology -
Characterization of continuous multi-objective problems by their landscape

2.b) Motivation

- To validate this PhD hypothesis, a metric that describe how similar two continuous multi-objective problems are without the need of any domain expertise.
- Landscape analysis is the field that studies the topological and structural characteristics of optimization problems.

2.b) Landscape of optimization problems



ZDT1

Eckart Zitzler, Kalyanmoy Deb, Lothar Thiele; Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. Evol Comput 2000; 8 (2): 173–195. https://doi.org/10.1162/106365600568202

2.b) Landscape of optimization problems

Challenging test problems for multi- and many-objective optimization,
Swarm and Evolutionary Computation,
Volume 81, 2023,
https://doi.org/10.1016/j.swevo.2023.101350

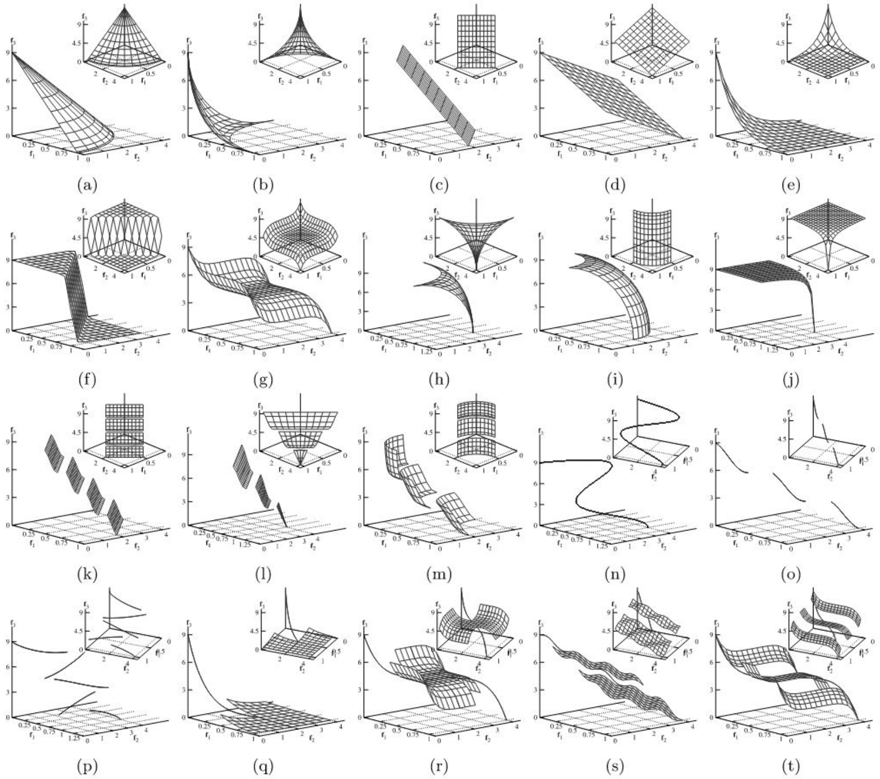


Fig. 2. Pareto fronts of ZCAT test problems using $M = 3$ for: (a) ZCAT1, (b) ZCAT2, (c) ZCAT3, (d) ZCAT4, (e) ZCAT5, (f) ZCAT6, (g) ZCAT7, (h) ZCAT8, (i) ZCAT9, (j) ZCAT10, (k) ZCAT11, (l) ZCAT12, (m) ZCAT13, (n) ZCAT14, (o) ZCAT15, (p) ZCAT16, (q) ZCAT17, (r) ZCAT18, (s) ZCAT19, and (t) ZCAT20.

2.b) Contributions

- moorphology - Software package for landscape analysis in continuous multi-objective optimization problems
- Design of a similarity metric between multi-objective problems based on landscape analysis.
- Evaluation of the tool over known benchmark problems.

2.b) Landscape characteristics

Characterize both the variable and objective search spaces.

Type	Name	Description
Problem-dependent	name	Name of the problem being sampled
	n_var	Number of variables
	n_obj	Number of objectives
	n_cons	Number of constraints
	sample_size	Number of samples used to extract the landscape characteristics
Global		
	dist_x_avg	Average distance among solutions in the variable space
	dist_x_max	Maximum distance among solutions in the variable space
	dist_f_avg	Average distance among solutions in the objective space
	dist_f_max	Maximum distance among solutions in the objective space
	nd_n	Proportion of non-dominated solutions
	dist_x_nd_avg	Average distance among non-dominated solutions in the variable space
	dist_x_nd_max	Maximum distance among non-dominated solutions in the variable space
	rank_avg	Average rank w.r.t. non-dominated sorting
	rank_max	Maximum rank w.r.t. non-dominated sorting
	rank_ent	Entropy of the number of solutions per rank w.r.t. non-dominated sorting
Evolvability		
	sup_avg_neig	Average proportion of dominating neighbours
	inf_avg_neig	Average proportion of dominated neighbours
	inc_avg_neig	Average proportion of incomparable neighbours
	lnd_avg_neig	Average proportion of locally non-dominated neighbours
	lsupp_avg_neig	Average proportion of supported locally non-dominated neighbours
	dist_x_avg_neig	Average distance from neighbours in the variable space
	dist_x_max_neig	Maximum distance from neighbours in the variable space
	dist_f_avg_neig	Average distance from neighbours in the objective space
	dist_f_max_neig	Maximum distance from neighbours in the objective space
Ruggedness		
	dist_x_cor_neig	Neighbour's correlation of the average distance from neigh. in the variable space
	dist_f_cor_neig	Neighbour's correlation of the average distance from neigh. in the objective space

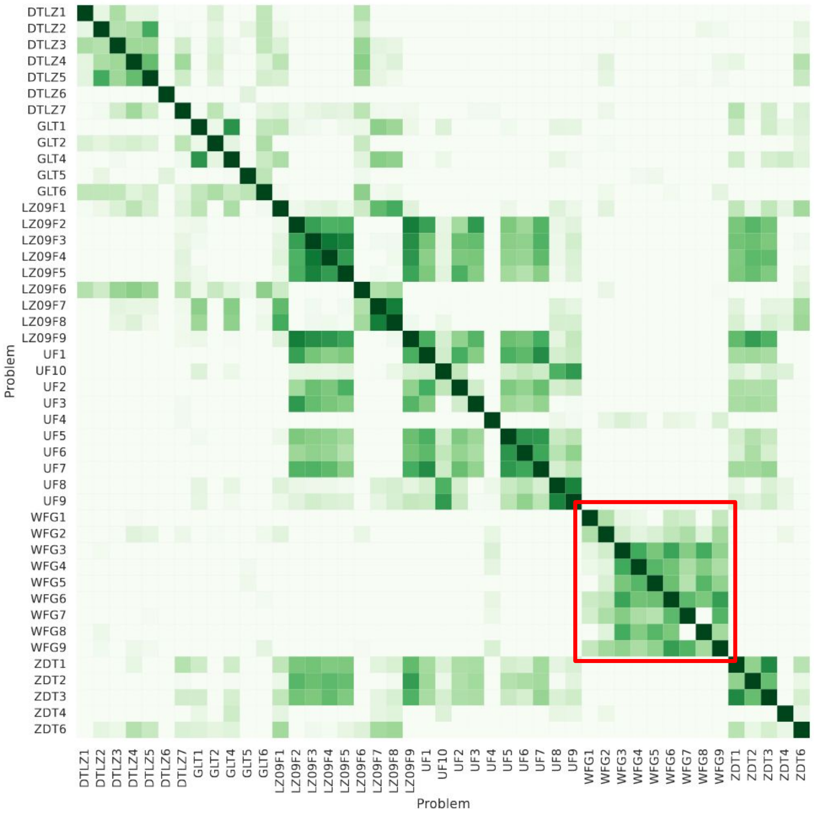
2.b) Landscape characteristics

Analyzed:

- 26 characteristics
- for 45 problems
- across 6 families

Results:

- Found clusters inside some families
- Verified similarities between families due to authors re-using functions



2) PhD contributions

a. A Semantic Approach to Standardizing Multi-Objective Optimization
b. Similarity Between Multi-Objective Problems Via Landscape Analysis
c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
e. Algorithmic Recommendations Based on Semantic Knowledge

A Study About Meta-Optimizing the NSGA-II Multi-Objective Evolutionary Algorithm

José F. Aldana-Martín[2], Antonio J. Nebro[1,2], Juan J. Durillo[3], and María del Mar Roldán García[1,2]

[1] Departamento de Lenguajes y Ciencias de la Computación. University of Málaga, 29071 Málaga, Spain
[2] ITIS Software, University of Málaga, 29071, Málaga, Spain
jfaldanam@uma.es, ajnebro@uma.es, mrgarcia@uma.es
[3] Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Germany
durillo@lrz.de

In 9th International Conference on Metaheuristics and Nature Inspired Computing (META 2023).

SoftwareX 24 (2023) 101551

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

ELSEVIER

SoftwareX

Check for updates

Original software publication

Evolver: Meta-optimizing multi-objective metaheuristics

José F. Aldana-Martín [a,b,*], Juan J. Durillo [c], Antonio J. Nebro [a,b]

[a] ITIS Software, Edificio de Investigación Ada Byron, University of Málaga, Málaga, 29071, Spain
[b] Dept. de Lenguajes y Ciencias de la Computación, University of Málaga, Málaga, 29071, Spain
[c] Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Garching bei Muenchen, Germany

Software tool
Evolver - Meta-optimizing multi-objective metaheuristics

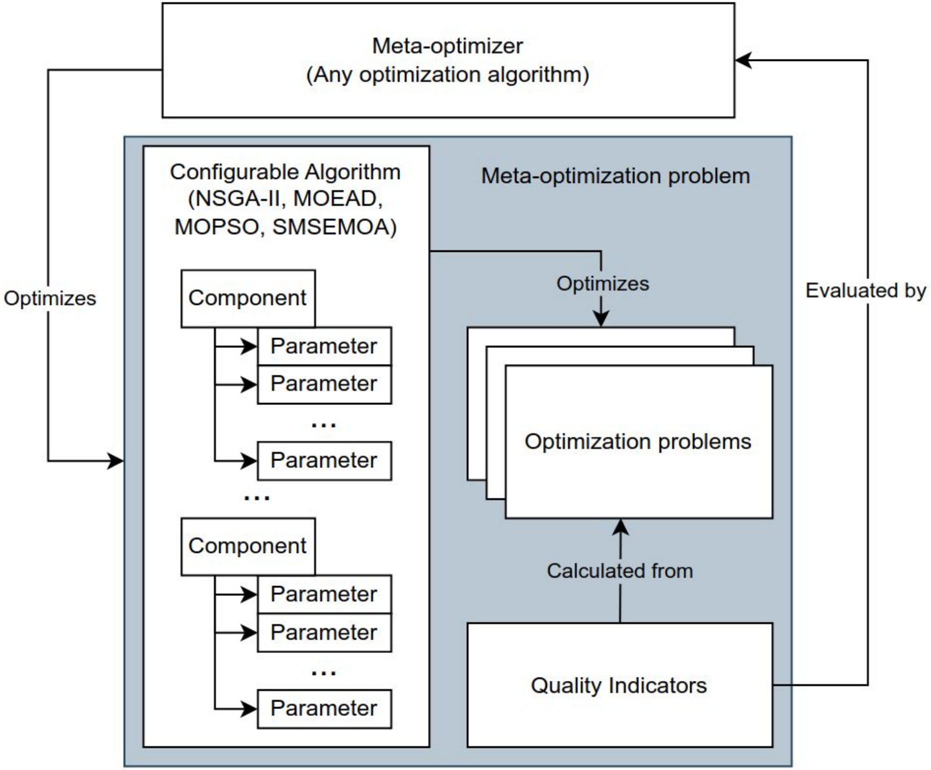
2.c) Motivation

- To populate the knowledge graph that will power the recommendation system, an automatic way to generate good configurations is required.
- To fill the knowledge graph, state-of-the-art algorithms for the auto-configuration of algorithms were explored, like irace
  - This tools were usually very slow to generate a large number of configurations. So we develop a novel meta-optimization approach.

2.c) Contributions

- A study is conducted about the use of NSGA-II to find configurations of NSGA-II, i.e., using NSGA-II as a meta-optimizer.
  - The basic idea is to consider the auto-design of NSGA-II as a multi-objective problem, where the decision variables represent parameters and components and the objectives can be combinations of quality indicators
- A auto-configuration framework tool, named Evolver, implemented within the jMetal framework

2.c) Meta-optimization approach



2.c) Experimental study

Experimentally, we were able to replicate the results of a previous study with the state-of-the-art method irace + jMetal

	NSGAII	SMPSO	AutoNSGAII
WFG1	$4.35e-01_{1.8e-01}$	$1.17e-01_{8.0e-03}$	$6.34e-01_{1.8e-05}$
WFG2	$5.61e-01_{2.3e-03}$	$5.61e-01_{1.6e-03}$	$5.64e-01_{9.2e-05}$
WFG3	$4.92e-01_{8.3e-04}$	$4.92e-01_{6.1e-04}$	$4.95e-01_{6.1e-05}$
WFG4	$2.17e-01_{3.7e-04}$	$2.03e-01_{2.4e-03}$	$2.18e-01_{1.5e-03}$
WFG5	$1.95e-01_{3.5e-04}$	$1.96e-01_{7.8e-05}$	$1.96e-01_{9.8e-05}$
WFG6	$2.01e-01_{1.3e-02}$	$2.09e-01_{5.0e-04}$	$2.02e-01_{1.4e-02}$
WFG7	$2.09e-01_{5.5e-04}$	$2.09e-01_{2.7e-04}$	$2.11e-01_{3.5e-05}$
WFG8	$1.47e-01_{2.7e-03}$	$1.47e-01_{3.4e-03}$	$1.40e-01_{3.1e-03}$
WFG9	$2.37e-01_{1.8e-03}$	$2.35e-01_{5.8e-04}$	$2.39e-01_{2.0e-03}$
DTLZ1	$4.88e-01_{7.9e-03}$	$4.94e-01_{2.7e-04}$	$0.00e+00_{4.9e-01}$
DTLZ2	$2.09e-01_{4.7e-04}$	$2.10e-01_{1.6e-04}$	$2.11e-01_{3.6e-05}$
DTLZ3	$0.00e+00_{1.8e-02}$	$2.10e-01_{1.2e-01}$	$0.00e+00_{0.0e+00}$
DTLZ4	$2.09e-01_{2.1e-01}$	$2.10e-01_{8.7e-05}$	$2.11e-01_{7.2e-05}$
DTLZ5	$2.11e-01_{3.4e-04}$	$2.12e-01_{2.2e-04}$	$2.12e-01_{4.5e-05}$
DTLZ6	$1.82e-01_{3.6e-02}$	$2.12e-01_{8.1e-05}$	$2.12e-01_{4.2e-05}$
DTLZ7	$3.34e-01_{3.0e-04}$	$3.35e-01_{1.2e-04}$	$3.35e-01_{9.2e-05}$

(a) Current study results using as objective the hypervolume indicator.

	NSGAII	SMPSO	AutoNSGAII
WFG1	$4.49e-01_{7.6e-02}$	$1.16e-01_{7.7e-03}$	$6.34e-01_{2.6e-05}$
WFG2	$5.64e-01_{9.5e-04}$	$5.62e-01_{1.2e-03}$	$5.65e-01_{5.1e-05}$
WFG3	$4.41e-01_{3.8e-04}$	$4.41e-01_{2.2e-04}$	$4.42e-01_{1.1e-05}$
WFG4	$2.17e-01_{7.6e-04}$	$2.03e-01_{2.4e-03}$	$2.17e-01_{3.0e-03}$
WFG5	$1.95e-01_{2.9e-04}$	$1.96e-01_{7.5e-05}$	$1.96e-01_{1.0e-04}$
WFG6	$2.03e-01_{8.9e-03}$	$2.09e-01_{4.3e-04}$	$2.08e-01_{1.3e-02}$
WFG7	$2.09e-01_{3.5e-04}$	$2.09e-01_{3.2e-04}$	$2.11e-01_{3.1e-05}$
WFG8	$1.48e-01_{2.3e-02}$	$1.48e-01_{1.0e-03}$	$1.39e-01_{2.3e-03}$
WFG9	$2.37e-01_{2.9e-03}$	$2.35e-01_{8.2e-04}$	$2.39e-01_{1.9e-03}$
DTLZ1	$4.66e-01_{1.6e-01}$	$4.94e-01_{1.9e-04}$	$0.00e+00_{0.0e+00}$
DTLZ2	$2.09e-01_{2.7e-04}$	$2.10e-01_{1.5e-04}$	$2.11e-01_{4.1e-05}$
DTLZ3	$0.00e+00_{0.0e+00}$	$2.10e-01_{6.3e-02}$	$0.00e+00_{0.0e+00}$
DTLZ4	$2.10e-01_{7.1e-04}$	$2.10e-01_{1.5e-04}$	$2.11e-01_{4.2e-05}$
DTLZ5	$2.11e-01_{3.5e-04}$	$2.12e-01_{1.3e-04}$	$2.12e-01_{4.1e-05}$
DTLZ6	$1.89e-05_{1.4e-03}$	$2.12e-01_{6.9e-05}$	$2.12e-01_{5.6e-05}$
DTLZ7	$3.29e-01_{2.8e-04}$	$3.30e-01_{9.8e-05}$	$3.30e-01_{7.3e-05}$

(b) Results obtained from [115] with irace using as objective the hypervolume indicator.

[115] Antonio J. Nebro, Manuel López-Ibáñez, Cristóbal Barba-González, and José García-Nieto. “Automatic Configuration of NSGA-II with JMetal and Irace”. In: Proceedings of the Genetic and Evolutionary Computation Conference Companion. GECCO ’19. Prague

2.c) Software capabilities

- Developed inside the jMetal family
- Implements 4 configurable algorithms
  - NSGA-II (dominance-based) - 19 components/parameters
  - MOEA/D (decomposition-based) - 25 components/parameters (Can implement the MOEAD/D-DE variant)
  - SMS-EMOA (indicator-based) - 18 components/parameters
  - MOPSO (also dominance-based) - 24 components/parameters
- Any jMetal algorithm can be used as meta-optimizer
- Available as maven project and Docker

2.c) Software capabilities

- Includes a graphical user interface
- Evaluated on a real world engineering problem

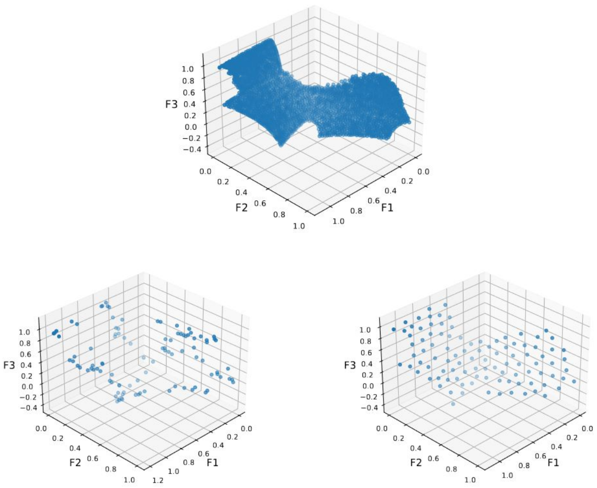
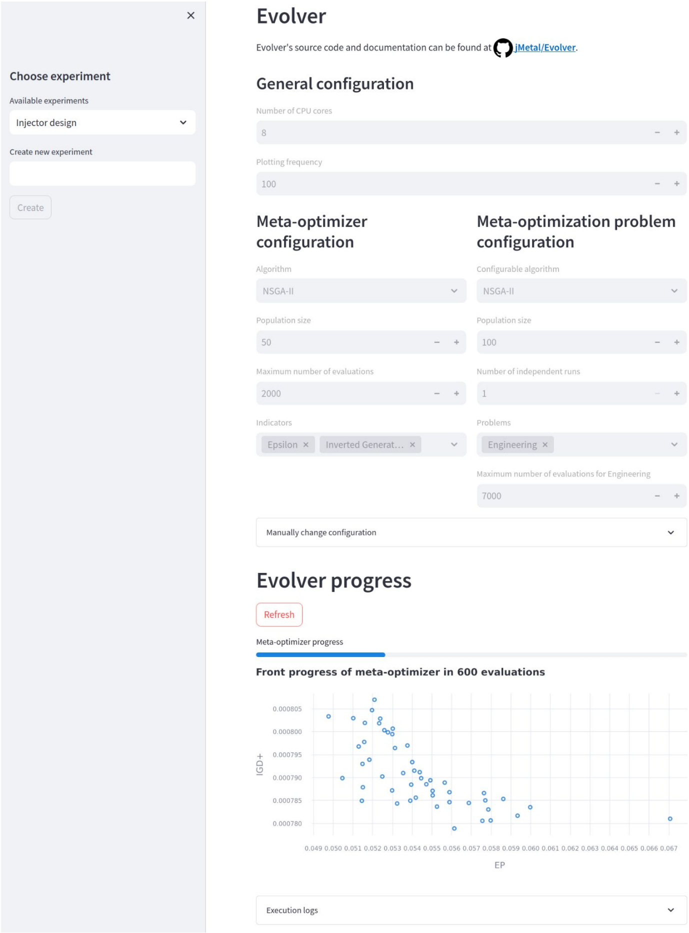


Figure 5.9: Reference front of the engineering problem (top), front obtained by NSGA-II with standard setting (bottom left), and front obtained by the auto-configured NSGA-II (bottom right).



2.c) Quality-Diversity: An alternative approach

- Ensemble theory
- Measure behavioural diversity by:
  - trajectory to the final population
  - Measured by normalized quality indicator

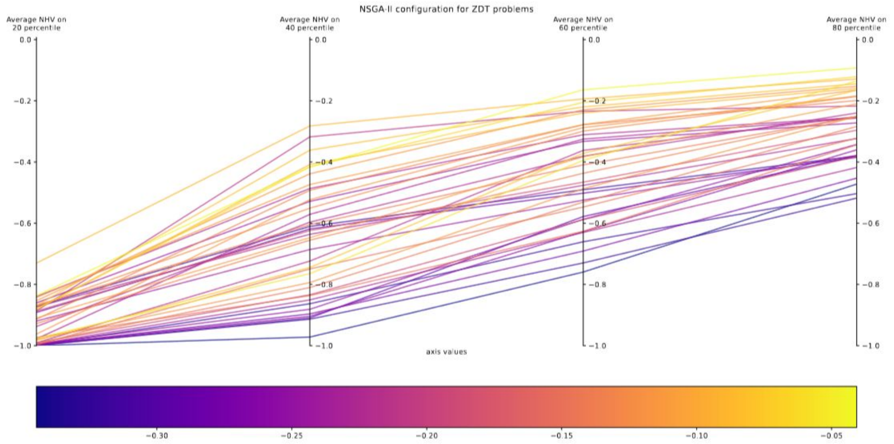


Figure 5.11: Parallel coordinates plot of the contents of a Quality-Diversity archive with 4 behavior characteristics measuring the progress of the NHV of the population of a multi-objective algorithm.

2.c) Quality-Diversity: An alternative approach

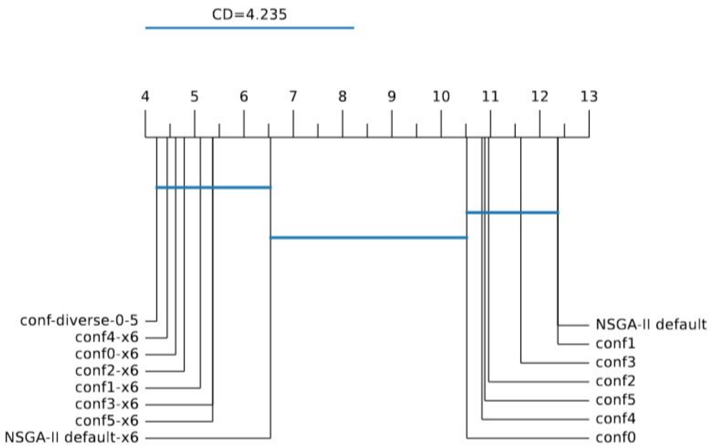


Figure 5.12: Critical distance plot ranking the obtain configurations and ensembles from the Quality-Diversity optimization process.

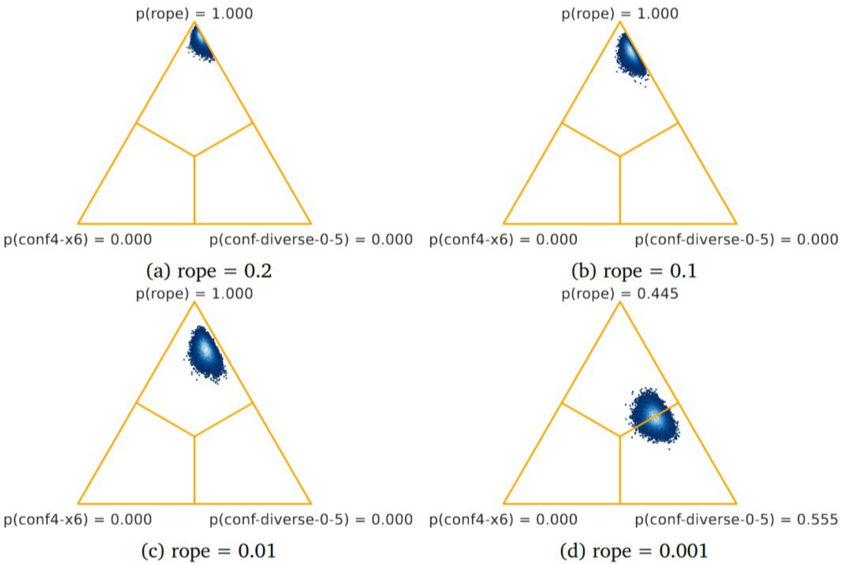


Figure 5.13: Posterior plot of the normalized hypervolume indicator using a Bayesian sign test for the top two performer ensembles with several different rope values.

2) PhD contributions

a. A Semantic Approach to Standardizing Multi-Objective Optimization
b. Similarity Between Multi-Objective Problems Via Landscape Analysis
c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
e. Algorithmic Recommendations Based on Semantic Knowledge

Leveraging Large Language Models for the Automatic Implementation of Problems in Optimization Frameworks

José F. Aldana-Martín[*,1,2][0000–0002–4845–762X], Juan J. Durillo[3][0000–0002–8023–6392], María del Mar Roldán-García[1,2][0000–0002–1470–2017], and Antonio J. Nebro[1,2][0000–0001–5580–0484]

[1] Departamento de Lenguajes y Ciencias de la Computación. University of Málaga, 29071 Málaga, Spain
[2] ITIS Software, University of Málaga, 29071, Málaga, Spain
jfaldanam@uma.es, mrgarcia@uma.es, ajnebro@uma.es
[3] Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Germany
durillo@lrz.de

Under review

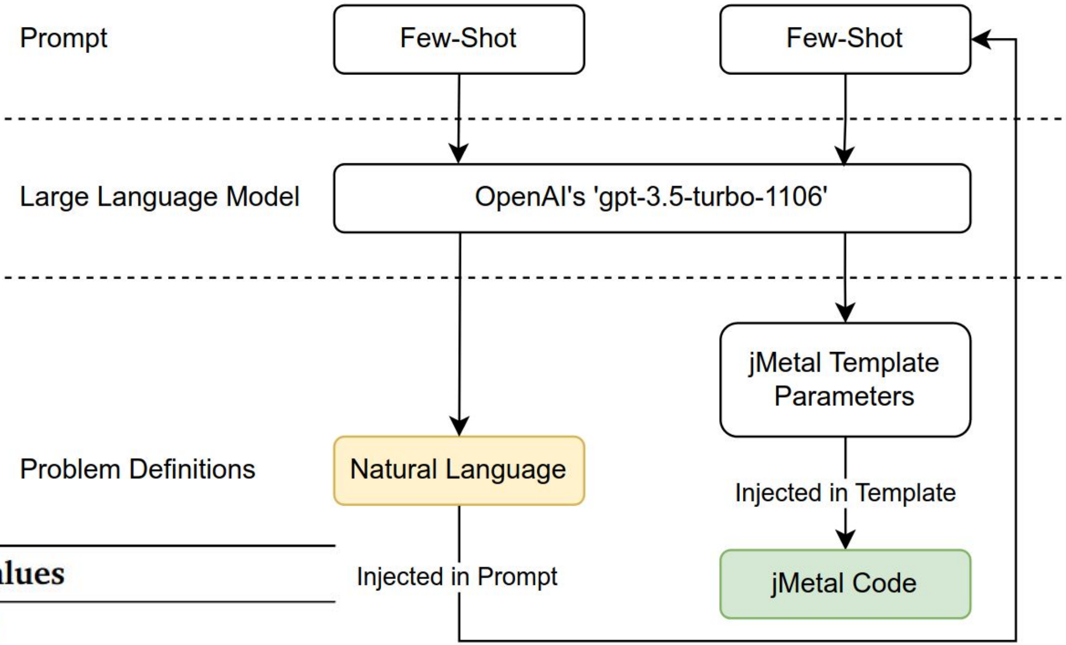
Software tool
SyntheticAI -
Synthetic generator for multi-objective problems

Software tool
moostral -
Automatic implementation of multi-objective optimization problems leveraging LLMs

2.d) Motivation

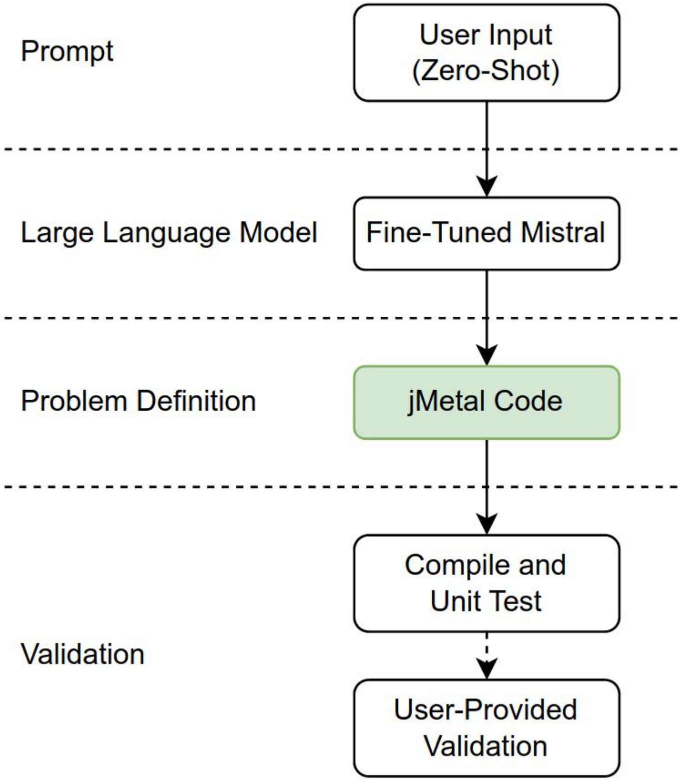
- Optimization frameworks provide software packages with everything required for working with metaheuristics.
- Domain experts often lack the technical skill to implement their optimization problems according to the specific rules of an optimization framework.

2.d) Synthetic problem generation



Parameter	Possible values
Number of objectives	[2, 5]
Number of variables	[2, 7]
Function type	Linear, quadratic, polynomial, rational, exponential, logarithmic, trigonometric or hyperbolic

2.d) Automatic problem implementation



2.d) Evaluation: On real world problems

Problem	Discipline	Source
Design of honeycomb heat sinks	Structure engineering	[146]
Vehicle crashworthiness design	Structure engineering	[92]
Development of liquid-rocket single element injector	Structure engineering	[56, 153]
Production of synthesis gas	Chemical engineering	[52]
Performance of packed bed latent heat thermal storage	Materials engineering	[53]
Ultrahigh-strength steel milling parameters	Materials engineering	[163]
Development of water and oil repellent fabric	Materials engineering	[1]
Machining of wire electrical discharge	Electronics engineering	[119]
Design of ultra-wideband antenna	Electronics engineering	[26]

Got 9 correct, failed on [53] as the formulas that define the problem are very long (>2200 tokens)

2.d) Evaluation: Natural described language

User input:

Implement a multi−objective optimization problem where the first objective is the negative square root of the first variable, a second objective that is the sum of the sine of each of the variables minus six and a half, and a third objective that is the first variable minus the third one. The range of feasible values for each variable is between $-\pi$ and π.

Showcases the strong reasoning capabilities of Large Language Models

Generated implementation:

```
package xyz.jfaldanam;

import java.util.List;
import org.uma.jmetal.problem.doubleproblem.impl.AbstractDoubleProblem;
import org.uma.jmetal.solution.doublesolution.DoubleSolution;

public class ProblemY extends AbstractDoubleProblem {

  /** Constructor */
  public ProblemY() {
    int numberOfVariables = 3;
    numberOfObjectives(3);
    numberOfConstraints(0);
    name("ProblemY");

    List<Double> lowerLimit = List.of(-Math.PI, -Math.PI, -Math.PI);
    List<Double> upperLimit = List.of(Math.PI, Math.PI, Math.PI);

    variableBounds(lowerLimit, upperLimit);
  }

  /** Evaluate() method */
  @Override
  public DoubleSolution evaluate(DoubleSolution solution) {
    double x1 = solution.variables().get(0);
    double x2 = solution.variables().get(1);
    double x3 = solution.variables().get(2);

    solution.objectives()[0] = -Math.sqrt(x1);
    solution.objectives()[1] = (Math.sin(x1) + Math.sin(x2) +
        Math.sin(x3)) - 6.5;
    solution.objectives()[2] = x1 - x3;

    return solution;
  }
}
```

2) PhD contributions

a. A Semantic Approach to Standardizing Multi-Objective Optimization
b. Similarity Between Multi-Objective Problems Via Landscape Analysis
c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
e. Algorithmic Recommendations Based on Semantic Knowledge

2) PhD contributions

a. A Semantic Approach to Standardizing Multi-Objective Optimization
b. Similarity Between Multi-Objective Problems Via Landscape Analysis
c. Automatic Generation of Quality Algorithmic Configurations for Metaheuristics
d. Leveraging Large Language Models for the Automatic Implementation of Optimization Problems
e. Algorithmic Recommendations Based on Semantic Knowledge

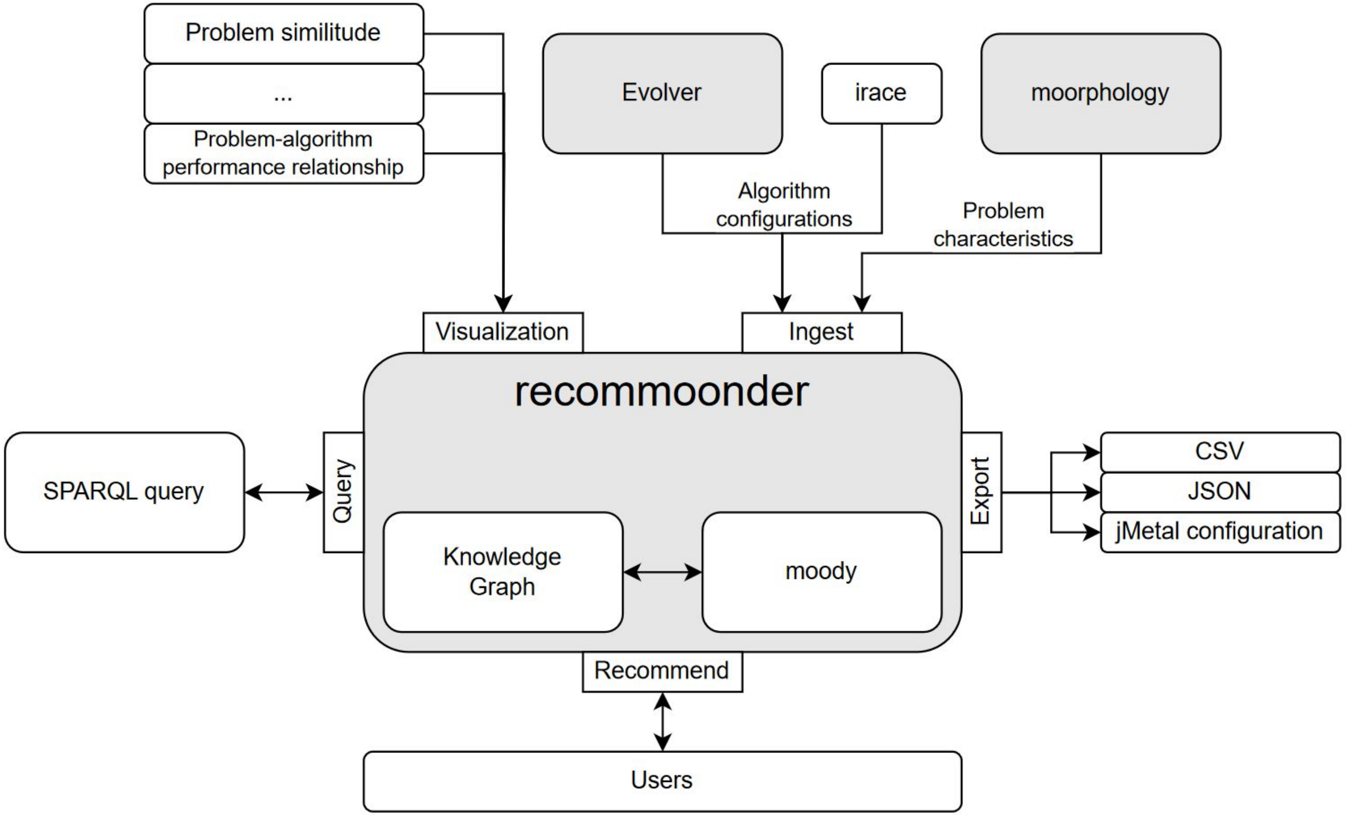
<u>Software tool</u>
recommoonder -
Algorithmic recommender for multi-objective optimization based on semantic technologies

2.e) Motivation

- Answering the main hypothesis formulated in this PhD thesis.

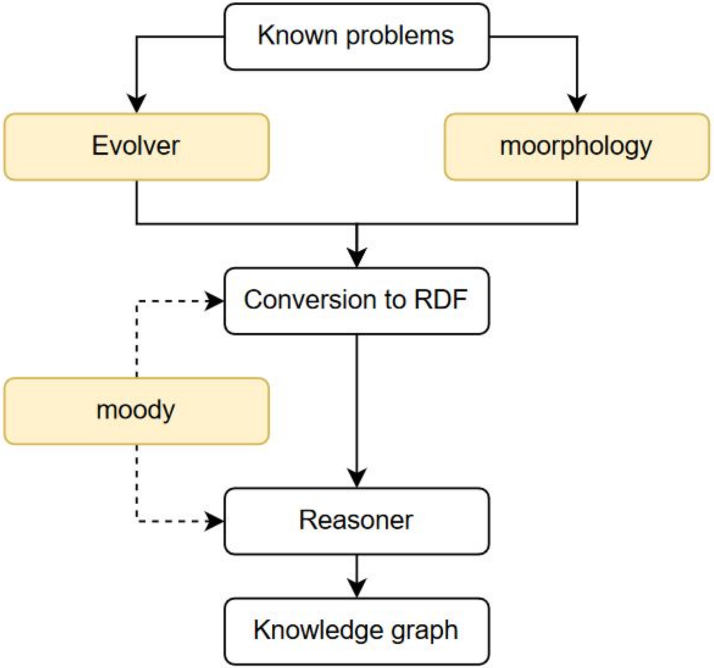
“Given previous knowledge on the relationship between a specific algorithmic configuration and the quality of the result of said algorithm solving a problem and given a similitude metric between two problems, it is possible to provide recommendations to non-expert users to choose an algorithmic configuration to efficiently solve a specific problem.”

2.e) Software architecture

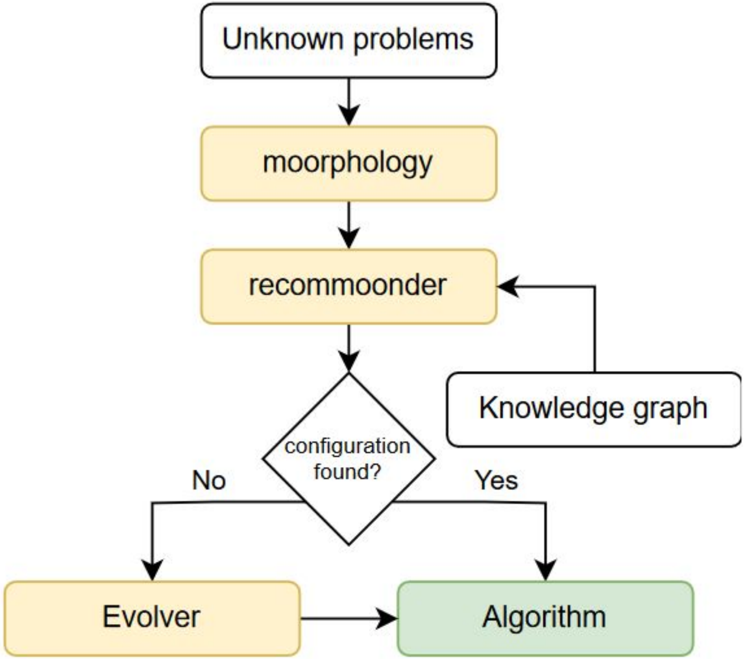


2.e) Software interfaces

Data ingestion



Algorithmic recomendation



2.e) Evaluation of the recommendation system

Evaluated on known problems

- If we ask for a previously known problem, will I get the best configuration available?

Evaluated on unknown problems

- If we ask for never seen before problems, will I get the a configuration than beats the default configuration?

2.e) Evaluation of the recommendation system

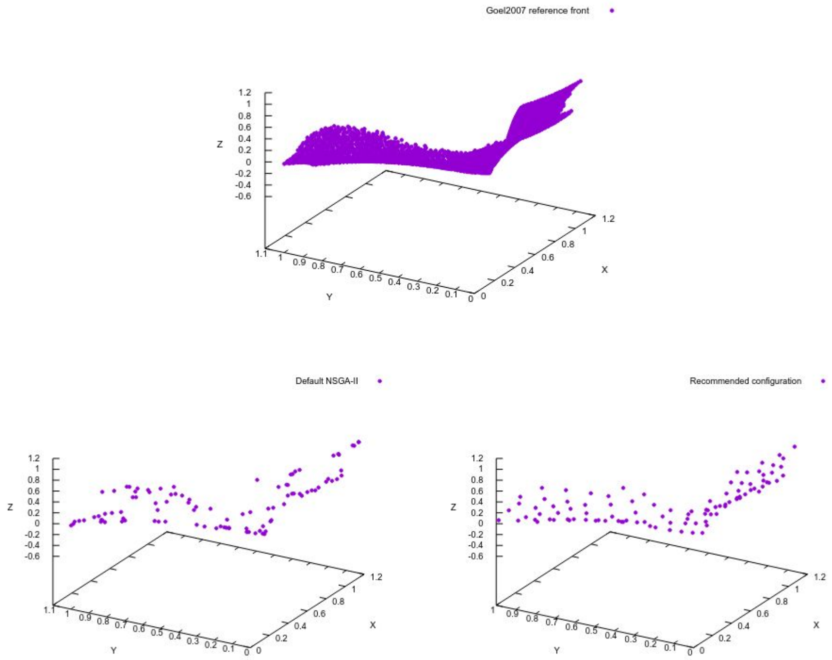


Figure 7.6: Reference front of Goel2007 (top), front obtained by NSGA-II with standard setting (bottom left), and front obtained by *recommoonder* (bottom right).

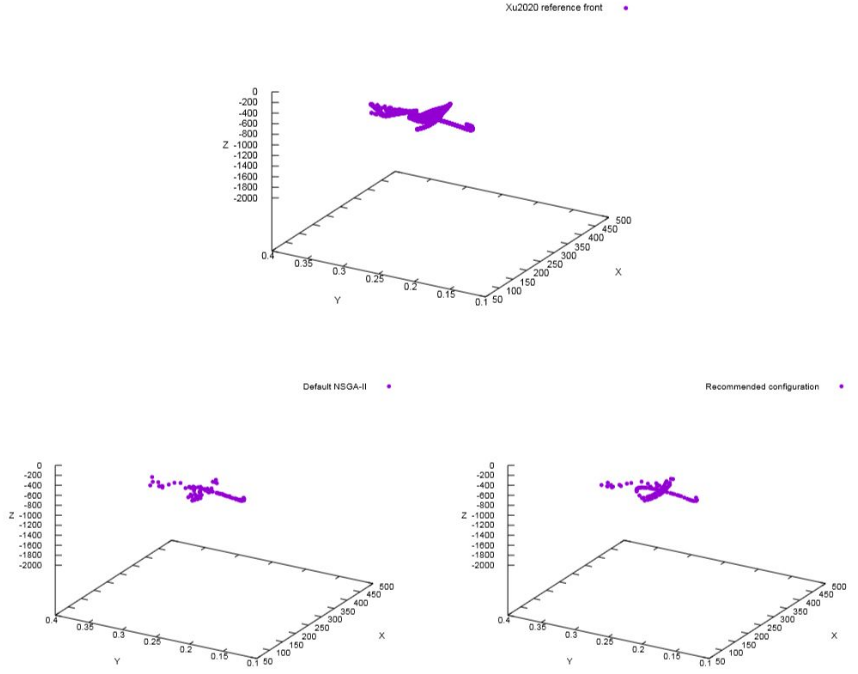


Figure 7.8: Reference front of Xu2020 (top), front obtained by NSGA-II with standard setting (bottom left), and front obtained by *recommoonder* (bottom right).

3) Conclusions and future work

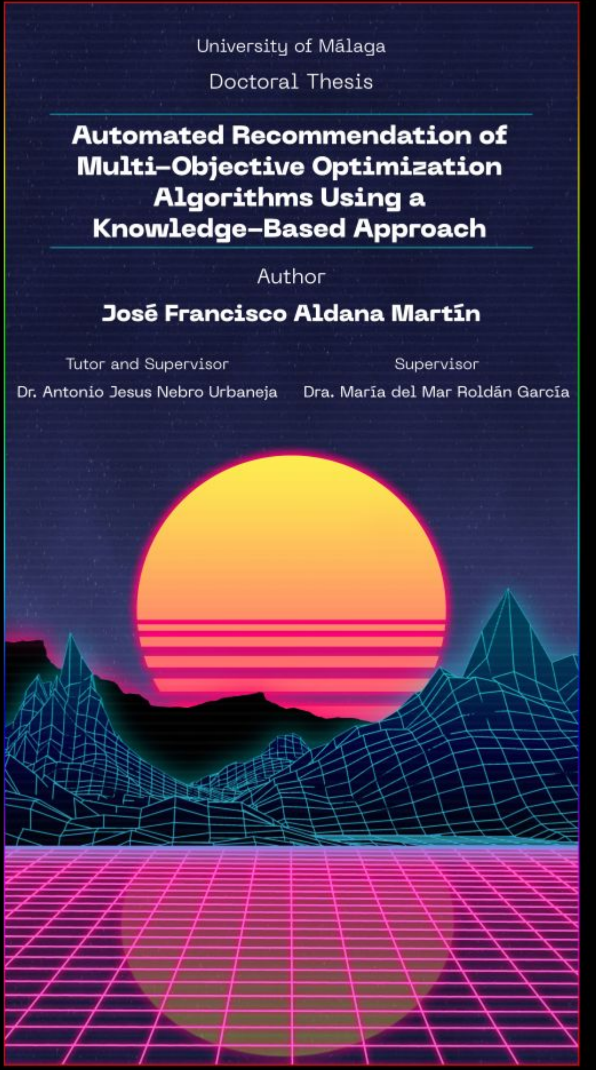
a. Conclusions
b. Future work

3.a) Conclusions and future work: Conclusions

- All 6 objectives for this PhD have been completed
- 2 papers in journal and 1 in international conference
  - 1 more paper journal article under review
- 7 open-source software packages
  - 1 open-source fine-tuned LLM
- A 3-month International research stay at the Big Data Artificial Intelligence team (BDAI) at the Leibniz Supercomputing Centre (LRZ) in Munich.
- 5+ research projects at Khaos Research
  - 3 more articles in journals and 1 book chapter

3.b) Conclusions and future work: Open research lines

- moody’s extension by including new metaheuristics and problems including new branches like discrete optimization and real world problems.
- In moorphology, an in-depth analysis on what is the optimal set of characteristic for algorithmic recommendation and how they relate to the similitude between problems.
- On the auto-configuration of algorithms, more experimentation on reducing the computational budget while still obtaining configuration with good performance when evaluated on a realistic budget.
- Adapting Evolver to support optimization problems with unknown Pareto front.
- Improving moostral by training the LLM to generate output of other popular frameworks like PlatEMO o Pagmo.
- Continuing the evaluation of Quality-Diversity optimization on real world problems.



Automated Recommendation of Multi-Objective Optimization Algorithms Using a Knowledge-Based Approach

Doctoral thesis

José Francisco Aldana Martín

Any questions?

UNIVERSIDAD DE MÁLAGA

khaos RESEARCH

SOFTWARE UNIVERSIDAD DE MÁLAGA