An Empirical Study

Meta and Hyper Heuristic Search for Multi-Objective Release Planning

Yuanyuan Zhang
CREST, UCL, UK

Mark Harman
CREST, UCL, UK

Gabriela Ochoa
University of Stirling, UK

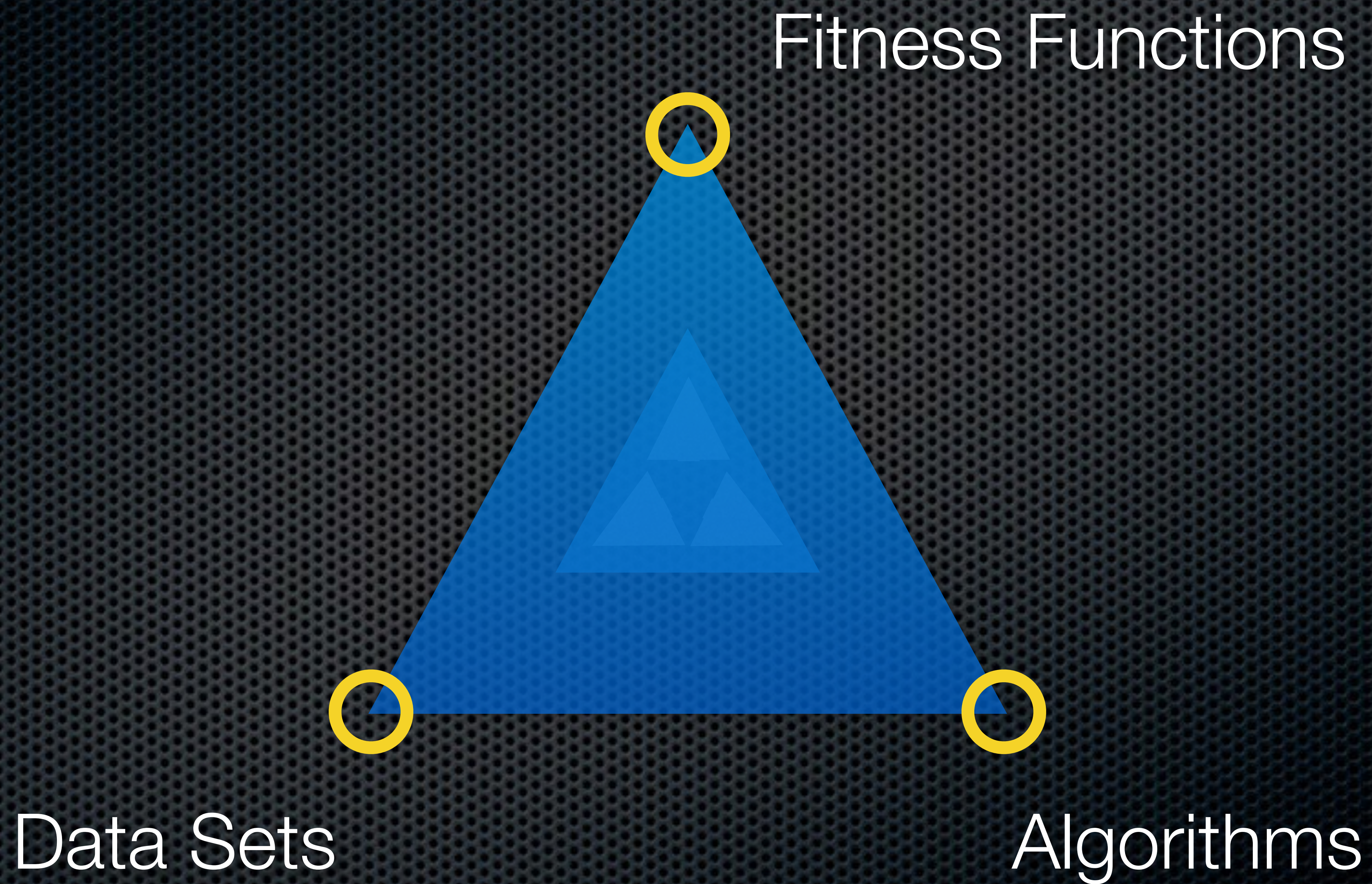
Guenther Ruhe
University of Calgary, Canada

Sjaak Brinkkemper
Utrecht University,
The Netherlands

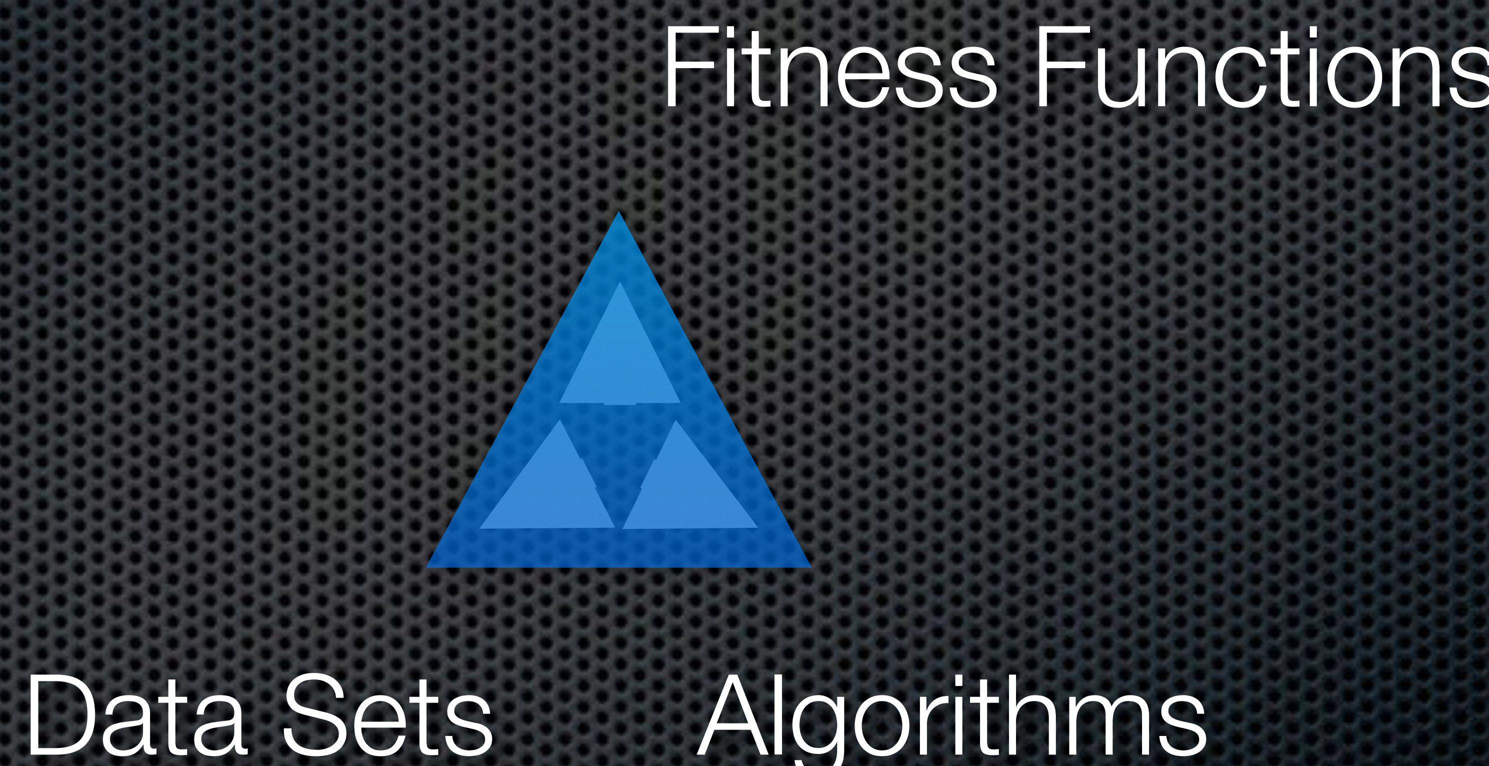
Agenda

- Contributions
- Background
- Data sets
- Fitness functions
- Algorithms
- RQs
- Results & analysis

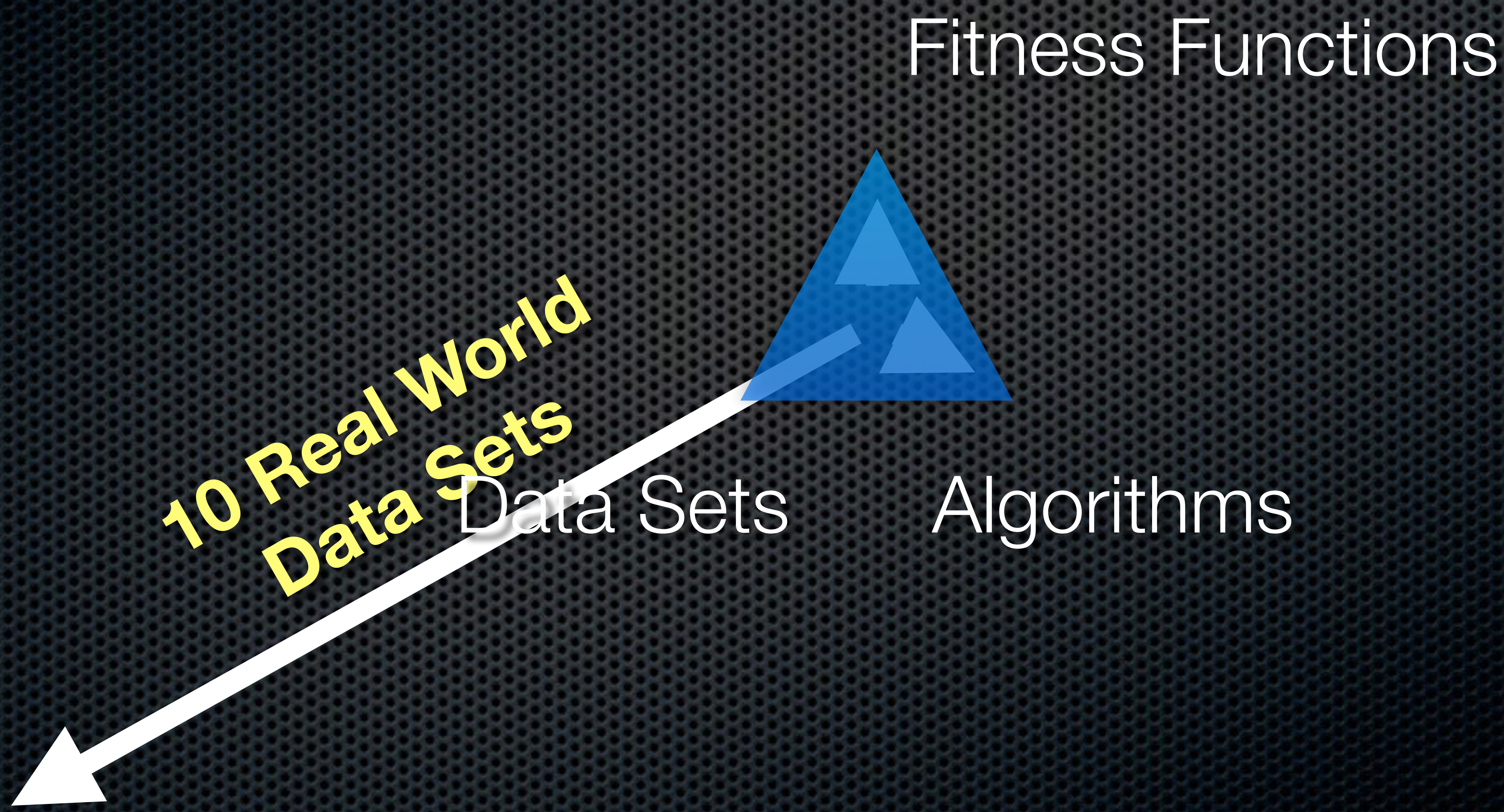
A Thorough Empirical Study



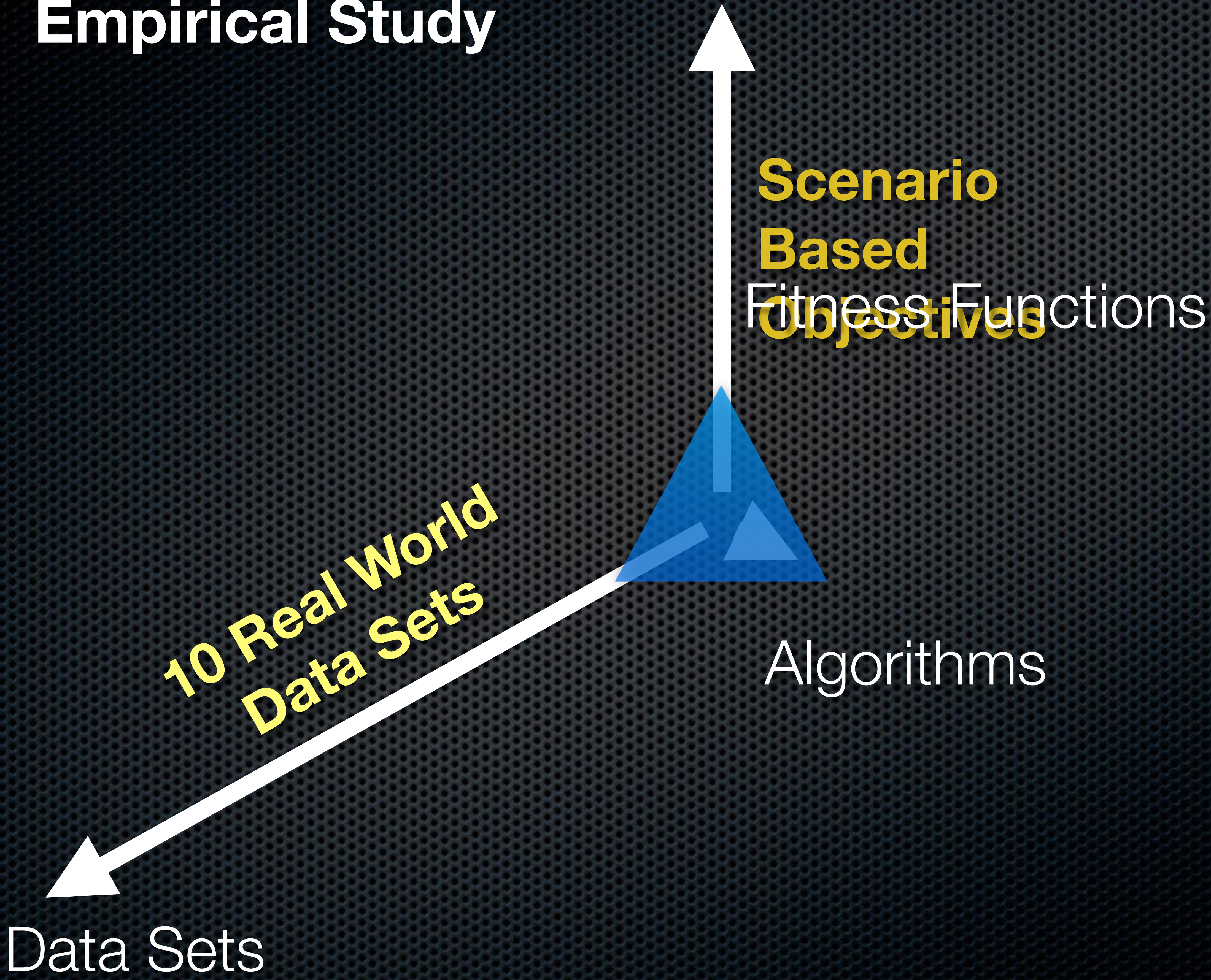
A Thorough Empirical Study



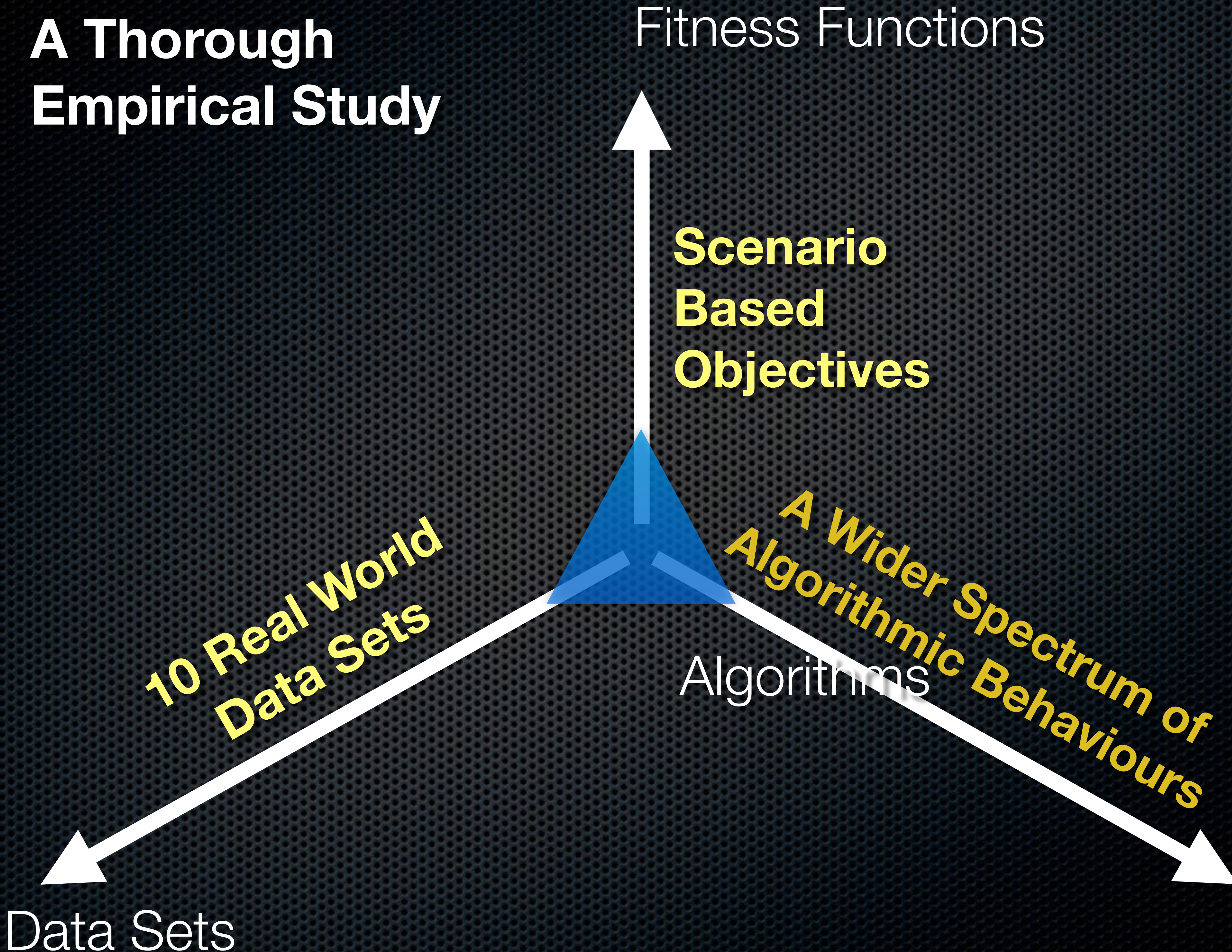
A Thorough Empirical Study



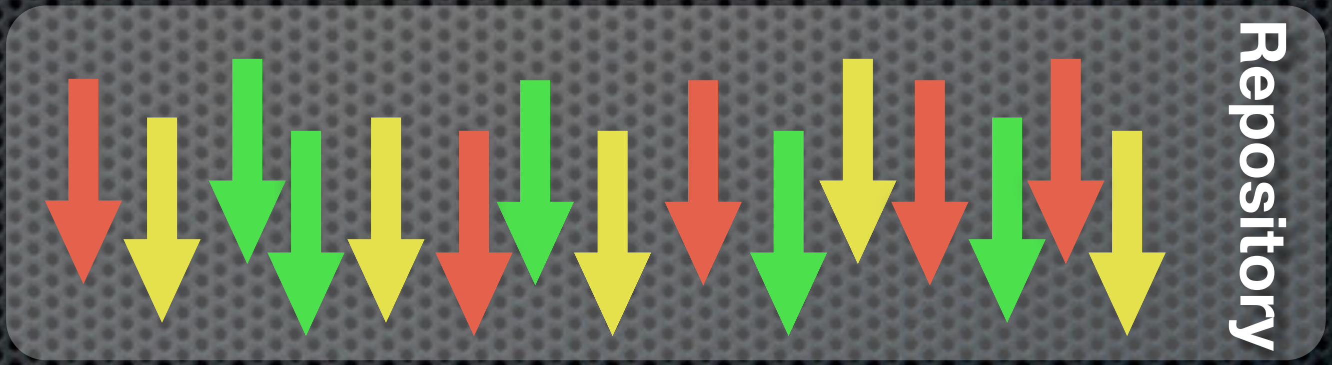
A Thorough Empirical Study



A Thorough Empirical Study

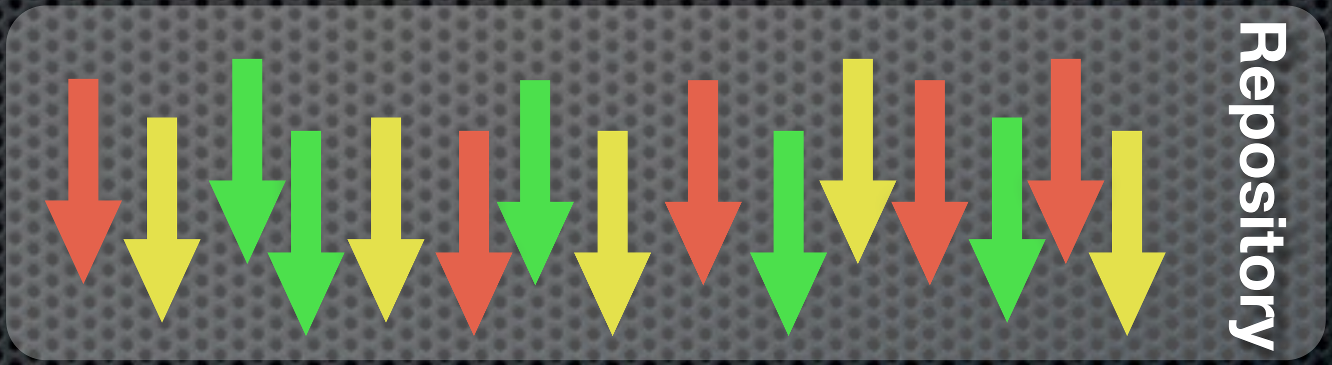


Release Planning



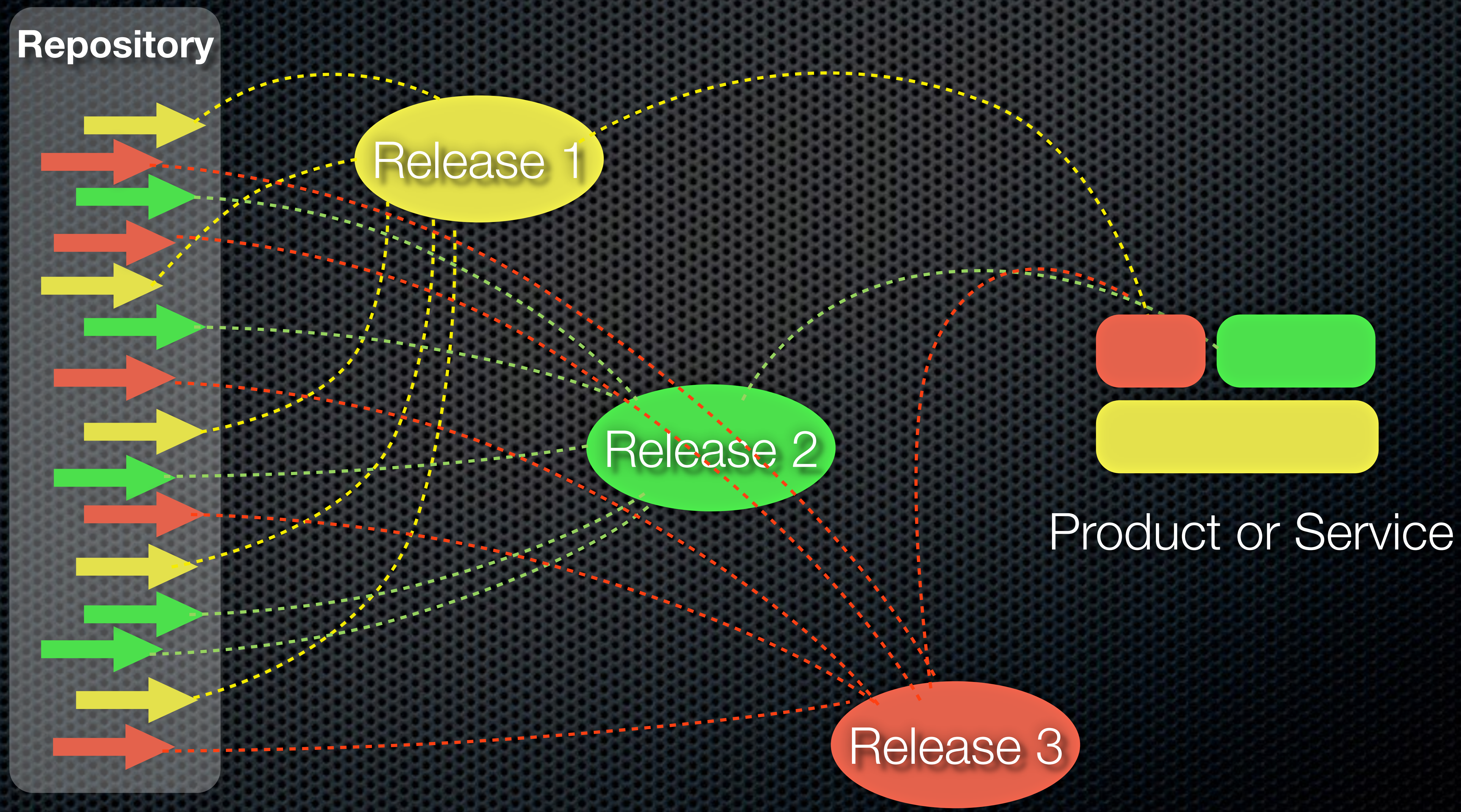
requirements and change requests

Release Planning



requirements and change requests

Release Planning



Strategic Release Planning (SRP)

- SRP is concerned with how to select and assign requirements to multiple subsequent releases.

Operational Release Planning (ORP)

- ORP deals with how to assign developers to the tasks to be performed.

Models

Stakeholders

Stakeholders Number (M)

Stakeholders Weight (W)

Models

Requirements

Cost (C)

Value (V)

Time to market (T)

Risk (R)

Frequency of use (F)

Models

Requirements

Dependence (D)

And

Or

Precedence

Value-related

Cost-related

Models

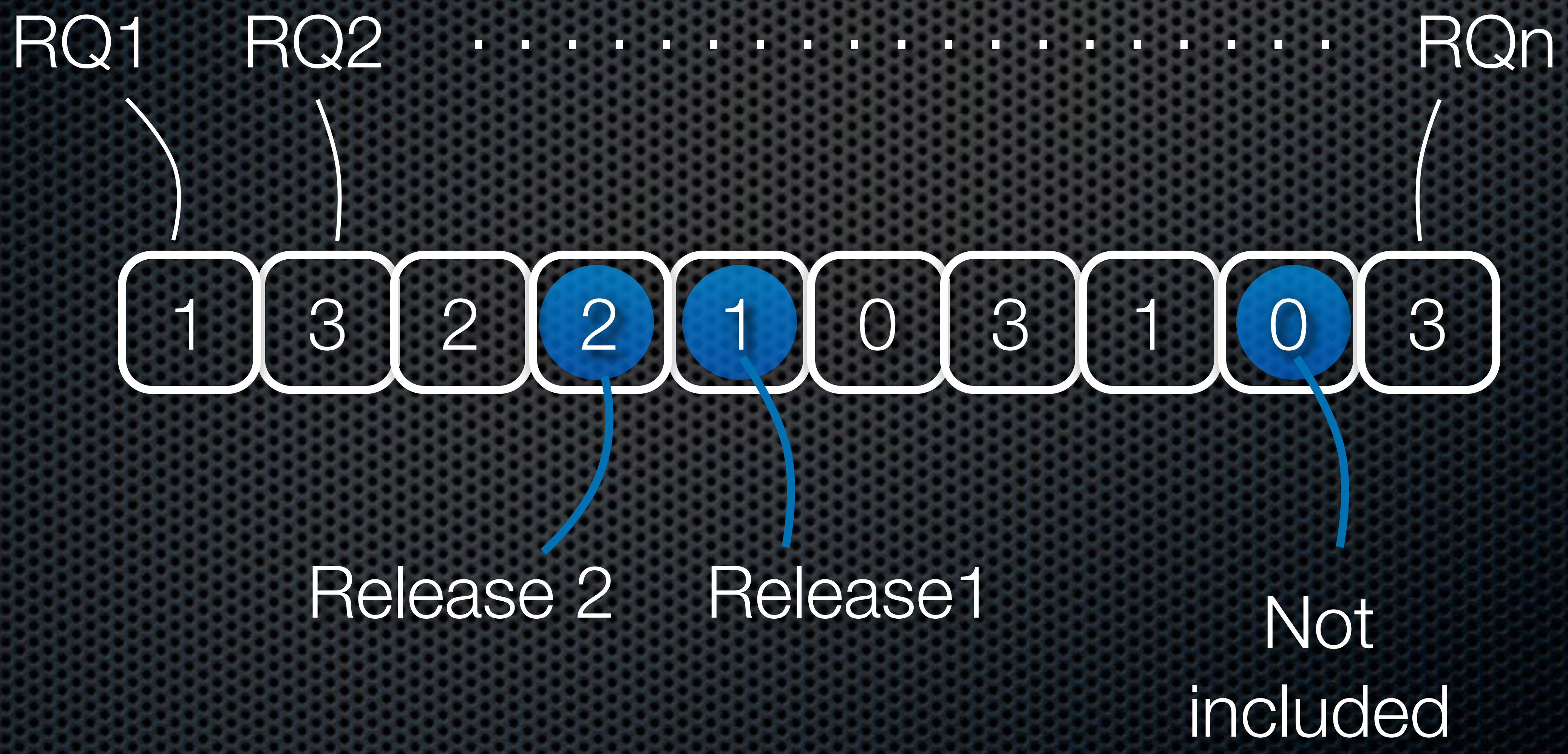
Releases

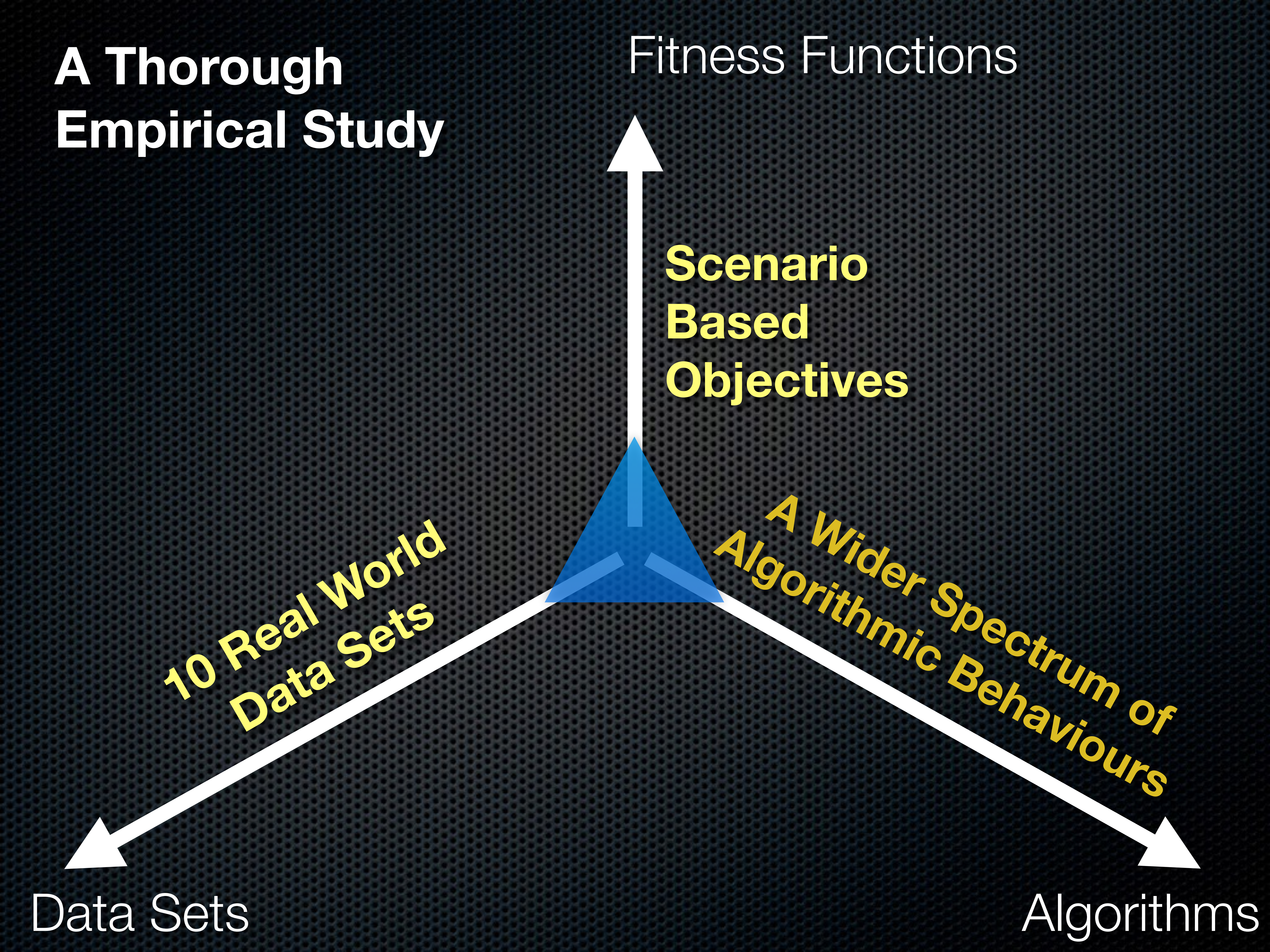
Release Number (K)

Release Importance (I)

Data Representation

A set of requirements





A Thorough
Empirical Study
Fitness Functions
Scenario
Based
Objectives
10 Real World
Data Sets
A Wider Spectrum of
Algorithmic Behaviours
Data Sets
Algorithms

10 Real World Data Sets

Data Sets	# Requirements	# Stakeholders	Objectives	
			Maximise	Minimise
Baan	100	17	Revenue	Cost
StoneGate	100	91	Sales Value	Impact
Motorola	35	4	Revenue	Cost
RalicP	143	55	Revenue	Cost
RalicR	143	79	Revenue	Cost
Ericsson	124	14	Importance for today & the future	Cost
MS Word	50	4	Revenue	Risk
Eclipse	3502	536	Importance	Cost
Mozilla	4060	768	Importance	Cost
Gnome	2690	445	Importance	Cost

Scenario-based Fitness Functions

FREQUENCY, IMPORTANCE, …

$$\text{Maximize} \quad f(\vec{x}) = \sum_{i=1}^{N} \boxed{\text{VALUE}}_{i,k} \bullet I_k$$

$$\text{Minimize} \quad f(\vec{x}) = \sum_{i=1}^{N} \boxed{\text{COST}}_{i,k}$$

IMPACT, RISK, …

A Wider Spectrum of Algorithmic Behaviours

Local	In-Between	Global
Hill Climbing	Simulated Annealing	NSGA-II

A Wider Spectrum of Algorithmic Behaviours

Local	In-Between / Meta	Global
Hill Climbing	Simulated Annealing	NSGA-II

Hyper

HHC	HSA	HNSGA-II

A Wider Spectrum of Algorithmic Behaviours

Meta-heuristics	Hyper-heuristics
Hill Climbing	HHC
Simulated Annealing	HSA
NSGA-II	HNSGA-II

Random

10 Hyper-Heuristic Operators

Ruin & Recreate

1 Random

2 Swap

3 Delete_Add

4 Delete_Add_Best

5 Delete_Worst_Add

6 Delete_Worst_Add_Best

7 Delay_Ahead

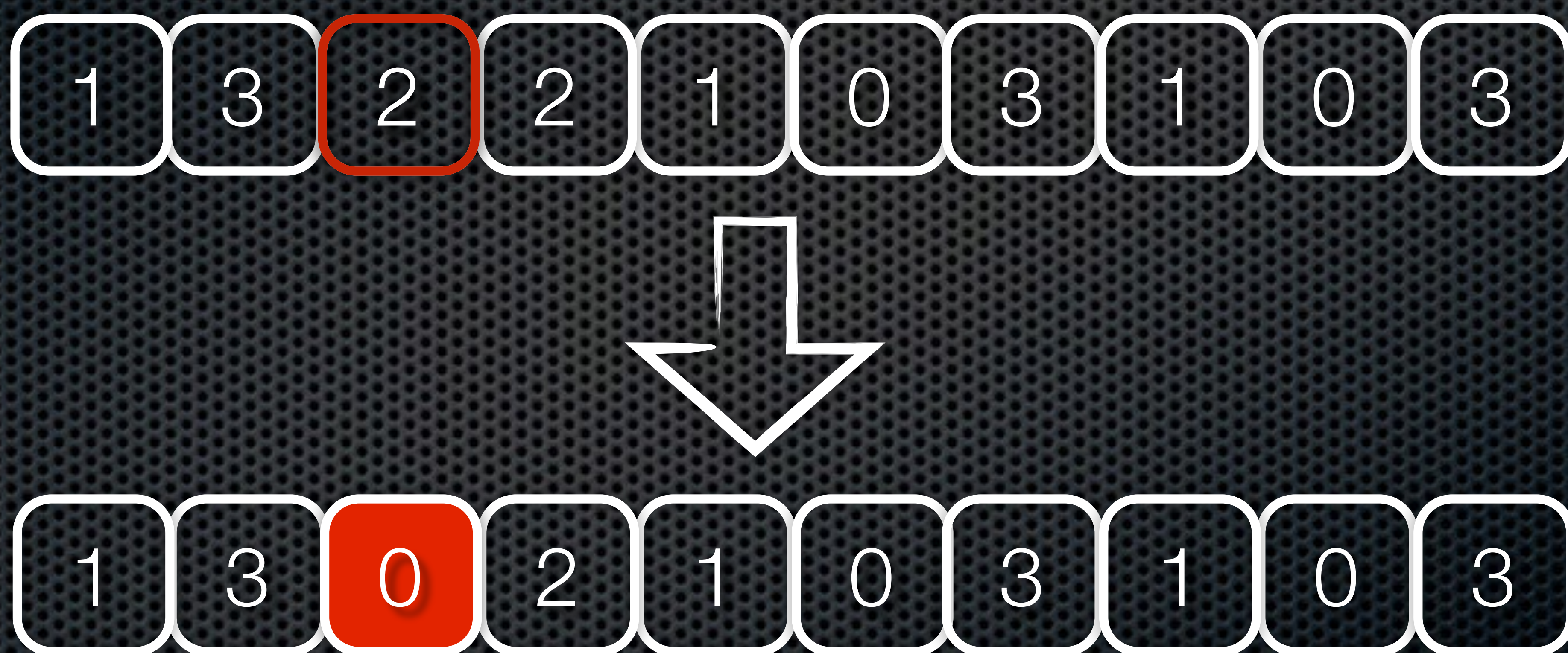
8 Delay_Ahead_Best

9 Delay_Worst_Ahead

10 Delay_Worst_Ahead_Best

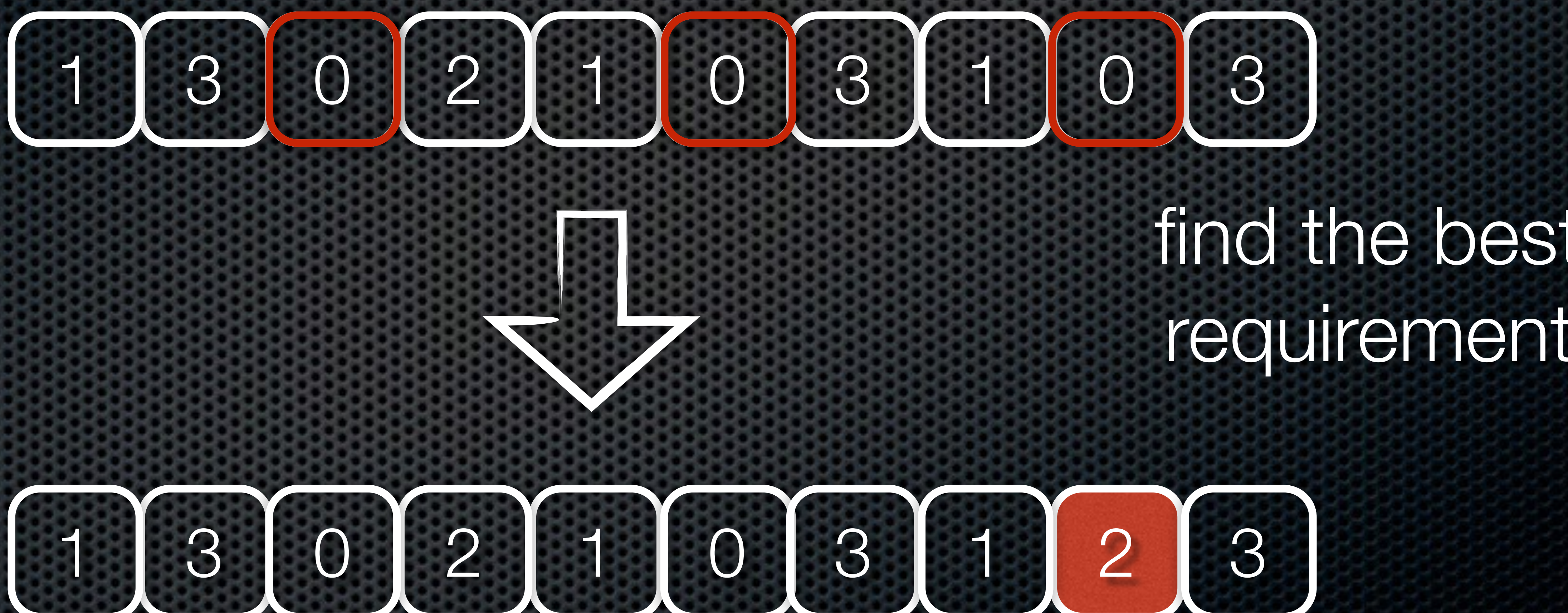
Operator: Delete_Add_Best

delete a requirement from the release with uniform probability



Operator: Delete_Add_Best

add the best requirement (based on one of fitness values) to one release



Adaptive Operator Selection

- Credit assignment
  - Extreme value credit assignment
  - Fitness improvement: hypervolume difference
  - Reference value: the fitness of the parents
- Operator selection
  - Probability matching

Performance Metrics

- Quality
  - Convergence
  - Hypervolume
  - Contribution
  - Unique Contribution
- Diversity is only interesting if the algorithm's quality is strong
- Speed

All the metrics were normalised between 0.0 and 1.0 and converted to 'Maximising metrics'.

Research Questions

RQ 1 - Quality: Which algorithm performs best?

RQ 2 - Diversity: What is the diversity of the solutions produced by each algorithm?

RQ 3 - Speed: How fast can the algorithm produce the solutions?

RQ 4 - Scalability: What is the scalability of each algorithm with regard to solution quality, diversity and speed?

Results & Analysis

RQ 1 - Quality

RQ 1 - Quality

Data Sets		Meta-heuristics			Hyper-heuristics		
		HC	SA	NSGA-II	HHC	HSA	HNSGA-II
7 smaller datasets	2 Fits						
	3 Fits						
3 larger datasets							

For the meta-heuristic algorithms, NSGA-II performs best overall for quality on smaller datasets

SA performs noticeably better on the three larger datasets

The three hyper-heuristic algorithms outperform their meta-heuristic counterparts;
HNSGA-II is beaten by its meta-heuristic counterpart only on the Ericsson dataset.

Results & Analysis

RQ 2 - Diversity

RQ 2 - Diversity

Data Sets		Random	Meta-heuristics			Hyper-heuristics		
			HC	SA	NSGA-II	HHC	HSA	HNSGA-II
7 smaller datasets	2 Fits							
	3 Fits							
3 larger datasets								

Random search perform very well, but the solutions are largely suboptimal

Of the Hyper-heuristic algorithms, HNSGA-II exhibits the best diversity

NSGA-II significantly outperforms HNSGA-II for Ericsson dataset

HNSGA-II significantly outperforms NSGA-II on Mozilla and Gnome

Results & Analysis

RQ 3 - Speed

RQ 3 - Speed

Data Sets		Random	Meta-heuristics			Hyper-heuristics		
			HC	SA	NSGA-II	HHC	HSA	HNSGA-II
7 smaller datasets	2 Fits							
	3 Fits							
3 larger datasets								

The speed of random search is worse than all other algorithms for the larger datasets

HNSGA-II is fastest overall

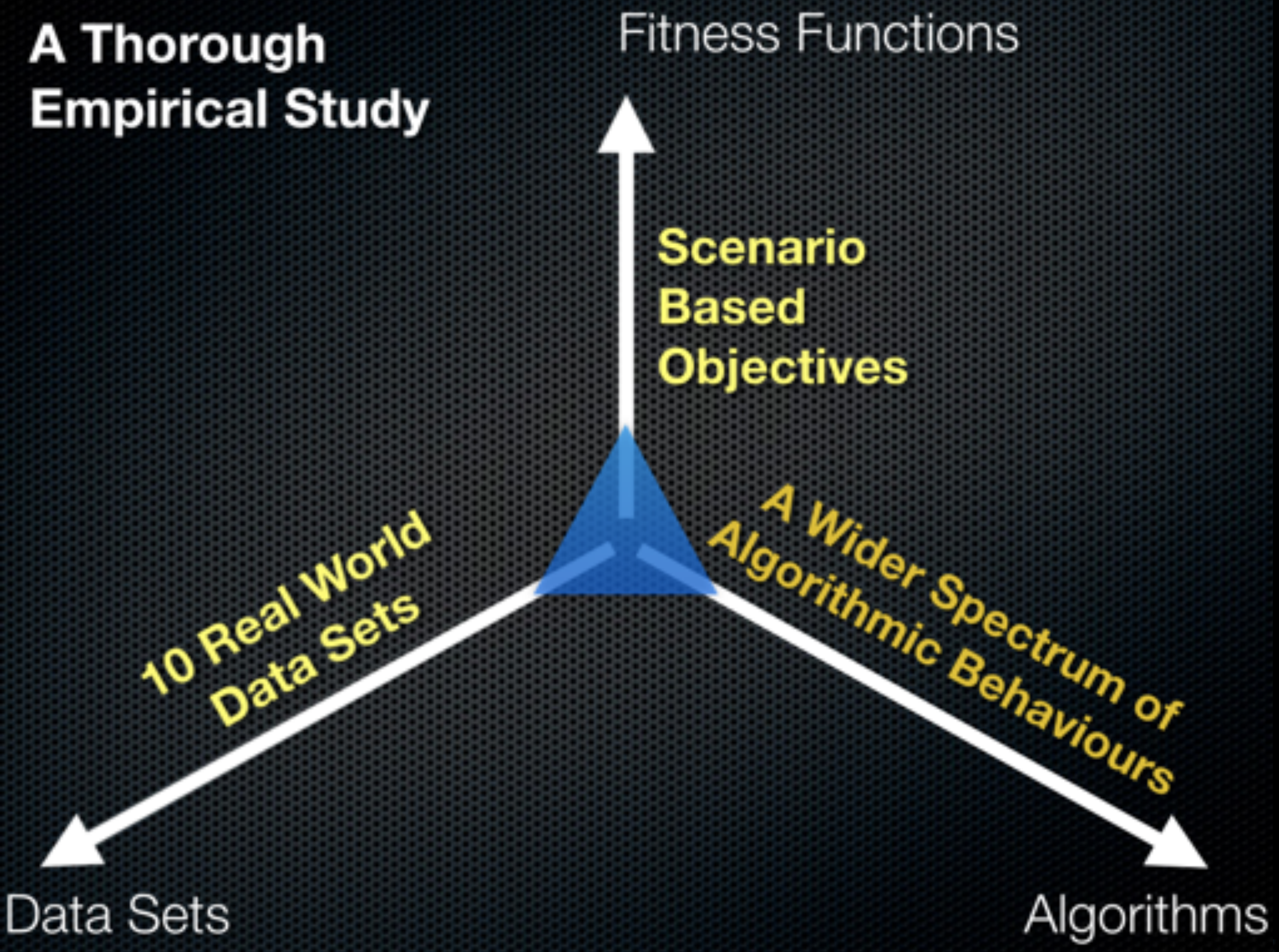
Results & Analysis

RQ 4 - Scalability

RQ 4 - Scalability

- The quality of solutions NSGA-II produced decrease as the problem size increase
  - NSGA-II's contribution to the reference front decrease, as the number of requirements increase
  - A negative correlation between the number of requirements and convergence of NSGA-II
- For the other algorithms, there is no negative correlation between problem size and solution quality
- The algorithms increase their diversity as the scale of the problem increase

A Thorough Empirical Study



Data Sets	# Requirements	# Stakeholders	Objectives	
			Maximise	Minimise
Baan	100	17	Revenue	Cost
StoneGate	100	91	Sales Value	Impact
Motorola	35	4	Revenue	Cost
RalicP	143	55	Revenue	Cost
RalicR	143	79	Revenue	Cost
Ericsson	124	14	Importance for today & the future	Cost
MS Word	50	4	Revenue	Risk
Eclipse	3502	536	Importance	Cost
Mozilla	4060	768	Importance	Cost
Gnome	2690	445	Importance	Cost

10 Hyper-Heuristic Operators

Ruin & Recreate

1 Random
2 Swap
3 Delete_Add
4 Delete_Add_Best
5 Delete_Worst_Add
6 Delete_Worst_Add_Best
7 Delay_Ahead
8 Delay_Ahead_Best
9 Delay_Worst_Ahead
10 Delay_Worst_Ahead_Best

Research Questions

RQ 1 - Quality: Which algorithm performs best?

RQ 2 - Diversity: What is the diversity of the solutions produced by each algorithm?

RQ 3 - Speed: How fast can the algorithm produce the solutions?

RQ 4 - Scalability: What is the scalability of each algorithm with regard to solution quality, diversity and speed?