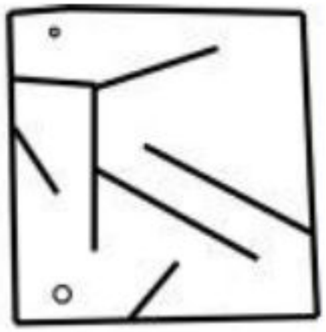
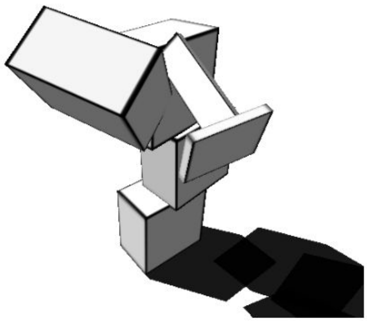
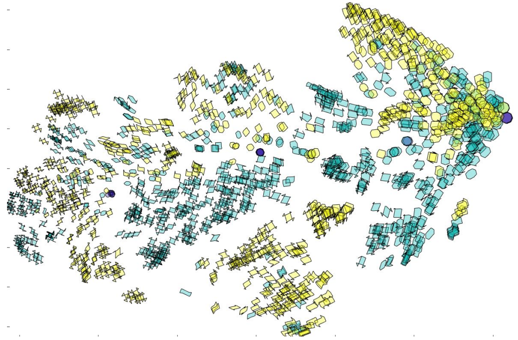
Quality Diversity

Incorporating the (extended) phenotype into optimization







Short bio

Alexander Hagg (born in Delft, now in Germany)

Bachelor CS Bonn-Rhein-Sieg Univ. of Appl. Sc.
Master Autonomous Systems (Robotics, H-BRS)
External PhD candidate (LIACS)
Institute of Technology, Resource and
Energy-efficient Engineering (TREE)

- RoboCup (DM, WM), Fraunhofer robotics
- AErOmAt project@TREE (QD for aerodynamics)
- KISs-BiS project@TREE (AI in prof. sports)
- Teaching (Genetic Algorithms, Neuroevolution, Algebra, Robotics Prog.)

Main interest: Computer Aided Ideation, Interaction in Divergent Optimization

Questions? alex@haggdesign.de

Universiteit Leiden

Hochschule Bonn-Rhein-Sieg University of Applied Sciences
Fachbereich Informatik

Institut für Technik, Ressourcen und Energieeffizienz

Initial provocation

(Hidden) assumptions in optimization:

- Solving a problem is solving its objective function
- Maximize, maximize, maximize.
- Objectives will not change

- The engineer just wants a solution
- A Pareto set is a diverse set
- We as computer scientists can solve engineering problems

- But does it capture the actual problem?

- But if X, then I am willing to...
- But they will in the real world of engineering

- They want insights too.
- Not always. Will show later.
- No. Engineers can solve them. **If** they get enough iteration time with the real world.

We can only help them gain insights

Introduction

Computers can do many things in parallel.

So instead of finding the best solutions, we should develop algorithms that

- find many (= different) good solutions
- provide insight into the problem and its solutions
- allow engineers to change their minds

-> this lecture is about

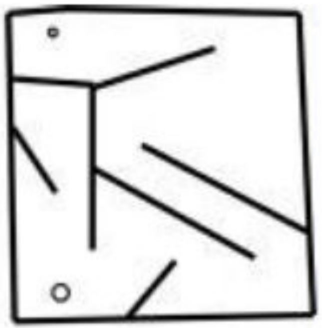
Extended phenotypes, multimodal divergent optimization -> Quality Diversity algorithms

Overview

1. Novelty Search
2. Quality Diversity
    a. First Algorithms
    b. The Extended Phenotype
    c. Archives
    d. Selection Procedure
3. MOO vs QD: a small experiment
4. Search for Features
5. Insights
6. Performance Metrics
7. Applications
8. Conclusion

1. Novelty Search

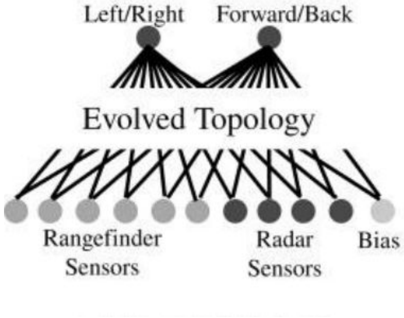
Abandon all objectives

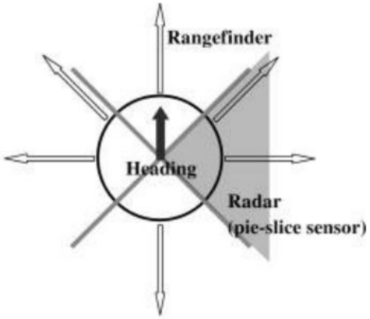


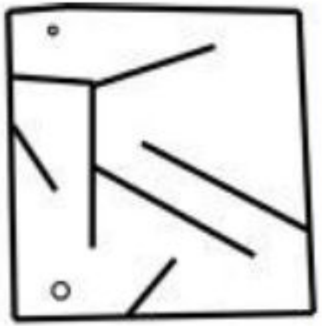
Deceptive Fitness Landscapes

Evolve neural networks that traverse the maze.

Fitness function: distance to the goal







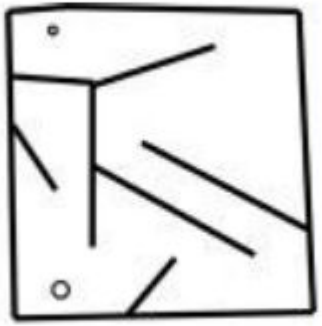
deceptive maze

Deceptive Fitness Landscapes

Goal is hard to reach due to local optimum.

The fitness landscape is **deceptive**:

- Many evolutionary paths lead to local optimum
- In order to reach the global optimum, a fitness *valley* has to be crossed



deceptive maze

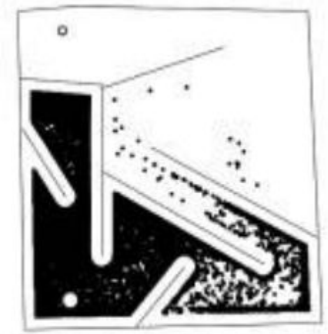
Deceptive Fitness Landscapes

Objective based search (NEAT[1]) runs into local optimum.

Even though NEAT uses *speciation* for diversity maintenance

Intuition: controllers are needed that might perform worse, but choose different paths.

-> **behavioral diversity**



Objective based search

[1] Stanley, K. O., & Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. Evolutionary Computation.

Abandoning Objectives

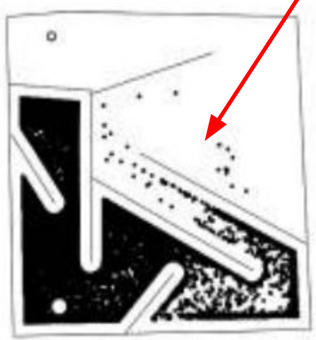
Replace fitness function with **novelty metric**[1].

Behavior characterization (BC): position at the end of robot trajectory

Novelty[2] = sparsity **ρ** in BC space, k nearest neighbors

$$\rho(x) = \frac{1}{k} \sum_{i=0}^{k} \mathrm{dist}(x, \mu_i),$$

high sparsity



Objective based search

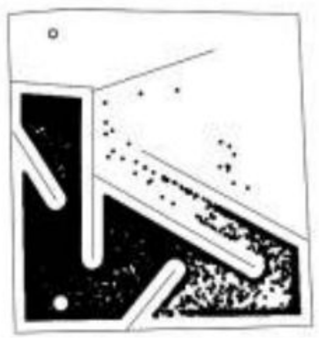
[1] Lehman, J., & Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. Evolutionary Computation.

[2] *Novelty* might be seen as a misnomer, as we are able to measure how "novel" a solution is. If we are able to measure it, is it really novel?

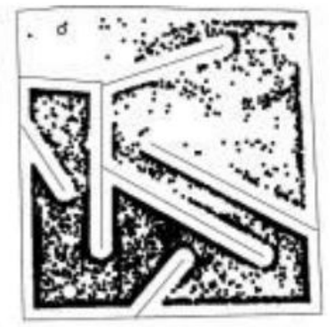
Abandoning Objectives

Solutions are much more diverse

Question: what do you think might happen when we remove the borders of the maze? Consequences for novelty based search?



Objective based search



Novelty based search

Novelty Search

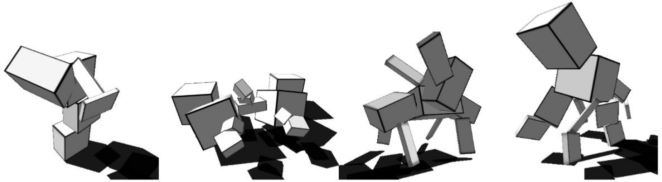
Apply to

- deceptive fitness landscapes
- multimodal landscapes
- intuitive behavioral characterizations
- and domain constraints on possible expressible behaviors

Lehman, J., & Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. Evolutionary Computation.

2. Quality Diversity

*“But we **are** interested in optimality!”*

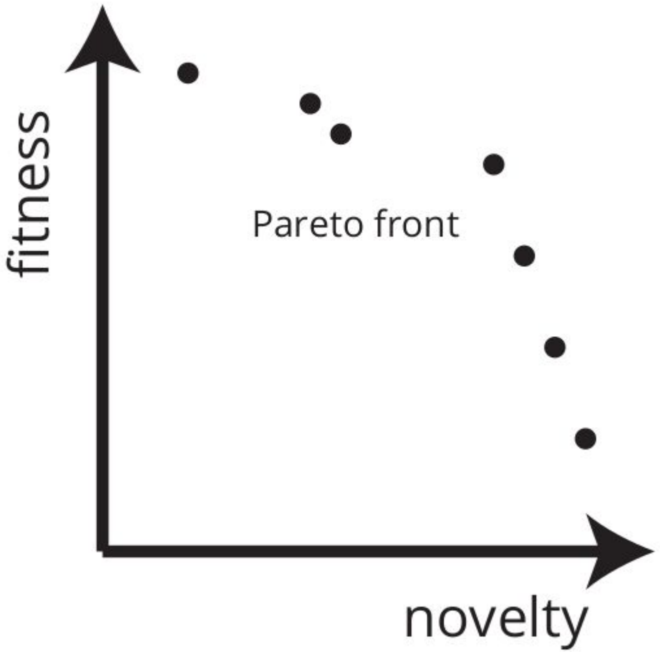


(a) Hopper (b) Crab (c) Quadruped (d) Tailed Quadruped

Reintroducing Fitness

Combine fitness and novelty using multiobjective optimization
-> Novelty Search with Global Competition

fails to exploit the fact that **some niches may naturally support different levels of fitness than others**.



Reintroducing Fitness

Novelty Search with Local Competition (NSLC) [1]

Solutions are only compared in their behavioral niche[2].

Combine novelty and **local** fitness criteria (# of neighbors with lower fitness)

[1] Lehman, J., & Stanley, K. O. (2011). Evolving a diversity of virtual creatures through novelty search and local competition. GECCO.
[2] niching itself is not new of course

Quality Diversity

-

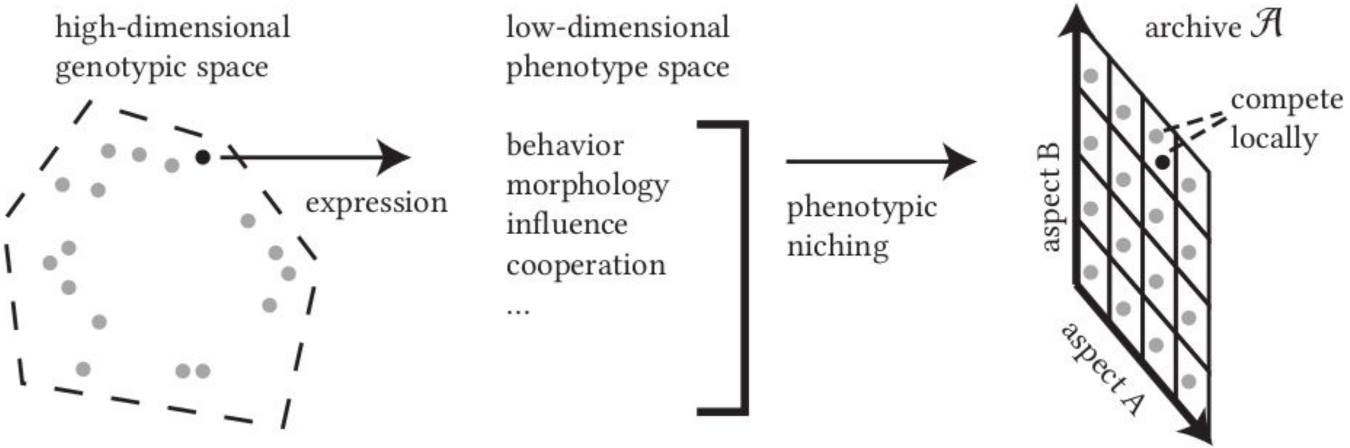


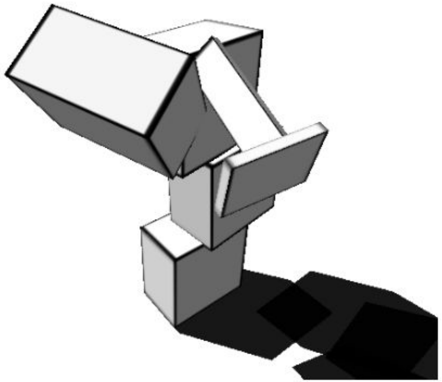
Fig. 2 Phenotypic aspects of solutions genotypic space determine similarity and niche assignment. Candidate solutions only compete within the phenotypic niches they occupy.

Hagg, A. (2019). Discovering Modes using Quality Diversity Algorithms. (tbp)

2a. First algorithms

Novelty Search with Local Competition

Fitness: distance travelled



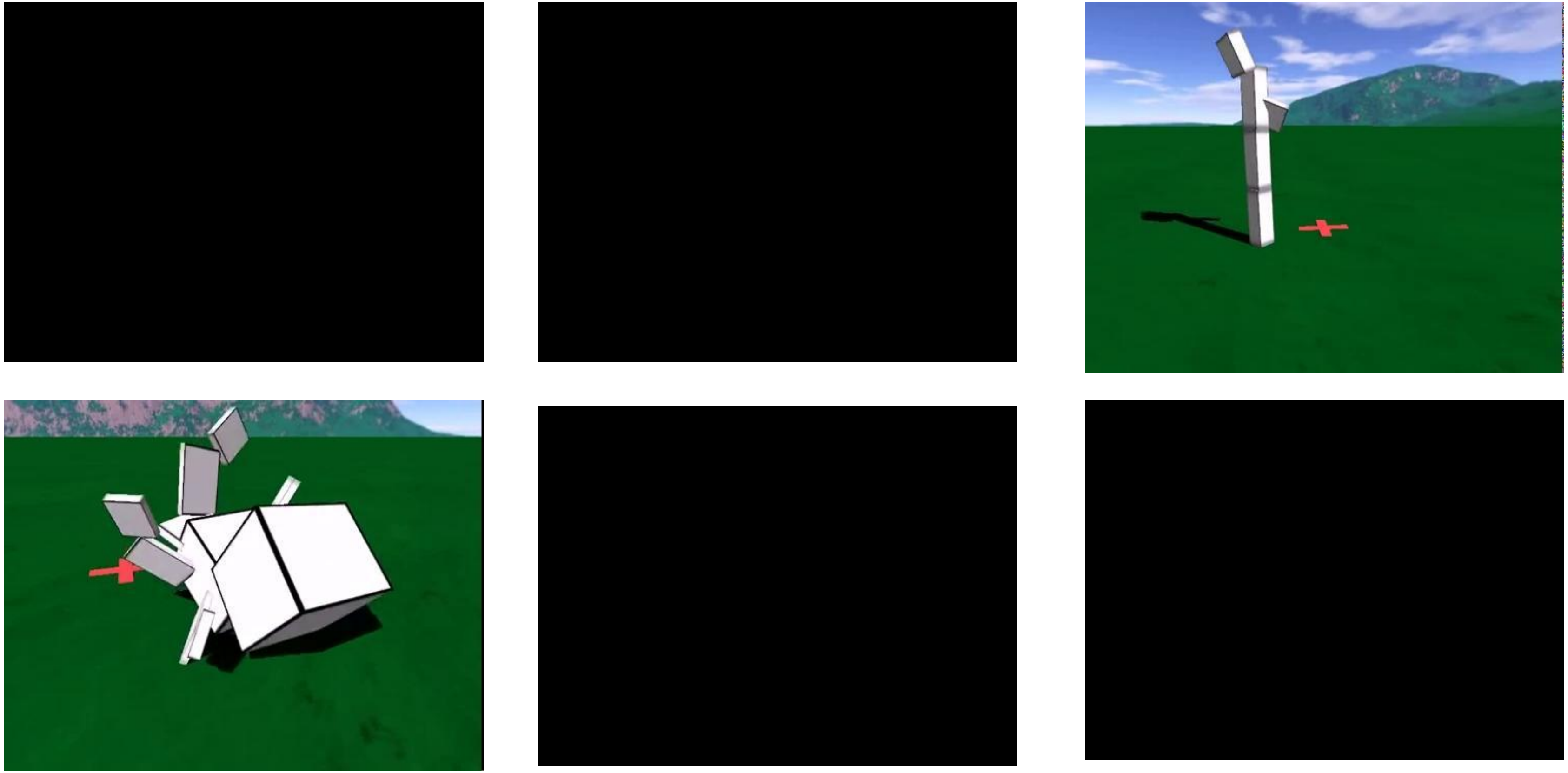
Behavior characterization (3 dimensions):

- height
- mass
- # active joints

Lehman, J., & Stanley, K. O. (2011). Evolving a diversity of virtual creatures through novelty search and local competition. GECCO.

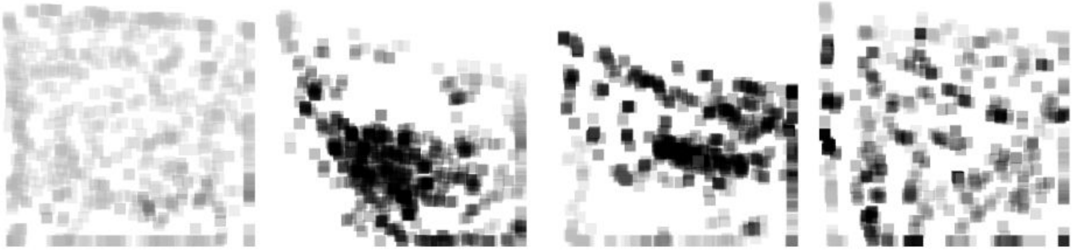
[1] niching itself is not new of course

Diversity of Virtual Creatures

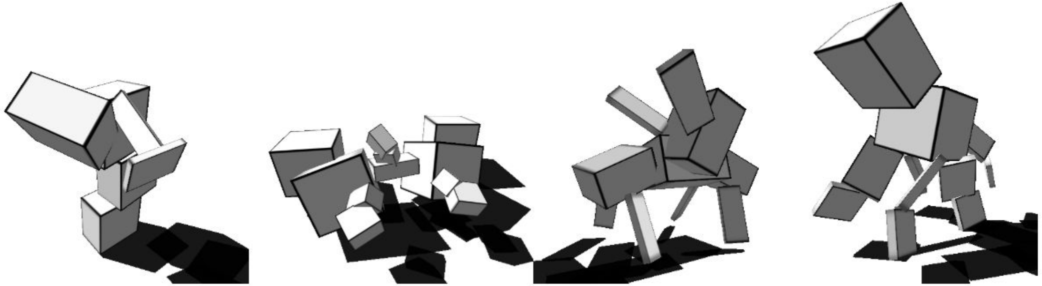


Lehman, J., & Stanley, K. O. (2011). Evolving a diversity of virtual creatures through novelty search and local competition. GECCO.

I. Novelty Search with Local Competition



(a) Novelty Only (b) Fitness Only (c) Global Competition (d) Local Competition



(a) Hopper (b) Crab (c) Quadruped (d) Tailed Quadruped

Figure 7: **Diverse competent morphologies discovered within a typical *single* run of local competition.** Various creatures are shown that have specialized to effectively exploit particular niches of morphology space. These creatures were all found in the final population of a typical run of local competition. The hopper (a) is a unipedal hopper that is very tall, (b) is a heavy short crab-like creature, and (c) and (d) are distinct quadrupeds. Creature (c) drives a large protrusion on its back to generate momentum, and (d) has a tail for balance.

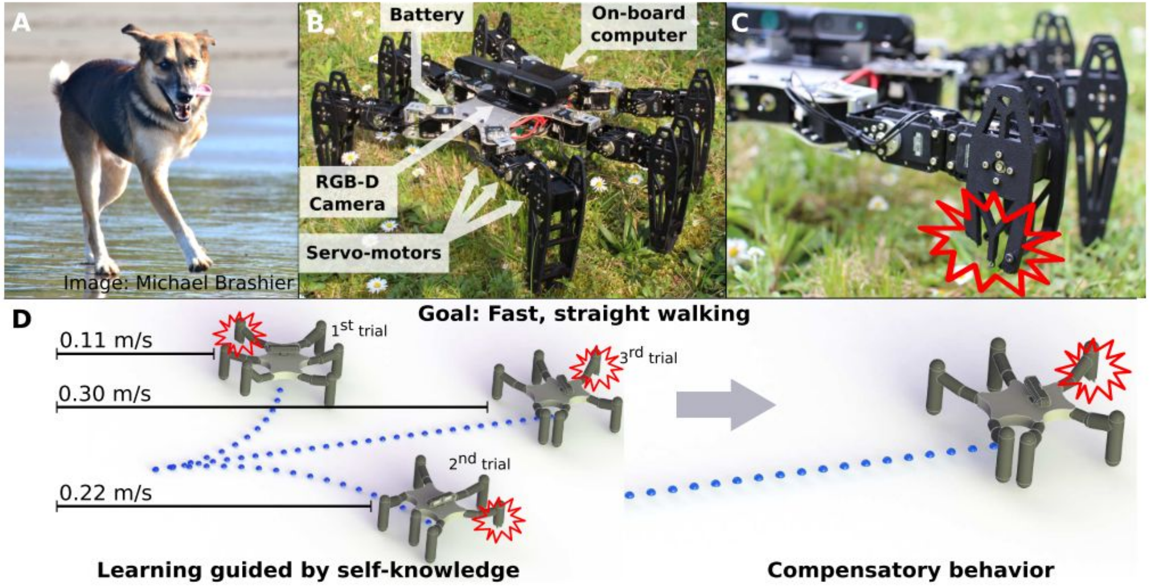
Lehman, J., & Stanley, K. O. (2011). Evolving a diversity of virtual creatures through novelty search and local competition. GECCO.

Multidimensional Archive of Phenotypic Elites

Learn an archive of walking gaits to allow adaptation to damage.

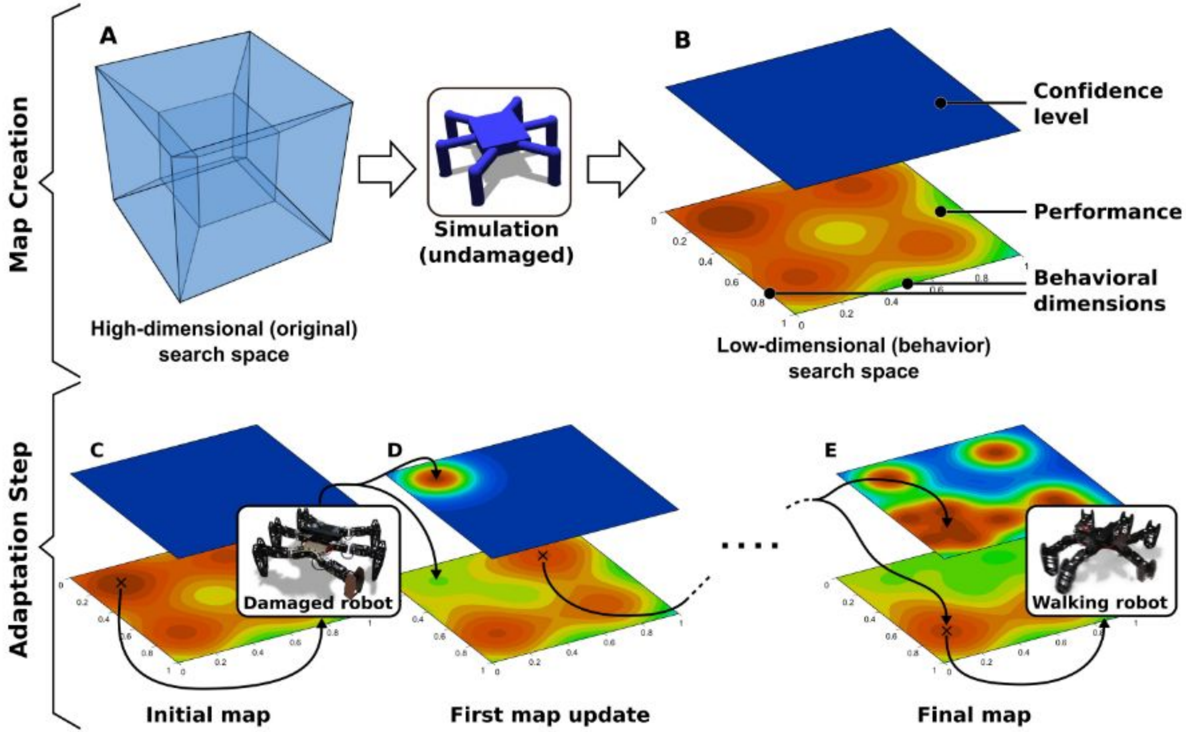
-

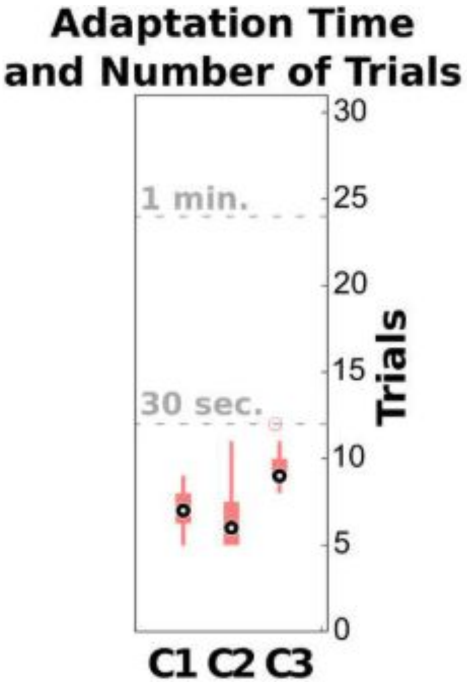
Phenotypic aspect:
number of legs used.



Cully, A., Clune, J., Tarapore, D., & Mouret, J. B. (2015). Robots that can adapt like animals. Nature.

II. Multidimensional Archive of Phenotypic Elites





Cully, A., Clune, J., Tarapore, D., & Mouret, J. B. (2015). Robots that can adapt like animals. Nature.

Quality Diversity Algorithms

-

Algorithm 1 Quality Diversity

Define and formalize phenotypic descriptors
Initialize population
Initialize archive $\mathcal{A}$
for iter = 1 $\rightarrow$ generations budget **do**
    Select parents to form new offspring population-based on scoring scheme
    Evaluate performance and phenotypic descriptors of offspring
    Add individuals (potentially) to niches in archive $\mathcal{A}$
    Update selection scores
end for

2b. The Extended Phenotype

A Conversation about the Extended Phenotype

The most important kind of replicator is the gene ... Replicators are not, of course, selected directly, but by proxy; they are judged by their phenotypic effects
… the replicator should be thought of as having **extended phenotypic effects, consisting of all its effects on the world at large**, not just its effects on the individual body in which it happens to be sitting.

(Dawkins 1982)

"There is a power and utility to regarding the gene as the unit of selection, but equally there is value to **seeing the organism as the unit of niche construction.**"

(Laland 2004)

"**The beaver's dam is as much an adaptation as the beaver's tail**. In neither case have we done the necessary research to show that it results from gene selection. In both, we have strong plausibility grounds to think it is."

(Dawkins 2012)

Dawkins, R. (1982). The Extended Phenotype.
Laland, K. N. (2004). Extending the Extended Phenotype,
Dawkins, R. (2012). Extended Phenotype - But Not Too Extended. A Reply to Laland, Turner and Jablonka.

The Extended Phenotype

-

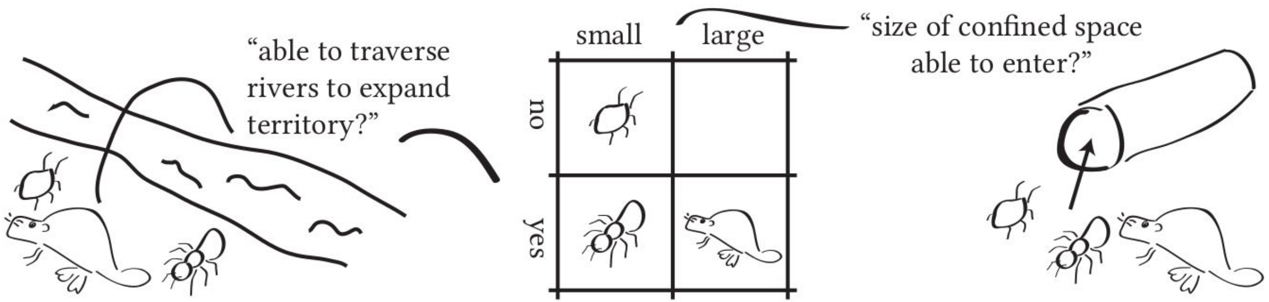


Fig. 1 Some phenotypic aspects, like the ability to traverse a river, put the fire ant and the beaver into the same phenotypic niche. Aspects like body size allows separation between the two species into separate niches. Stink bugs however can enter similar confined spaces like ants do, but they cannot swim.

Dawkins, R. (1982). The Extended Phenotype.
Hagg, A. (2019). Discovering Modes using Quality Diversity Algorithms. (tbp)

“Behavior Characterization”

In NS/NSLC we saw the following BC dimensions:

- Final position of a robot
- Height of virtual creature
- Mass of virtual creature
- Number of active joints of virtual creature

- Result of a behavior
- Creature morphology
- Creature morphology
- Creature morphology

“BC” can be seen as misnomer, as it does not only encompass behavioral aspects.
Misnomer due to the application to virtual creatures.

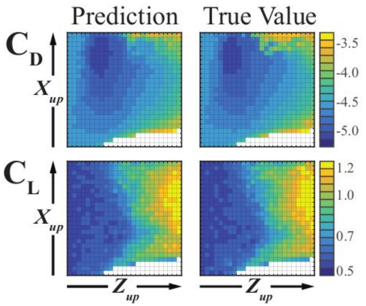
Behaviors are not the only part of a phenotype

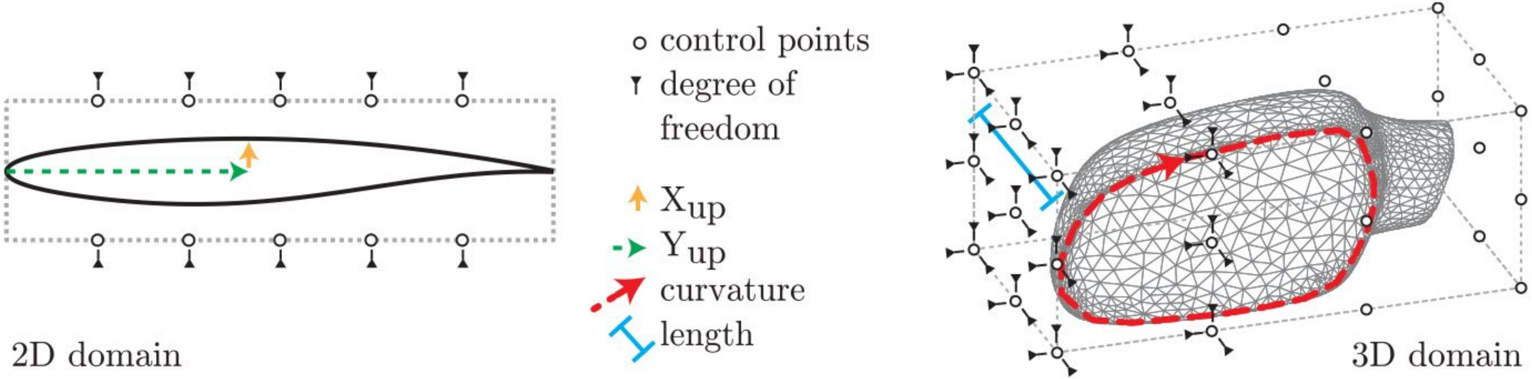
Morphology, behavior, influence, cooperation, extend the phenotype as far as necessary and appropriate for application.

Dawkins, R. (1982). The Extended Phenotype.

Morphological Features

Other examples of phenotypic features -

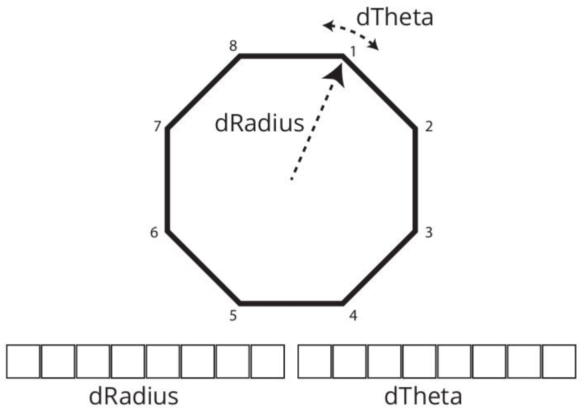


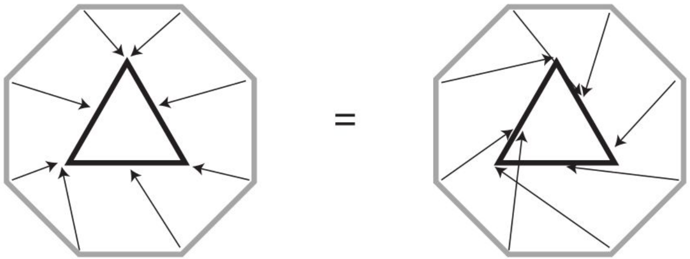


Gaier, A., Asteroth, A., & Mouret, J. B. (2017). Feature space modeling through surrogate illumination. GECCO.
Hagg, A., Asteroth, A., & Bäck, T. (2018). Prototype Discovery using Quality-Diversity. PPSN.

Difference Genotypic/Phenotypic Niching

- Depends on encoding (free form deformation vs direct encoding of polygon)
- Genetic neutrality: triangles versus rotated triangles
- So: are we interested in putting these triangles into their own niche? I think not. The engineer thinks not.
- Possibly take this further one step: “morphological” neutrality: different shapes can cause similar flow (-> idea: perform niching on the flow characteristics)





2c. Archives

Archive types

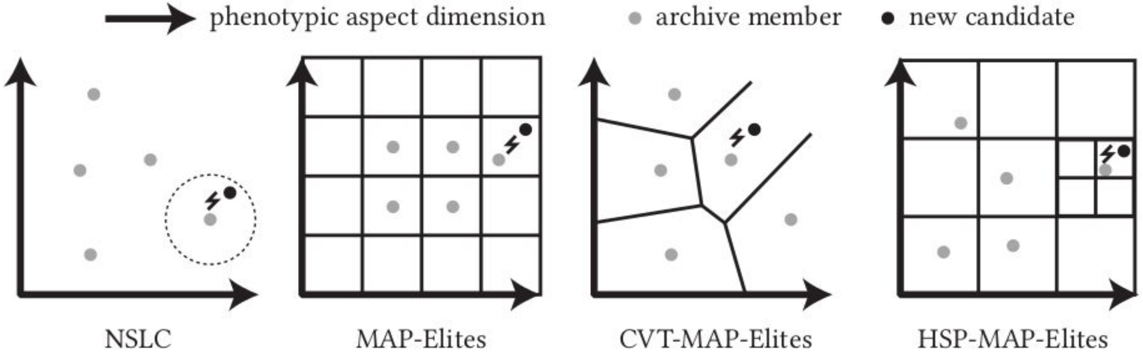


Fig. 3 The first two QD algorithms introduced an unbounded (NSLC) and fixed-grid bounded (MAP-Elites) version of the phenotypic archive. The fixed grid in the latter can be replaced with a Voronoi or a hierarchical subdivision of the phenotype space.

Figure from: Hagg, A. (2019). Discovering Modes using Quality Diversity Algorithms. (tbp)

Archive types

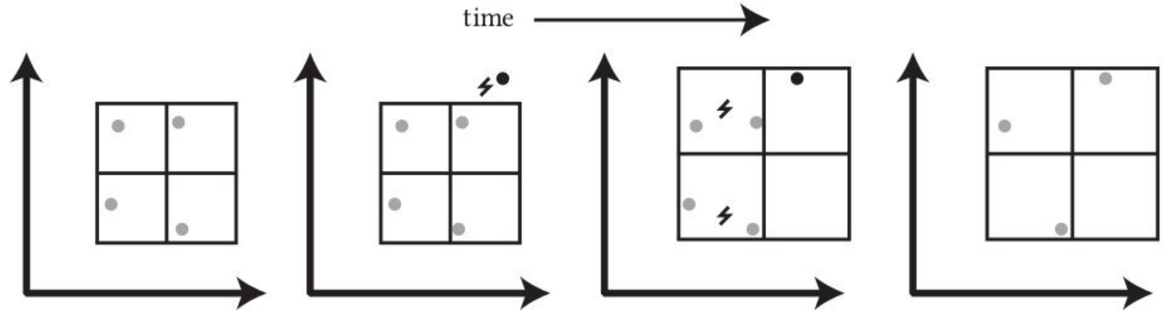


Fig. 4 Expansive MAP-Elites. The bounds of the MAP-Elites archive can be expanded over time in case the bounds cannot be predetermined [76].

Figure from: Hagg, A. (2019). Discovering Modes using Quality Diversity Algorithms. (tbp)

Archive types

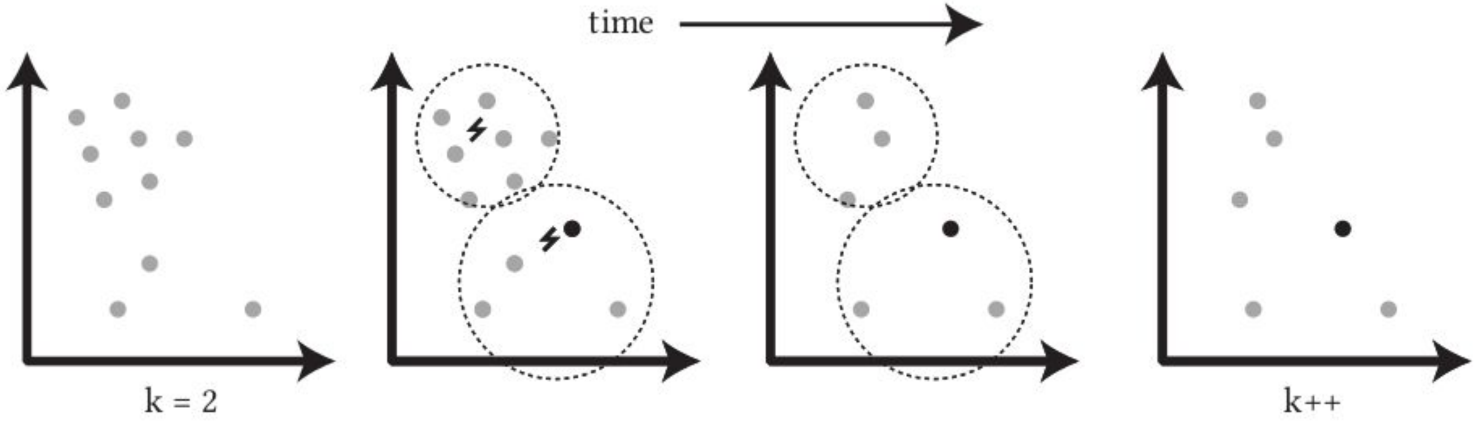


Fig. 5 Cluster-Elites. The subdivision of an unbounded archive can be created using clustering techniques. The number of clusters can then be increased over time [76].

Figure from: Hagg, A. (2019). Discovering Modes using Quality Diversity Algorithms. (tbp)

Archive types

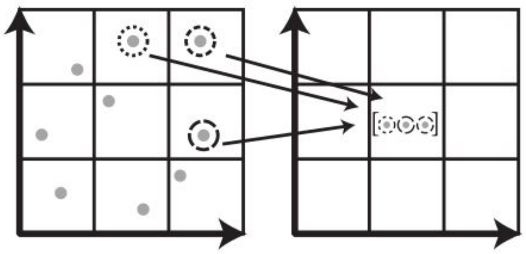


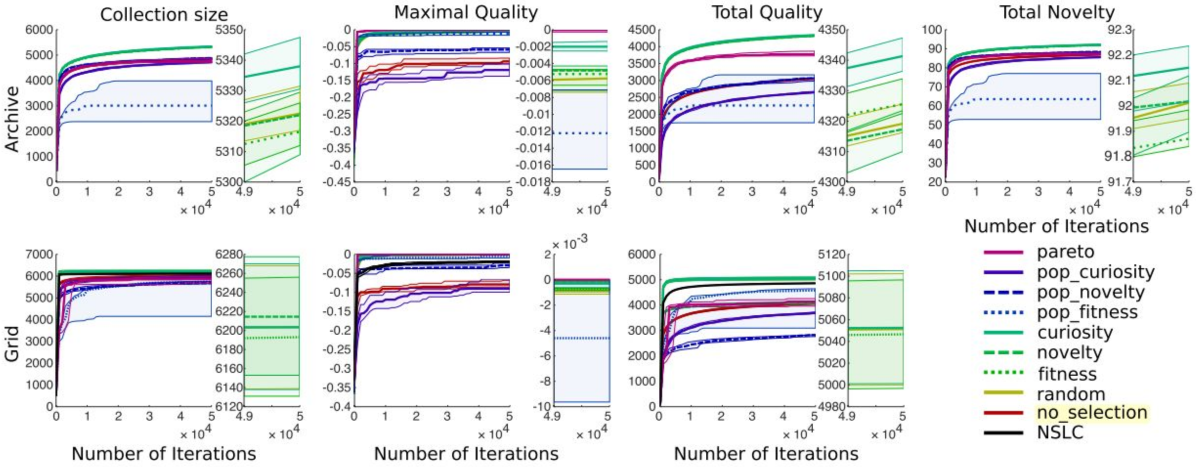
Fig. 6 Hierarchical behavior repertoire. By building up a hierarchy of stacked archives, each archive can be filled using compositions of primitives from another layer.

Figure from: Hagg, A. (2019). Discovering Modes using Quality Diversity Algorithms. (tbp)

2d. Selection Procedure

Selection Procedures

- Random
- Proportionate to score
  - Fitness
  - Novelty
  - Curiosity: propensity of individual to generate successful offspring
- ...

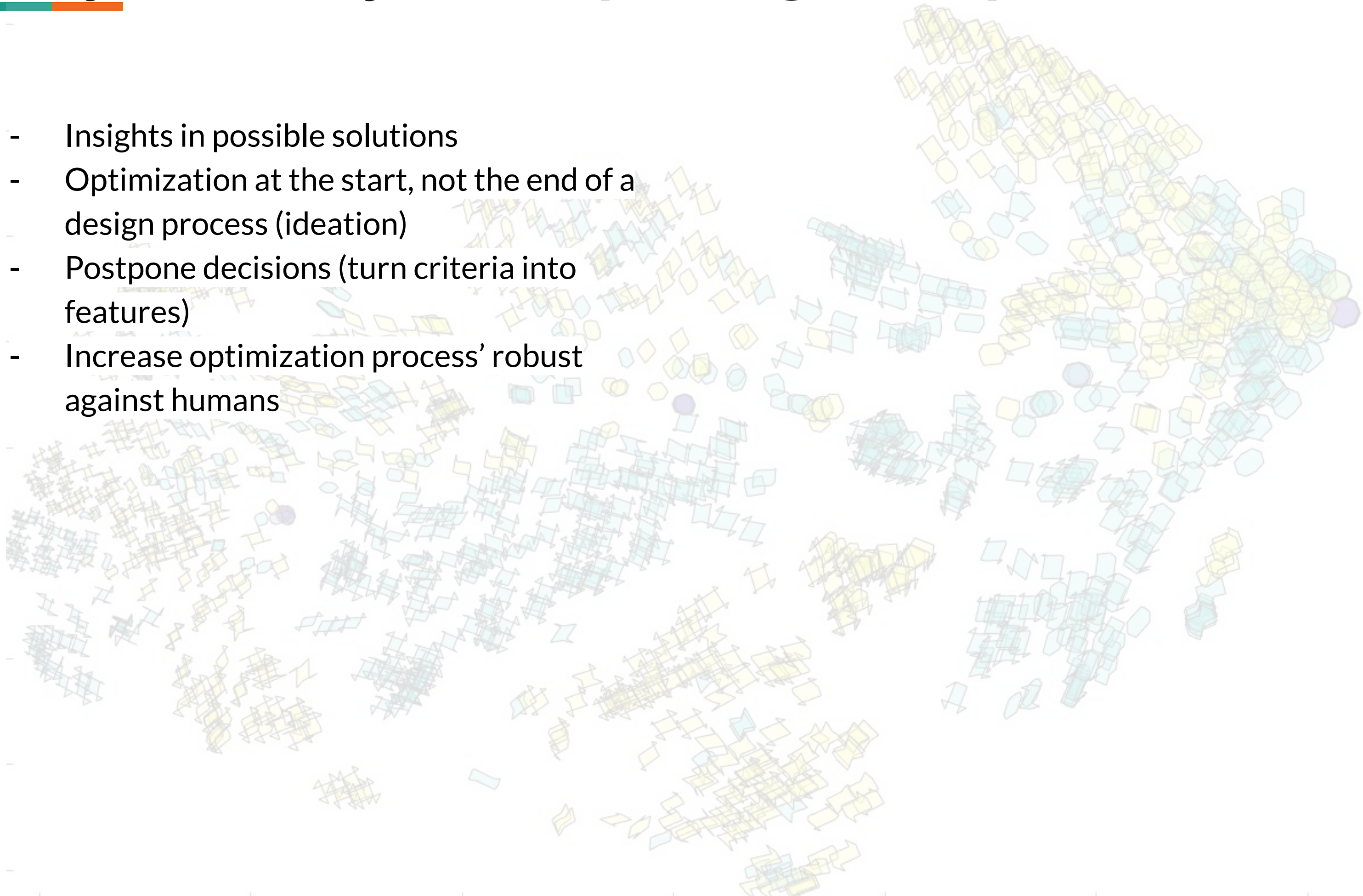


Cully, A., & Demiris, Y. (2017). Quality and Diversity Optimization: A Unifying Modular Framework. EC

3. MOO vs QD: a small experiment

Why Diversity in Morphological Optimization?

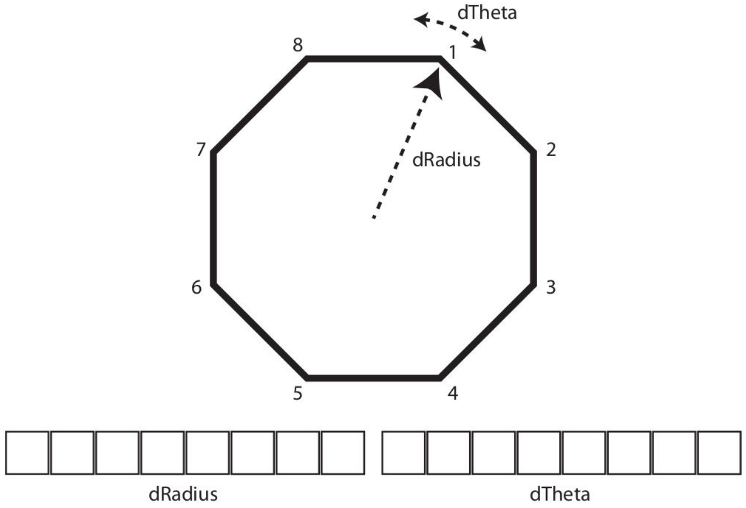
- Insights in possible solutions
- Optimization at the start, not the end of a design process (ideation)
- Postpone decisions (turn criteria into features)
- Increase optimization process’ robust against humans



Comparing Diversity between MOO and QD

Free form deformation of n-polygons

Criteria: **area**, **perimeter length**, **point symmetry**

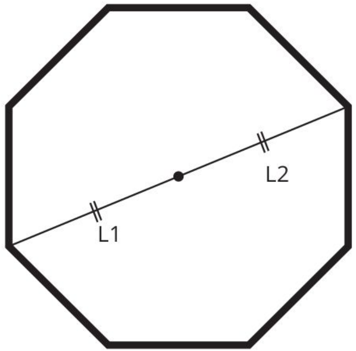


MOO: NSGA-II:

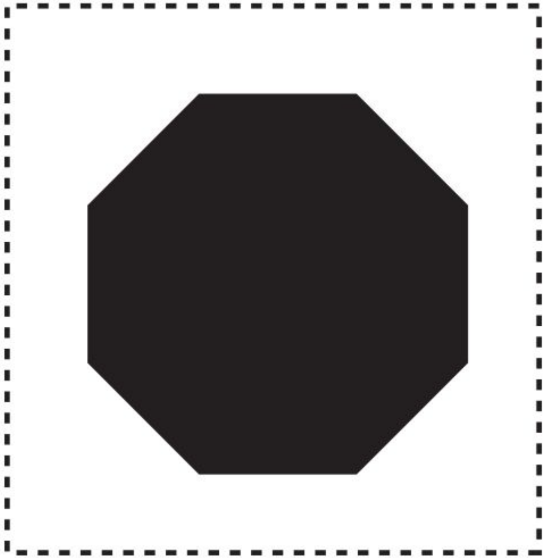
- three criteria/objectives
    - maximize point symmetry
    - min. perimeter length
    - max. area

QD: MAP-Elites:

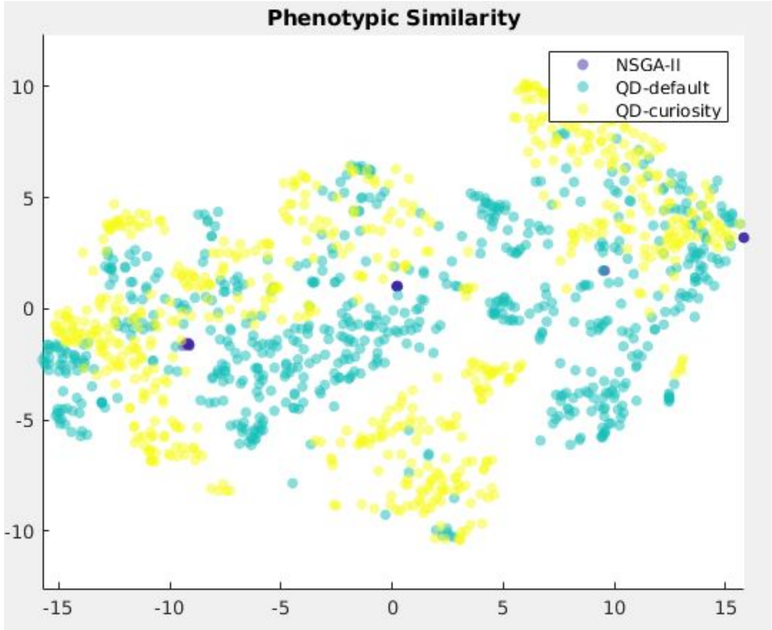
- Single objective: maximize point symmetry
- Features: area, perimeter length



Measuring Morphological Diversity



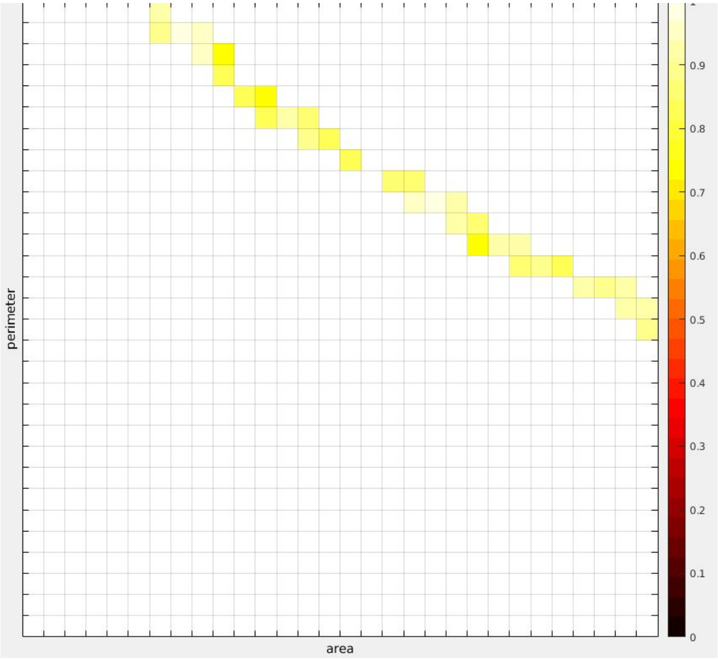
Phenotype: bitmap of filled polygon



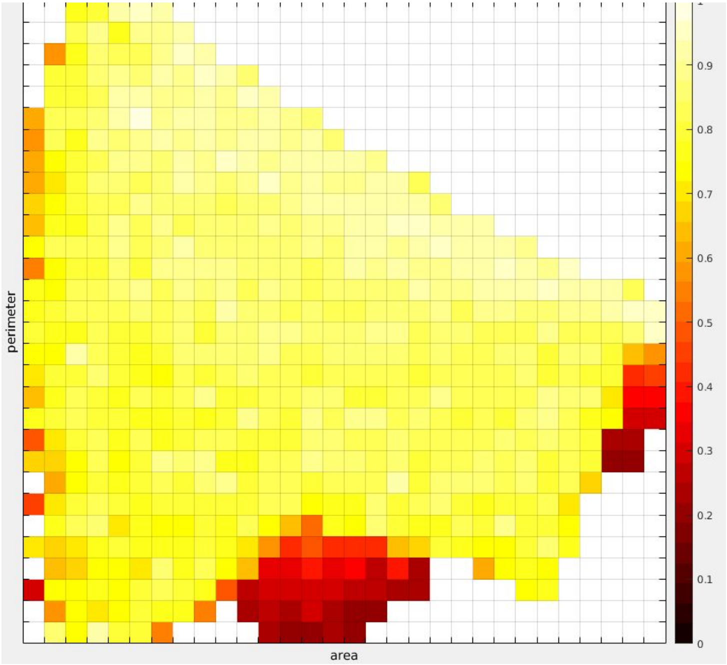
t-SNE[1] projection of phenotypes, comparing NSGA-II to two instances of a QD algorithm

[1] Van Der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing high-dimensional data using t-SNE. Journal of Machine Learning Research

Comparison in Archive Space



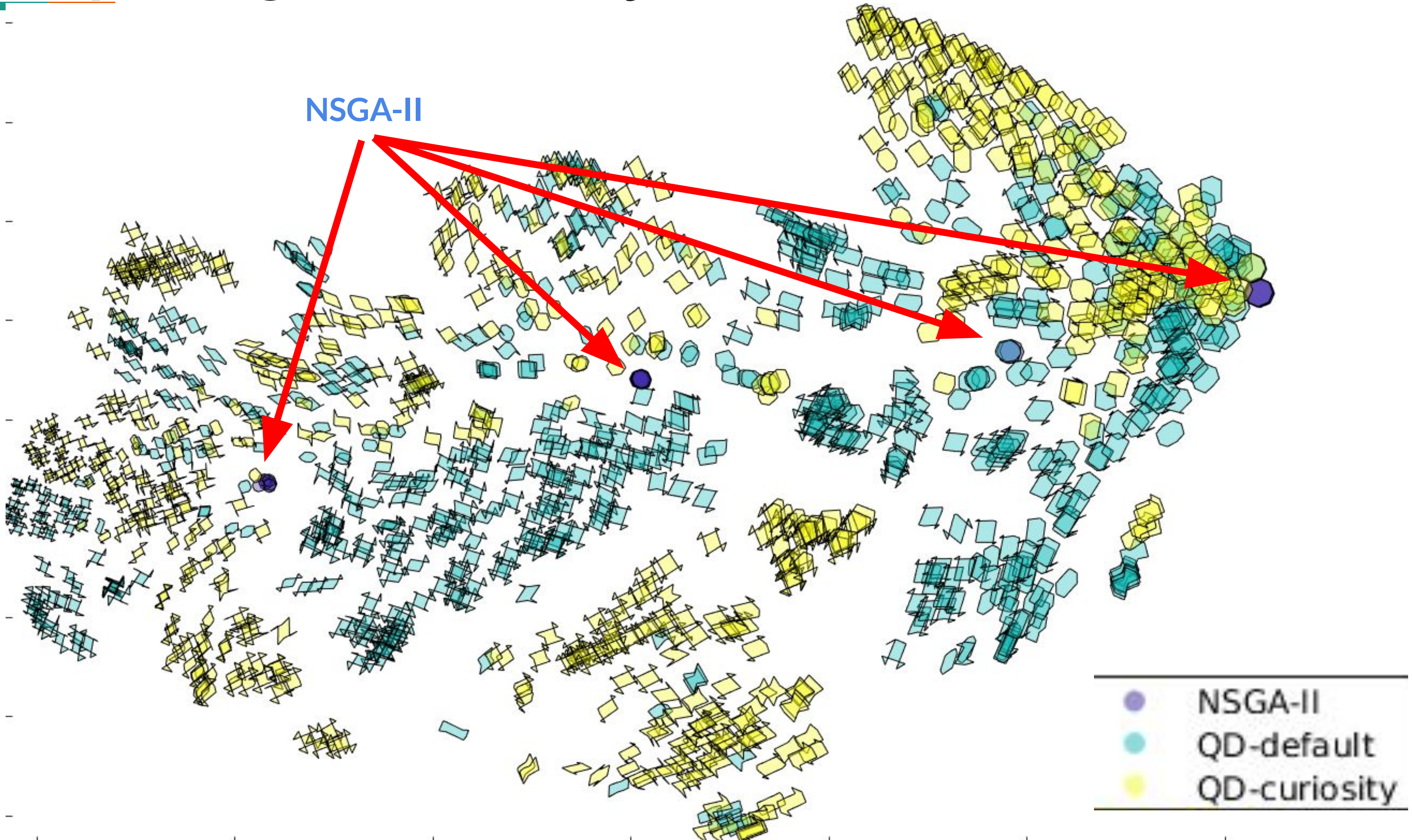
NSGA-II



MAP-Elites

This result is obvious, as in NSGA-II area and perimeter are optimization criteria.
But that is the crucial **difference of features and criteria.**

Morphological Diversity of QD > MOO



4. Search for features

Defining Phenotypic Features

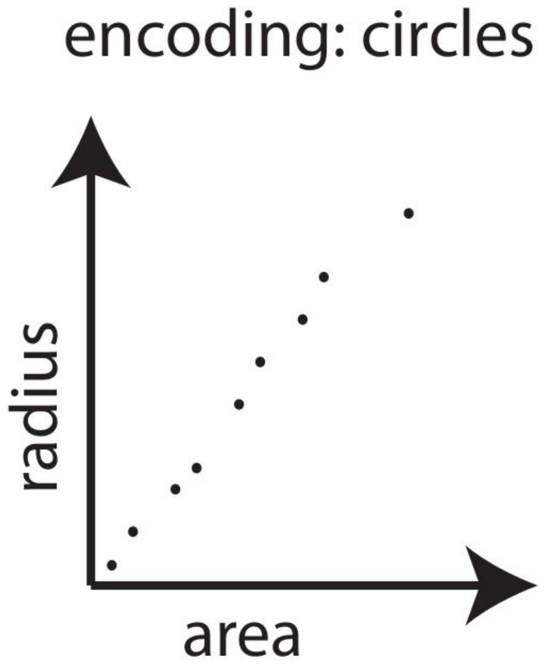
Defining phenotypic features is hard

- From experience: 3 features can already be hard
- Can we expect formal definitions of features from every user?
- **Question**: what are interesting features of airflow?
- Features might be highly correlated, and diversity collapses.
  Question: what would happen if the correlation between two features equals 1?

Defining Phenotypic Features

Defining phenotypic features is hard

- From experience: 3 features can already be hard
- Can we expect formal definitions of features from every user?
- **Question**: what are interesting features of airflow?
- Features might be highly correlated, and diversity collapses.
  Question: what would happen if the correlation between two features equals 1?

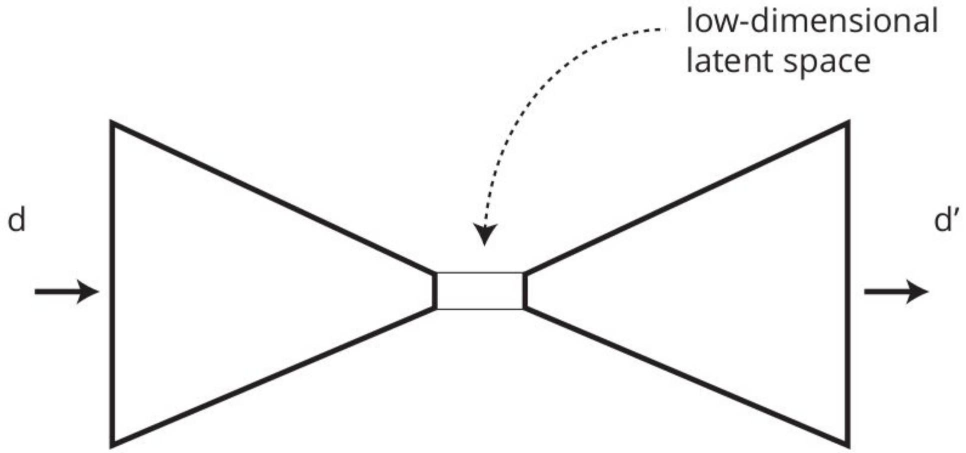


Automating search for features

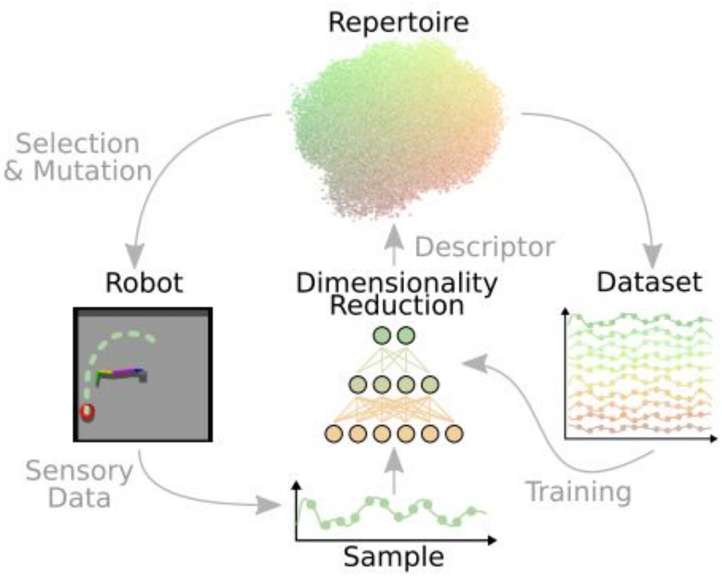
Can we automate finding phenotypic similarity?

- Idea: use autoencoders!
- High dimensionality of phenotypes (d)
- Find low-dimensional description
- Use latent space dimensions as features

(finally some deep learning… :|)

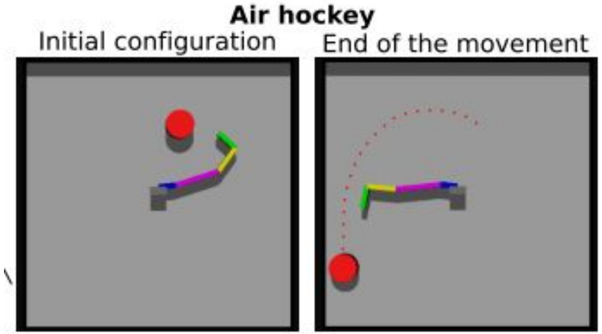


Behavioral Manifold



Some evidence:

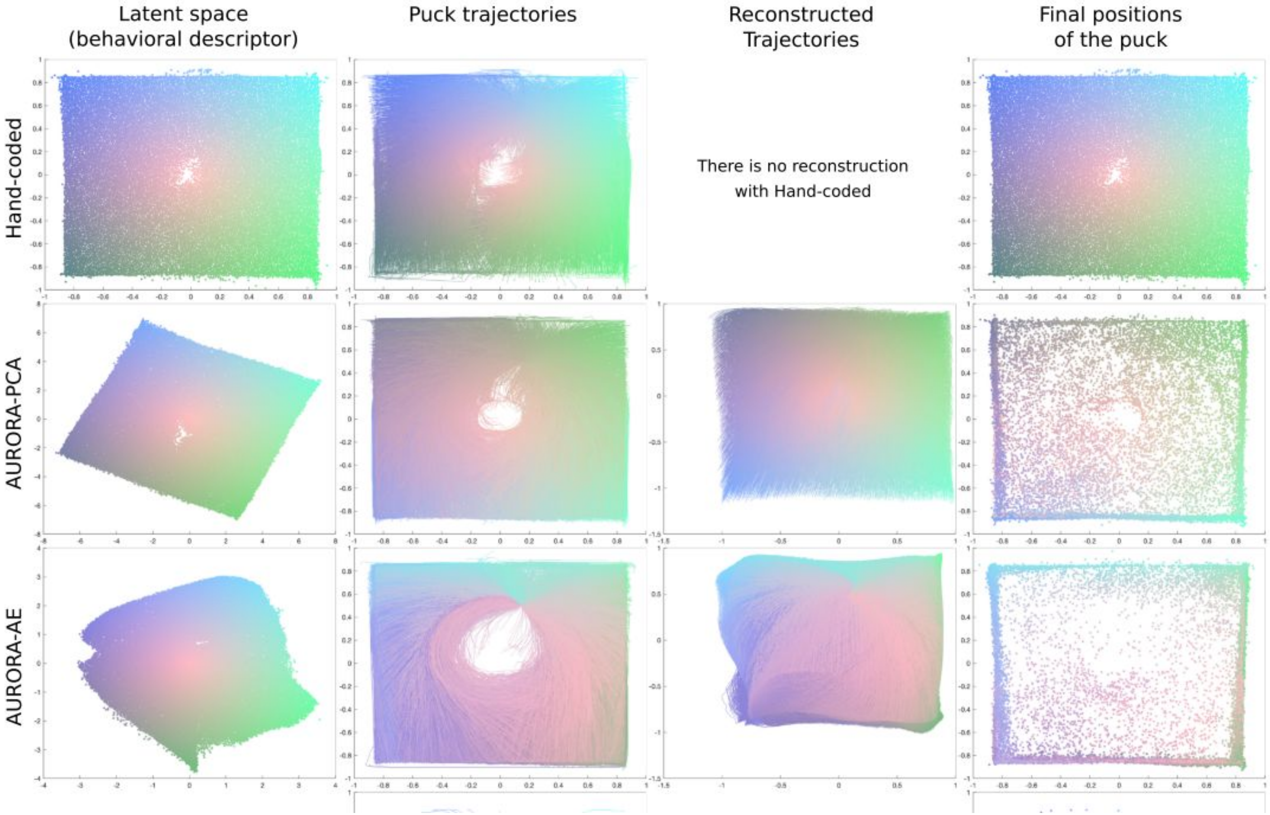
- [1] Paolo 2019
- [2] Cully 2019
- … and some of my work soon

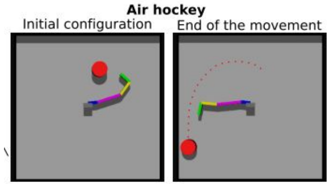


[1] Paolo, G., Laflaquière, A., Coninx, A., & Doncieux, S. (2019). Unsupervised Learning and Exploration of Reachable Outcome Space.
[2] Cully, A. (2019). Autonomous skill discovery with quality-diversity and unsupervised descriptors.

Behavioral Manifold

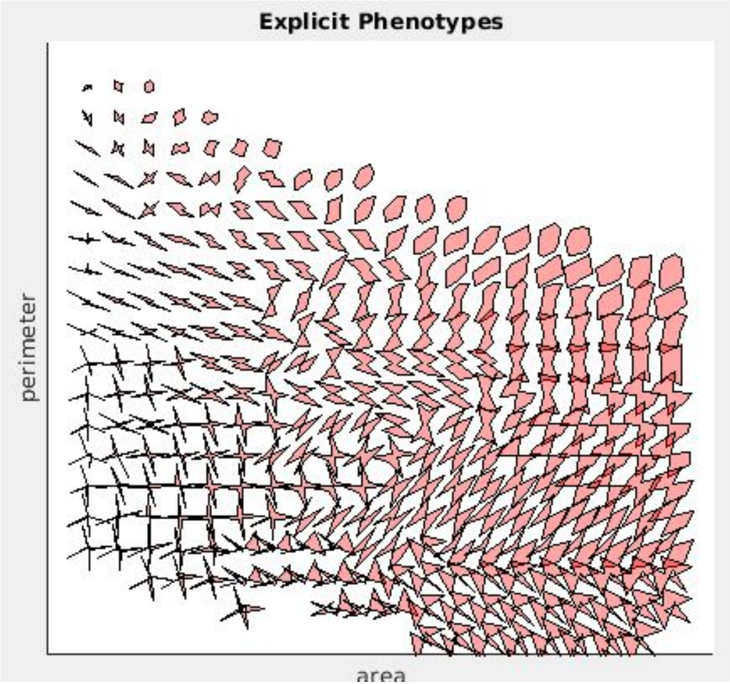
-





Cully, A. (2019). Autonomous skill discovery with quality-diversity and unsupervised descriptors.

Morphological Manifold*



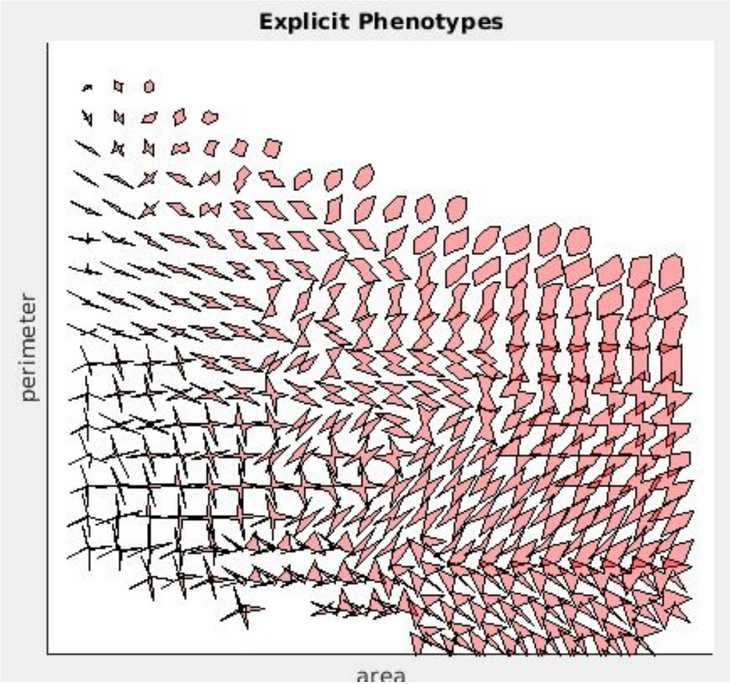
?

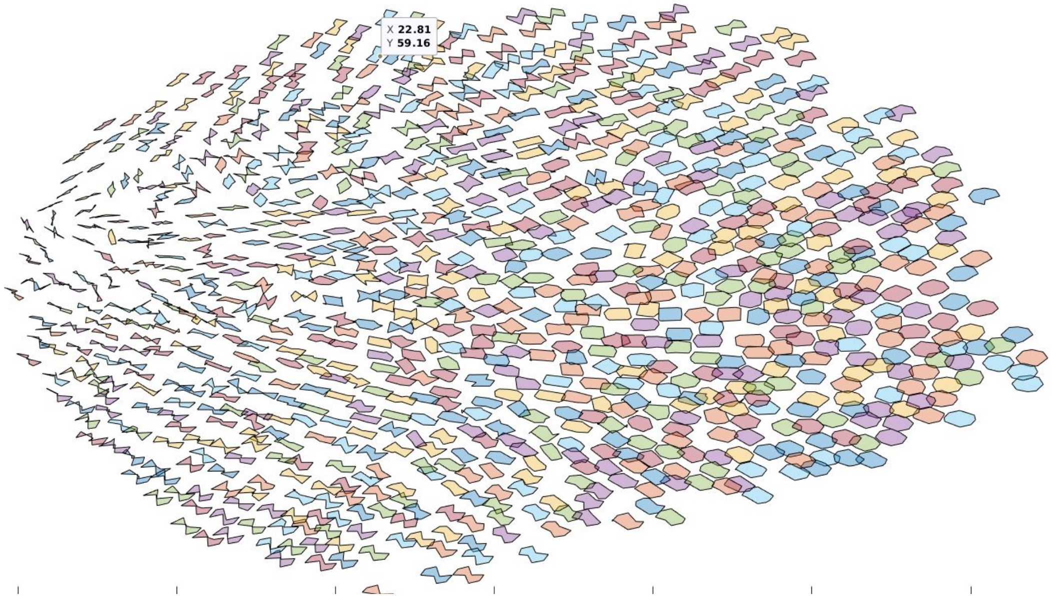
Manually defined features

Learned features (autoencoder)

* writing this up as we speak

Morphological Manifold





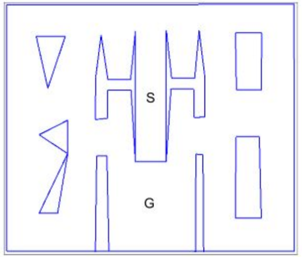
Manually defined features

Learned features (autoencoder)

5. Insights

Alignment of Quality and Diversity

QD algorithms generally perform better than NS alone when BC is uncorrelated to fitness



BC dimensions used:

1. High alignment: endpoint (x,y) coordinates
2. Low alignment: direction

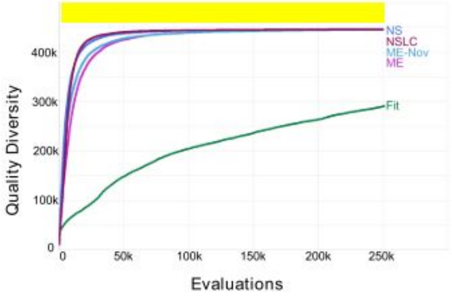


Figure 2: EndpointBC (very high alignment). In this

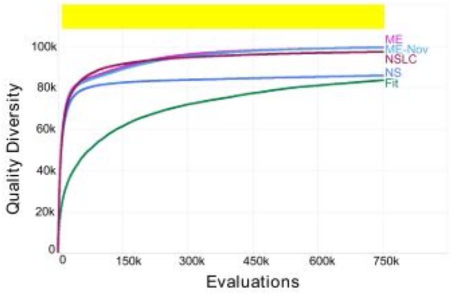


Figure 5: DirectionBC (low alignment). With an al-

Pugh, J. K., Soros, L. B., & Stanley, K. O. (2016). Searching for quality diversity when diversity is unaligned with quality. LNCS
Pugh, J. K., Soros, L. B., Szerlip, P. A., & Stanley, K. O. (2015). Confronting the challenge of quality diversity. GECCO

Stepping Stones

- Evolutionary paths often pass through several basins of attraction before ending up in the target basin.
- Decomposing a problem into subtasks is often necessary to reach complicated tasks.
- **stepping stones**: subtasks can be used as intermediate goals that allows a system or organism to be "guided" to a certain complex goal in a multimodal domain.

How does this effect occur in QD algorithms?

The order of the subtasks (simple to complex) is clearly of importance to the performance of an optimization algorithm.

QD algorithms **avoid ordering** altogether by allowing all subtask combinations to be explored using the QD archive. Simple subtasks are saved alongside of complex subtasks.

Huizinga, J., & Clune, J. (2018). Evolving Multimodal Robot Behavior via Many Stepping Stones with the Combinatorial Multi-Objective Evolutionary Algorithm
Meyerson, E., & Miikkulainen, R. (2017). Discovering evolutionary stepping stones through behavior domination. GECCO

Alignment of Genotype and Phenotype

Impact on QD performance:

- Genetic sensitivity: small perturbations in genome can lead to large changes in phenotype.
- Genetic neutrality: multiple species can be assigned to the same niche. Niches are defined by simplification of full phenotype. Comparison might lead to preference of one over the other..
- Ecological neutrality: morphological variations might show similar behavior in a niche.

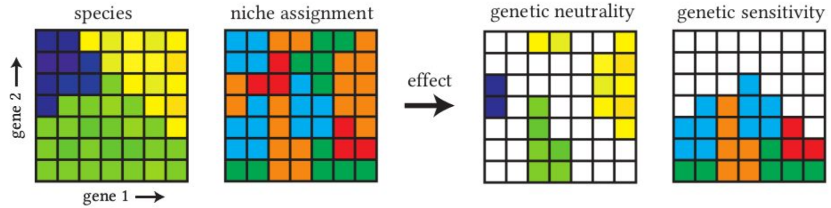


Fig. 10 Species and phenotypic niche borders are not necessarily aligned. Due to genetic neutrality, different genotypes can expose similar behavior and be put into the same niche. Genetic sensitivity can cause members of the same species to occupy different niches.

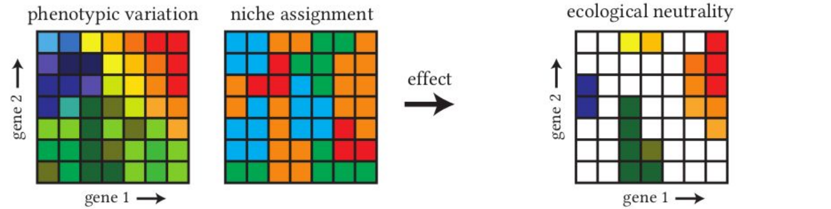


Fig. 11 Due to ecological neutrality, different phenotypes can, under given circumstances show the similar behavior. This can cause different species to occupy the same niche.

Influence on Exploitation and Exploration

- Mutation operator and strength
- Selection scheme
- Effect: erosion of archive
- Increasing Exploration.
    - Using novelty metrics as part of BC space.
    - Using multiple non-parallel archives
    - Using surprise search

6. Performance Metrics

Metrics

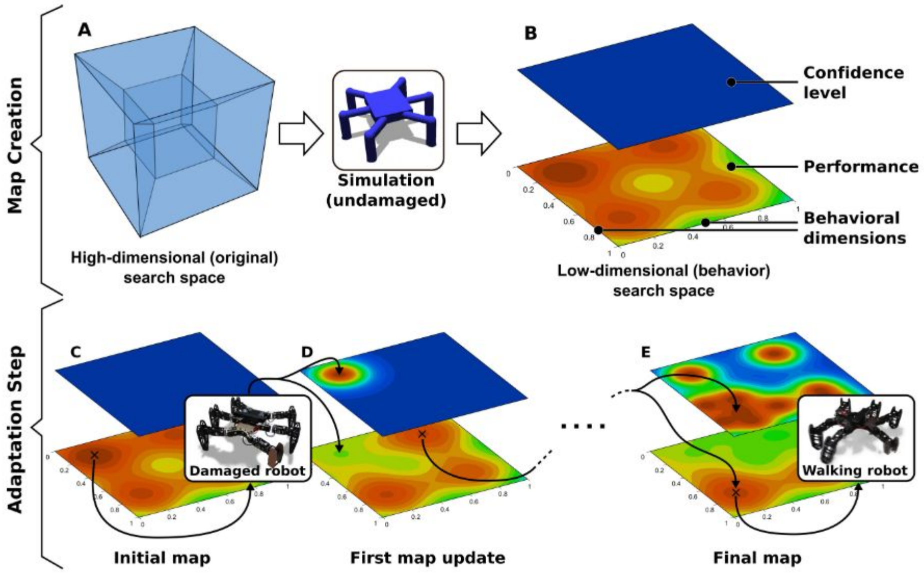
- Collection size
- Maximal quality
- Total quality: sum of all fitness values
    - increases when fitness of individuals
    - or individual gets added
- Total novelty

Or more classical diversity metrics:

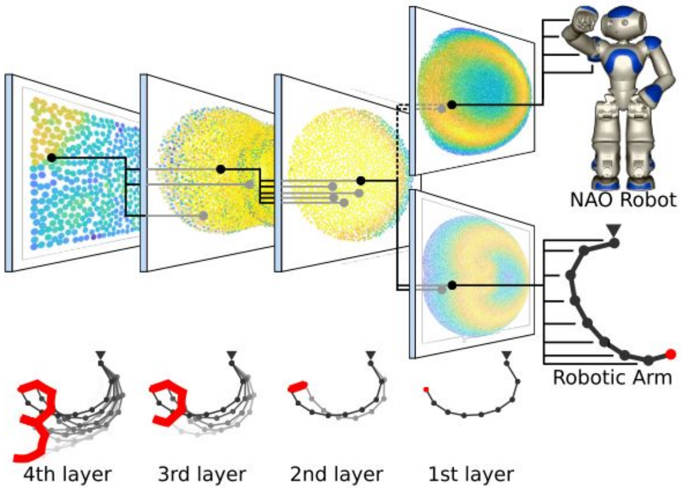
- Volume of genetic convex hull
- Distances in genetic space
- Sparsity

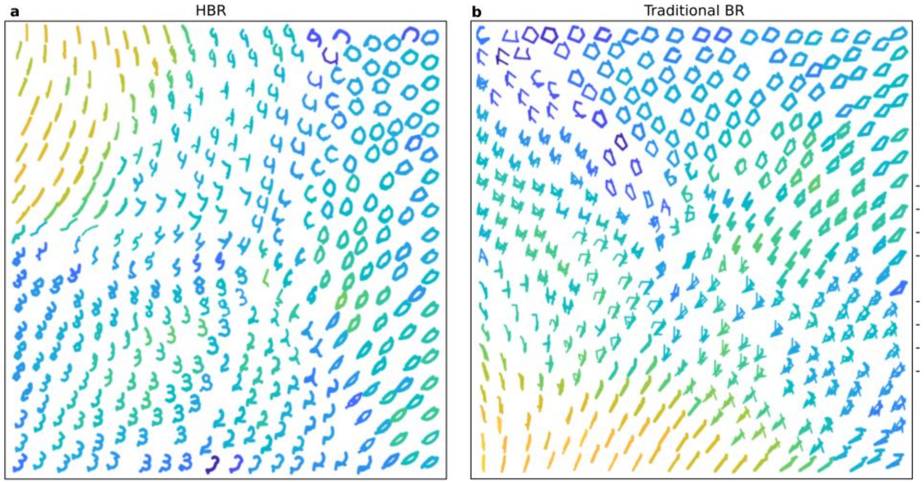
7. Applications

Walking Gaits



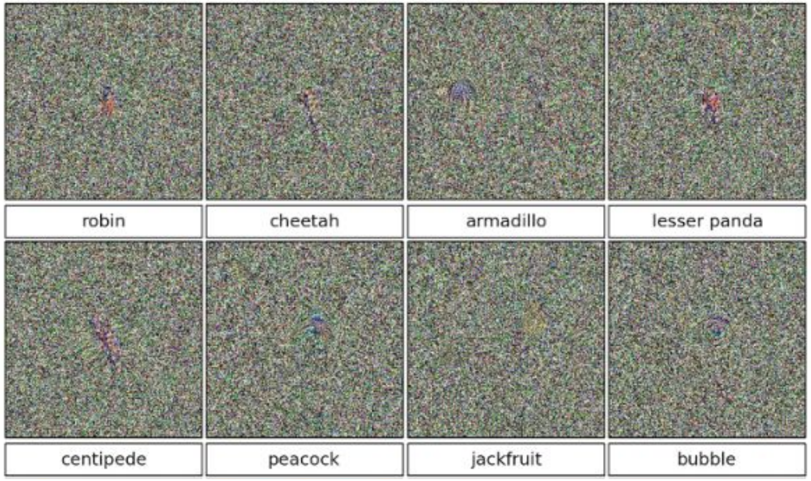
Hierarchical Behavior

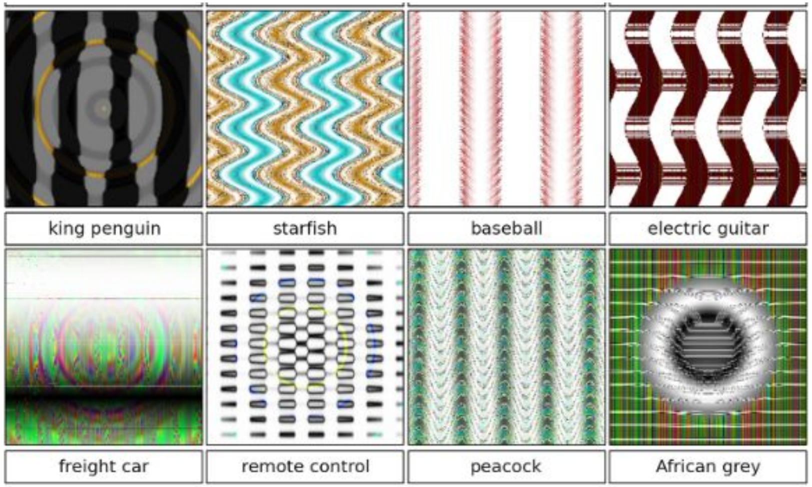




Cully, A., & Demiris, Y. (2018). Hierarchical Behavioral Repertoires with Unsupervised Descriptors.

Generating Adversarial Examples in DL

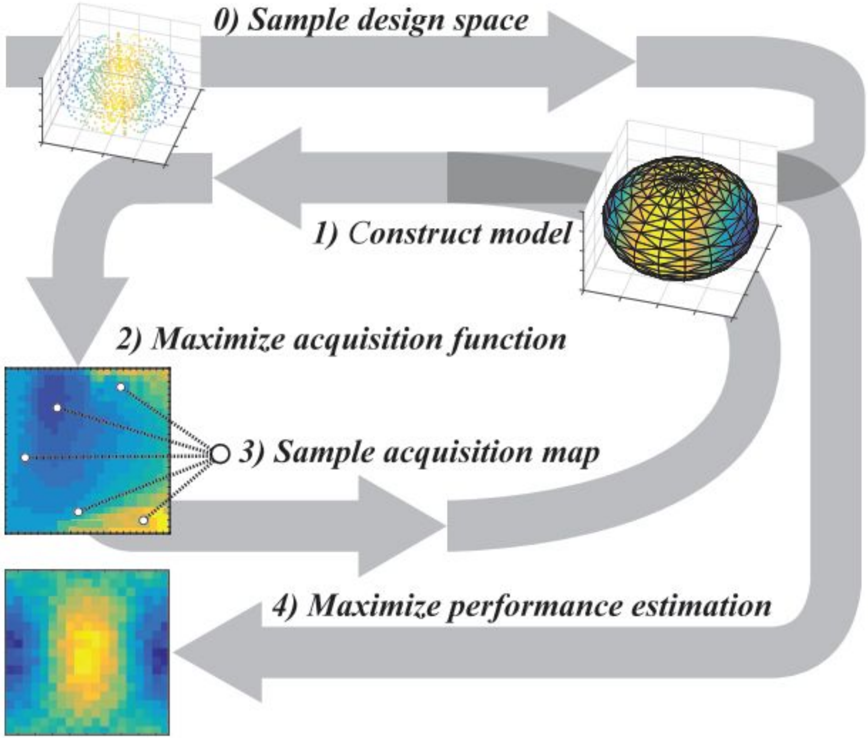




Nguyen, A., Yosinski, J. and Clune, J., (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. ICCV

Morphological Optimization w. Surrogate Assistance

- QD needs 100.000s-1.000.000s of evaluations
- Efficient Quality Diversity with standard Gaussian Process surrogate model
- Acquire expensive precise evaluations with upper confidence bound sampling



Gaier, A., Asteroth, A., & Mouret, J.-B. (2017). Data-Efficient Exploration, Optimization, and Modeling of Diverse Designs through Surrogate-Assisted Illumination

Morphological Optimization w. Surrogate Assistance

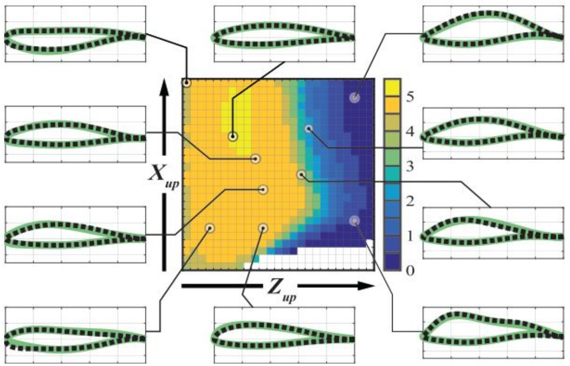


Fig. 3. *Design Space Overview with SAIL*
Prediction map produced by SAIL after 1000PE.
Border: Median performing designs found by SAIL in green, best designs found by CMA-ES in black.

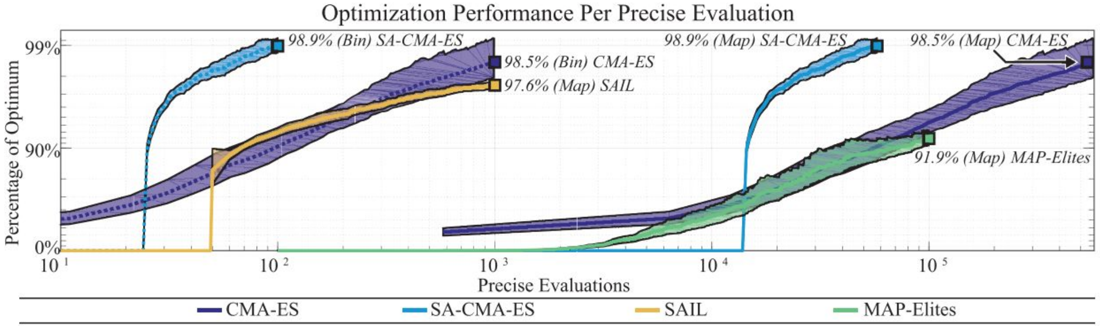
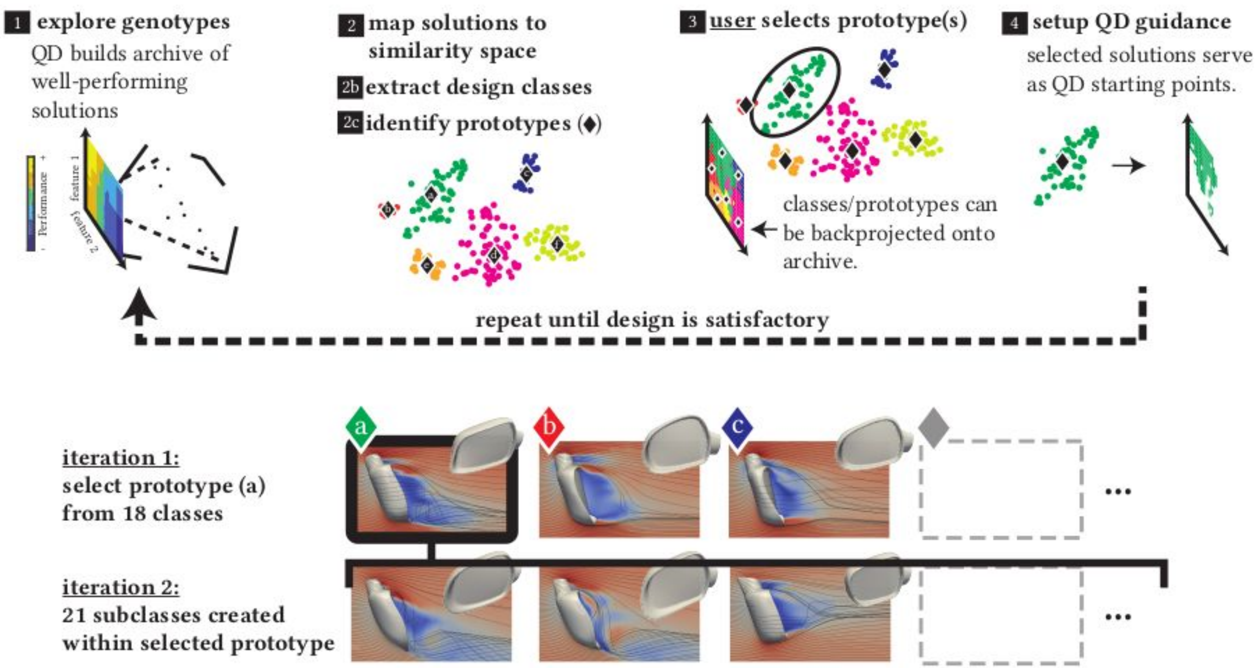


Fig. 7. *Optimization Efficiency in a Single Bin and Over the Entire Design Space*
Computational efficiency of CMA-ES, SA-CMA-ES, MAP-Elites, and SAIL in precise evaluations. *Bin*: median progress towards optimum in every bin. *Map*: performance of CMA-ES and SA-CMA-ES is median bin performance multiplied by number of bins. Performance of individuals produced to construct initial models is set to 0%. Bounds indicate one standard deviation over 20 replicates. PEs and performance in log scale.

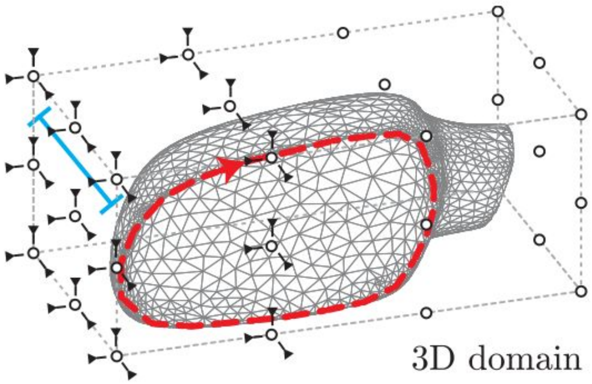
Computer Aided Ideation

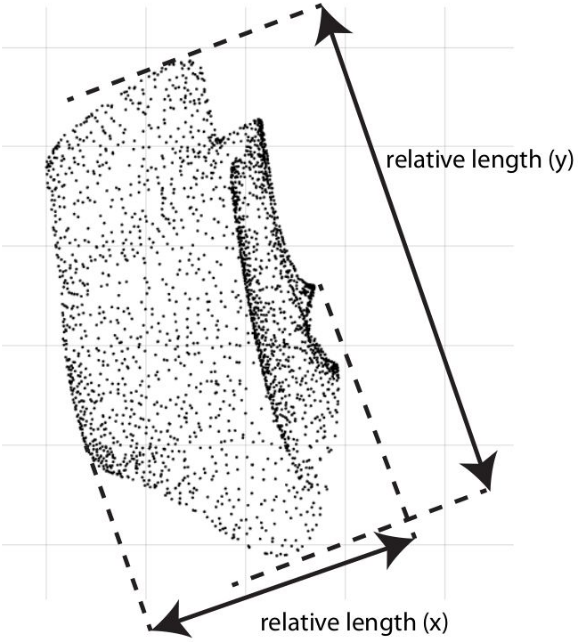


Hagg, A., Asteroth, A., & Bäck, T. (2018). Prototype Discovery using Quality-Diversity. PPSN

Expensive Morphological Optimization

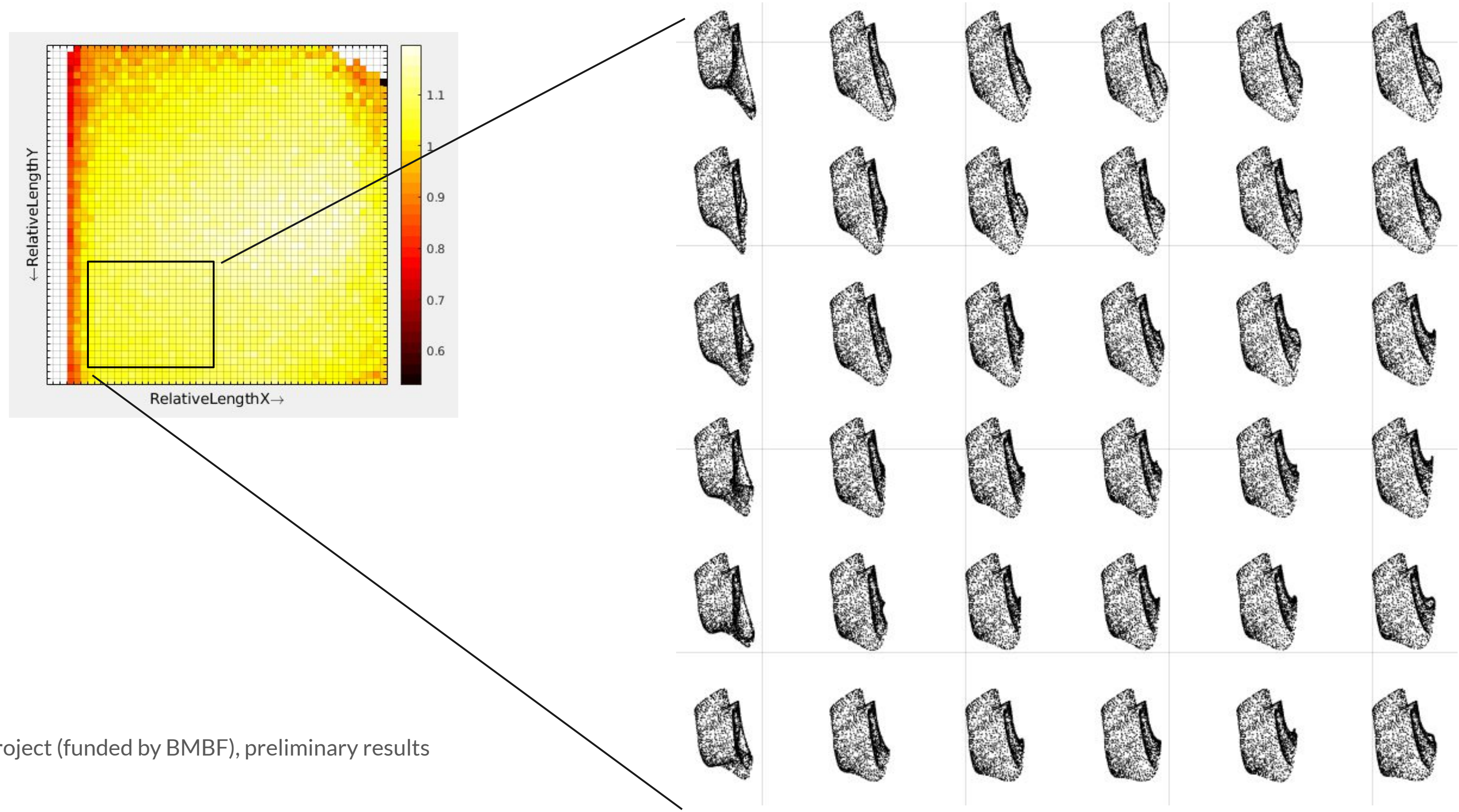
Free form deformation of mirror (TUM)
Minimize drag coefficient (OpenFOAM)





AErOmAt project (funded by BMBF), preliminary results

Expensive Morphological Optimization



AErOmAt project (funded by BMBF), preliminary results

8. Conclusion

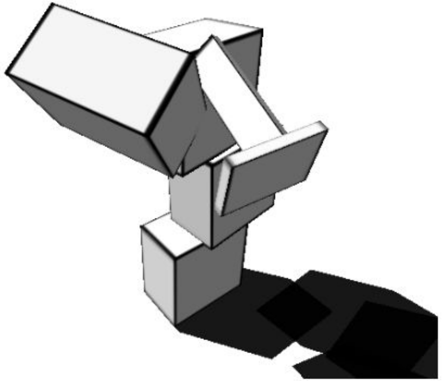
Summary

Quality Diversity:

- Phenotypic niching
- Behaviors, morphological features, etc.
- QD produces 100s, 1000s of optimal designs or behaviors

QD: gives us a notion of what is possible and high-performing, maximizing phenotypic diversity of a large solution set.

Maybe design criteria are not always what they are, or at least might be premature. Try turning criteria into features.



Thank you!

Q&A ?