Bayesian Quality Diversity Search with Interactive Illumination

Paul Kent
University of Warwick
UK
paul.kent@warwick.ac.uk

Juergen Branke
Warwick Business School
UK
Juergen.branke@wbs.ac.uk

ABSTRACT

This paper presents a novel way for interactively identifying a most preferable solution based on quality and behavioural characteristics. Our algorithm combines the principles of Quality-Diversity Search and Bayesian Optimization to create Gaussian Process surrogate models of the behaviour and fitness space. Unlike traditional Quality-Diversity methods which aim to find good solutions with different behavioural characteristics, we propose a three-step interactive approach that allows a decision maker to efficiently identify the most preferred solution(s). In the first stage, it uses an entropy-based acquisition function to generate an illumination model, followed by an interactive phase where the decision maker can specify regions of interest and a target behaviour. These preferences are then utilized by an improvement greedy acquisition function to guide the optimization process and quickly identify a solution close to the user-specified target. In a case study, with a simulated decision maker, we demonstrate that our approach can find better solutions much more quickly than by selecting the most preferred solution from an archive generated with MAP-Elites.

KEYWORDS

quality diversity, surrogate assisted, Bayesian optimization, illumination, interactive

ACM Reference Format:

Paul Kent and Juergen Branke. 2023. Bayesian Quality Diversity Search with Interactive Illumination. In *Genetic and Evolutionary Computation Conference (GECCO '23), July 15–19, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3583131.3590486

1 INTRODUCTION

Quality-Diversity optimization (QDO)[19] is a rapidly growing area of research that attempts to find a set of behaviourally diverse solutions that are high performing. By contrast, traditional optimization is focused on pure optimization and returns a single point.

QDO is an interdisciplinary field that draws on techniques from mathematics, computer science, engineering, and the natural sciences and often uses evolutionary algorithms for optimization. It has numerous applications in fields such as machine learning [14], robotics [3, 16], chemical discovery [21], engineering design [6] and game design [5, 7].

GECCO '23, July 15–19, 2023, Lisbon, Portugal

ACM ISBN 979-8-4007-0119-1/23/07.
https://doi.org/10.1145/3583131.3590486

A key outcome of QDO is an understanding of the relationship between a solution's behavioural characteristics and the maximum achievable fitness with those characteristics. The process of learning this, often complex, relationship is referred to as *illumination* and it creates key insight to end-users on how choices in the behavioural domains impacts the possible performance.

Recent methods using surrogate models have emerged from both the evolutionary optimization community [6, 8] and the field of Bayesian optimization [11]. These methods, while restricted to lower dimensional problems due to the Gaussian Process (GP) models employed, show vast improvements in sample efficiency and produce high quality models of the underlying problem in relatively few observations.

There are various motivations for searching for a diverse set of good solutions, e.g., different solutions are needed at different times (different robotic motions depending on the task), a problem actually requires a diverse set of solutions (a set of artificial players in game design), or a decision maker (DM) wants to be able to examine a diverse set of solutions before selecting one (in engineering design).

Over the past few years there has been interest in the area of interactive optimization, or Human In the Loop (HILO), optimization. This includes some works in the QDO literature [2] and the interactive process can be used to discover behavioural qualities that are of interest to the decision maker.

In the present work we propose an interactive method for identifying a most preferred solution based on fitness and behavioural characteristics. It utilises ideas from QDO and Bayesian optimization to rapidly build representative models of the mapping from behaviour to maximum achievable fitness before allowing a human DM to guide the search towards areas of interest and eventually a specified target behaviour.

The main contributions of this paper are

- The first (to the best of our knowledge) framework for interactive QDO in order to effectively and efficiently identify the solution most preferred by a DM.
- A new entropy based acquisition function for use with the BOP-Elites algorithm which balances model building and fitness improvement.
- An 'adaptive resolution' strategy which allows us to present information to the decision maker.
- A simple stopping criterion for when we believe our models are sufficiently insightful.
- A Monte Carlo acquisition function for searching for the best solution adhering to the user-specified target behaviour.

Compared to the standard approach of running a QDO algorithm to generate a behaviourally diverse archive of good solutions and picking the most preferred from the generated set, our approach has two key advantages. First, because it is interactive it can quickly focus search on the most promising regions, thereby saving many function evaluations unnecessarily illuminating less relevant regions. Second, any pre-computed solution archive can only provide a discrete set of solutions to choose from, while our interactive framework allows the DM to specify a target in continuous behavioural descriptor space, and we can very quickly locate solutions extremely close to the desired target behaviour.

The paper is structured as follows. After a formal problem definition in Section 2, we survey related work in Section 3, then describe our methodology in Section 4. We demonstrate the use of our interactive method in Section 5, and show that it is much more efficient than standard approaches. The paper concludes with a summary and some ideas for future work.

2 PROBLEM DEFINITION

We consider an n dimensional box-constrained search domain $\mathcal{X} \subset \mathbb{R}^n$ with a fitness function

$$f(x) : \mathcal{X} \to \mathcal{Y} \subset \mathbb{R}$$

describing the quality of a solution, and a function g which produces m additional descriptor values

$$g(x) : \mathcal{X} \to \mathcal{B} \subset \mathbb{R}^m$$

that define additional behavioral characteristics relevant to the DM. We consider the case where both functions are black-box and expensive and can only be evaluated simultaneously. Our interaction mechanism further assumes $m \leq 2$. In principle, the framework should be applicable also to higher descriptor spaces, although this would be much harder for a DM to analylze and would require appropriate interaction mechanisms.

Our goal is to help a human DM to identify their most preferred solution, taking into account the solution's quality (which we assume w.l.o.g. to be maximized) and behavioral descriptor values (on which the preferences of the DM are unknown a priori).

Note that combinations of descriptor values define an m dimensional descriptor space and for each point $b \in \mathcal{B}$ in descriptor space there exists a solution $x^*(b)$ with optimal fitness $y^*(b)$. A DM would only be interested in those best-fitness solutions for any point b, so ideally we would like to find the function

$$h : \mathcal{B} \to \mathcal{Y}^* \tag{1}$$

with

$$h(b) = \max_{x|g(x)=b} f(x). \tag{2}$$

A DM could then look at all $b \in \mathcal{B}$ and the associated best-fitness values $h(b)$ and pick their most preferred solution. Unfortunately, as h and g are expensive black-box functions, because of the max operator, and the non-injective mapping, it is not possible to derive function h analytically.

3 RELATED WORK

3.1 Quality-Diversity optimization

Quality Diversity optimization (QDO)[19] attempts to find a set of points that have a high quality but also have a high diversity in descriptor space. To this end, it partitions the descriptor space into regions $\mathcal{R} = \{r_1, r_2, \ldots, r_{|\mathcal{R}|}\}$. The task of a QD illumination algorithm is then to find for each descriptor region $r \in \mathcal{R}$ the *elite* point $e_r \in X$ with the maximum fitness, i.e.,

$$e_r = \arg\max_{x \in r} f(x). \tag{3}$$

These elite points are then stored in a *solution archive* $\mathcal{E}$, a data structure that holds the best identified solution for each discrete region in descriptor space. This archive has the nice characteristic of being easy to visualize in case of low descriptor dimensions and thus being very insightful for DMs.

In a standard QDO setting, the performance of the algorithm can be measured by the value of the set of elite points in the solution archive $\mathcal{E}$, and a common metric is the QD-Score [19].

$$\text{QD-Score} = \sum_{r \in \mathcal{R}} f(e_r), \tag{4}$$

where, without loss of generality, we assume that f is positive and $f(e_r)$ is defined as zero if the algorithm has not yet found any solution in region r.

Most quality-diversity research occurs in the context of the MAP-Elites algorithm [17] which is a powerful evolutionary approach. This algorithm leverages the best points discovered in the solution archive to create a new generation of candidate solutions, thus efficiently filling the archive and directing the search towards optima. Despite its computational efficiency, the MAP-Elites algorithm requires a substantial number of function evaluations to converge to a satisfactory outcome. In response to this challenge, recent developments in surrogate-assisted approaches have sought to enhance the sample efficiency of the search process.

3.2 Surrogate Assisted QDO

The Surrogate Assisted Illumination (SAIL) algorithm, described in [17], deviates from the traditional MAP-Elites approach by performing the optimization over a Gaussian Process (GP) surrogate model of the fitness function, rather than the computationally expensive original function. Utilizing the Upper Confidence Bound (UCB) acquisition function, the algorithm strikes a balance between exploring high-performing areas of the fitness landscape and reducing uncertainty in the fitness predictions. The SAIL algorithm, as well as the Surrogate Assisted Phenotype Niching (SPHEN) algorithm [8], which extends SAIL to work with black-box descriptors, demonstrate a marked improvement in sample efficiency in QDO problems of small enough dimensions.

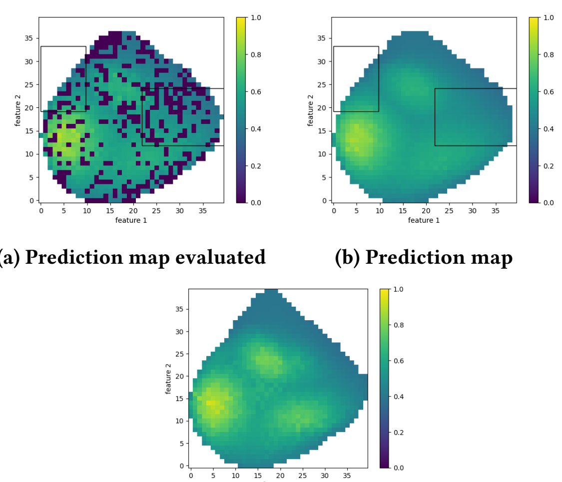
Surrogate Assisted QDO algorithms, in contrast to traditional quality-diversity optimization algorithms, output a *prediction archive* instead of a solution archive. The prediction archive constitutes a set of points, each paired with its predicted fitness, and in the case of the SPHEN algorithm, its predicted region membership. This archive offers valuable insights into the problem being addressed and can serve as a useful guide for future optimization efforts.

When the descriptor evaluations are black box, the region in descriptor space a solution is predicted to lie in may be wrong. This can be seen in Figure 1 (a), in which regions for which the predicted elite e_r did not actually fall into the region r are displayed as black squares. A simple method for evaluating the quality of the prediction archives in this case is to assign zero value to predicted solutions that actually lie in a region different from the predicted one:

$$\text{PredictedQD-Score} = \sum_{r \in \mathcal{R}} \theta(e_r) \tag{5}$$

$$\theta = \begin{cases} f(e_r) & \text{if } g(e_r) = r \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

While this metric is useful to test the performance of algorithms, calculating it requires evaluating each predicted solution and is therefore computationally expensive. Nevertheless, the prediction map itself can be used to provide helpful visual feedback to a DM, displaying the predicted values in each region, with a reminder that it is prone to error, see for example Figure 1 (b). A comparison with a high resolution map created by running MAP-Elites for 100,000 evaluations Figure 1 (c) shows that the prediction map provides a very good approximation to the true mapping and therefore provides useful insights to the DM, although the prediction map misses some lower performing regions in the top and right edge.



(a) Prediction map evaluated

(b) Prediction map

(c) High resolution map

Figure 1: The evaluated prediction archive (a) shows good performance and accuracy after only 209 optimization evaluations but some regions have been mispredicted leading to drop in performance (black squares). The predicted solution archive (b) smooths out the visualization, looking similar to a high resolution evaluated map (c) and providing useful insights.

An approach based on Bayesian Optimization was proposed in [11] and further developed in [12]. The Bayesian Optimization of Phenotypic Elites (BOP-Elites) uses an adapted version of the widely used Expected Improvement acquisition function (see next section) for choosing where to sample next. BOP-Elites was designed to efficiently maximize the QD-Score of the solution archive and it outperforms other surrogate based methods in this regard.

This work builds upon the framework of the BOP-Elites algorithm, implementing an adaptation to the acquisition function (see Section 4) which changes the focus from increasing the fitness of the solution set to building a good predictive model for illumination, interaction and eventual exploitation.

3.3 Bayesian optimization

Bayesian Optimization (BO) is a probabilistic model-based approach to optimization that uses Bayesian inference to model the unknown objective function and make informed decisions about the next evaluation point. BO is well-suited for black-box optimization problems where the objective function is expensive to evaluate and gradient information is not available. In BO, typically a Gaussian Process (GP) model is used to represent the objective function and make predictions about its behavior. The GP models are updated with each new evaluation, providing a more accurate representation of the objective function as more evaluations become available.

BO values points using a mathematically inspired heuristic calculation called an *acquisition function*. The acquisition function encodes the importance of exploration versus exploitation and the most widely used example is the Expected Emprovement (EI):

$$EI(x) \quad = \quad \mathbb{E}\left[\max(f(x) - y^*, 0)\right] \tag{7}$$

where y^* is the fitness of the best solution found so far. This may be calculated in closed form [10, 15]

$$EI(x) = (\mu - y^*)\Phi\left(\frac{\mu - y^*}{\sigma}\right) + \sigma\phi\left(\frac{\mu - y^*}{\sigma}\right), \tag{8}$$

where μ and σ are the posterior mean and variance of the GP model and ϕ and Φ are the PDF and CDF of the standard normal distribution.

In the context of QDO, BO can be used to efficiently search for a high-quality solution archive. QD algorithms aim to generate a diverse set of high-performing solutions for a given problem, rather than a single optimal solution. By using BO to search the QD solution space, it is possible to find regions of high performance in an efficient and effective manner.

A BO acquisition function for QDO exists in the form of Expected Joint Improvement of Elites (EJIE)[11, 12]:

$$EJIE(x) = \sum_{r \in \mathcal{R}} \mathbb{P}(x \in r | D) EI_r(x). \tag{9}$$

which takes the sum of Expected Improvements to each region of some point x, weighted by the probability that it is in this region.

3.4 Interactive optimization

Interactive optimization, also known as human-in-the-loop optimization, refers to a class of optimization algorithms that allow for human input in the optimization process. This type of optimization is especially useful in applications where the optimization objectives or constraints are subjective or difficult to quantify.

There has been growing interest in the development of interactive optimization algorithms. Researchers have proposed various approaches to incorporate human input in the optimization process, including interactive genetic algorithms and evolutionary computation [13, 20], interactive multi-objective optimization [1, 9, 18], Quality-diversity [2] and others. These algorithms allow humans to provide guidance to the optimization process by specifying preferences, constraints, or even fitness evaluations.

Interactivity is especially crucial in expensive quality-diversity domains, as although descriptor dimensions are predetermined, the actual quality and behavior of generated solutions may be unclear. Providing predictions of optimal achievable quality in different regions provides the DM with valuable insights that are useful in forming preferences over the descriptor space.

Identifying solutions the DM can trust involves two critical elements: the quality of the information provided to the DM and the autonomy over the final solutions returned. In this work, we concentrate on constructing a model of the underlying mapping from descriptor space to best achievable fitness, and provide visual feedback through the prediction maps. Interactivity takes the form of directly guiding the illumination of the algorithm before the DM takes total control of the exploitation stage by defining a target descriptor and tolerance over distance from the preferred target.

4 METHOD

We propose Bayesian Optimization through Interactive iLlumination (BOIL), a Bayesian QDO algorithm that allows for an interactive illumination process, eventually allowing the DM to understand the problem, and guide the search towards the most preferred solution.

More specifically, BOIL comprises of three stages:

(1) Initial illumination. This stage aims to quickly produce an approximate predictive map that allows the DM to understand likely trade-offs and identify regions in descriptor space that appear preferable.
(2) Focused illumination. This stage allows the DM to highlight preferable regions in descriptor space, so that the algorithm can refine its predictive map specifically for those regions.
(3) Target search. Finally, the DM can specify a target $\mathcal{P}$ in descriptor space, and the algorithm will attempt to identify the highest fitness solution with the targeted descriptor values.

In the following, we will describe each of these stages in turn.

4.1 Initial illumination

We approach the illumination problem using the BOP-Elites framework [11, 12] which approximates the fitness and descriptor functions using GP models.

The (EJIE) acquisition function (9) proposed in the original BOP-Elites paper [11] is an improvement-greedy method that effectively identifies high-performing points to add to its observations set. However, its focus on the highest-performing regions may result in weaker early model building. The authors have proposed an 'upscaling' method [12] to address this issue, but we additionally offer a modified version of the acquisition function that specifically aims to mitigate errors in region predictions.

These region mis-predictions are attributed to the posterior uncertainty in the descriptor models. As QDO methods with structured archives have a finite number of regions, the uncertainty in region membership can be quantified using entropy,

$$H(x) = \sum_r \mathbb{P}(r) \log\left(\frac{1}{\mathbb{P}(r)}\right), \tag{10}$$

where $\mathbb{P}(r)$ is the posterior probability of point x being in region r.

Given that the magnitude of the entropy increases with the number of regions, we normalize the entropy and incorporate it as a factor in the evaluation of the acquisition function, resulting in

$$EJIE^H = \sum_{r \in \mathcal{R}} EI_r(x)\mathbb{P}_r(x)\left(1 + \frac{H(x)}{H_{max}}\right). \tag{11}$$

This modified acquisition function will prioritize points that not only show potential for improving a specific region, but also hold the potential for reducing errors in region membership predictions. This steers the algorithm towards more informative points near region boundaries, while still exploring high-performing regions.

Following [12], we perform progressive upscaling of the solution archive. That is, although we wish to predict solutions for a 40x40 solution archive, we begin by running BOIL over a 5x5 archive. The region filling nature of the acquisition function rapidly fills the 25 regions [1] and as they are spread out in descriptor space, the algorithm benefits from a model built on diverse points.

Another novel contribution of this work is a proposed stopping criterion. Once all regions in the solution archive have been filled, the scale of the acquisition function changes and the entropy calculation pushes the algorithm to balance exploring region boundaries and improving the solution set. Once the boundaries are well explored and the models predict that the solutions in the archive are near-optimal, the maximum acquisition function value drops substantially. We use this drop as an indicator of a well filled solution archive, and increase the resolution of the solution archive (upscaling) or involve the DM when the maximum acquisition function value falls below a certain threshold (in this work we use a threshold 0.1, and as the training values for the GP models are standardised, this worked well in multiple domains).

The upscaling process involves transitioning from a 5x5 grid to a 10x10 grid, which is efficient because many of the region boundaries remain the same and each elite in the 5x5 archive corresponds to a solution in the 10x10 archive. Further upscaling to a 20x20 and, if necessary, a 40x40 grid is possible. However, it should be noted that a domain with 40x40 reachable regions may prove challenging for a GP based method to fill, due to limitations in the model's capacity to handle a large number of points. The objective of the method, however, is not to fill the regions with observations, but rather to make predictions, and this can be achieved with a prediction map based on a small set of points.

4.2 Focused illumination

Fine granular models generated based on points gained from relatively coarse solution archives can provide rich insights into the mapping from descriptor space to best achievable fitness. The prediction map, which is derived from these models, enables a DM to gain an understanding of the emerging behaviour-fitness landscape very early on (i.e., after only very few evaluations). When the maximal acquisition function value drops, indicating good model building, we offer the DM the option to select regions of the behavioural space for targeted illumination. This allows for the DM to focus the search effort to the most relevant regions. Figure 1

[1] This is due to the expected improvement calculation comparing against an empty cell.

provides an example. The left part of the figure shows the solution archive after having evaluated only 309 solutions (100 initial solutions plus 209 steps of BOP-Elites iterations). Based on this information, it is possible to generate an informative prediction map displayed in Figure 1 (right). The DM can analyze the two parts of the figure and identify regions that seem particularly interesting and worth more exploration (rectangular areas).

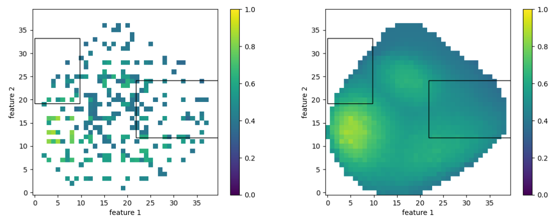


Figure 2: The observed solution archive (left) and the predicted solution archive (right) with the rectangular focus-areas that have been chosen for focused illumination (Section 4.2).

The BOP-Elites program continues to operate, but the calculations for $EJIE^{H}$ are limited to the areas specified by the user. A higher-resolution partition of regions is then created, usually at 20x20 resolution, consisting only of the regions within the selected focus areas. This way, BOP-Elites will continue to accumulate data points that will improve the model's performance specifically in these chosen areas of the descriptor space. The same termination criteria are applied, and model building stops when the acquisition function reaches the threshold. At this point, the DM is given the chance to evaluate the effect on the prediction map and determine whether to choose a new focus area, refine a current focus area, or move on to the next stage of the algorithm. Figure 3 compares the user-defined right focus-area from Figure 1 before focused illumination (left) and after (right). Only 57 new evaluations have been used to provide a much more granular illumination of the focus-area.

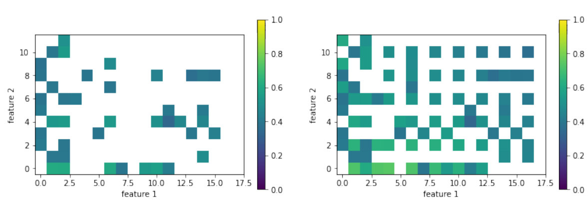


Figure 3: A chosen focus-area (the right focus-area from Figure 1) before optimization (left) and after 57 additional observations in the focus-area (right). The gaps are filled with high performing points, the prediction model will now be significantly improved in this region.

4.3 Target search

Once the DM has sufficiently explored the prediction map they may pick a solution with the desired predicted fitness and descriptor value combination, the target descriptor point $\mathcal{P}$. Our algorithm will then search for the best possible solution with the requested target behavior. However, since the behaviour is a characteristic of the solution and cannot be controlled explicitly, it is not possible to guarantee that we find a solution with a specific target behavior. Therefore, we additionally ask the DM for a trade-off parameter $\alpha \in [0, \infty]$ and optimize a linear combination of the solution's fitness and distance to the desired target behavior (in a sense, a *utility* $U(x)$), and we aim to maximize this utility:

$$\max_{x} U(x) = f(x) - \alpha||\mathcal{P} - g(x)||. \tag{12}$$

A larger α value will emphasize search for a point close to $\mathcal{P}$, while a value of $\alpha = 0$ assumes complete indifference regarding distance from the target and seeks a global fitness optimum.

If the DM does not provide a preference, we may solve the problem for a variety of α values and present the DM with a series of options with different trade-offs.

Unfortunately calculating the EI including the predicted distance calculation is not analytically tractable and so we must approximate this value by Monte Carlo sampling. We are able to sample from the fitness GP and descriptor GPs and posterior models at a point x are independent univariate Gaussians. In order to make each optimization run deterministic, we sample a number of z values from the standard normal distribution $Z \in \mathbb{N}(0, 1)$ and sample the posterior distributions at these z's.

$$\hat{y}(.) \sim \mathbb{N}(\mu_y, \sigma_y) \tag{13}$$

where μ_f and σ_f are the mean and standard deviation of the posterior predictive fitness model of the GP at x. We can take samples as:

$$\hat{y}(x) = z\sigma_y + \mu_y \tag{14}$$

The same approach is applied to the descriptor models. We take 1600 samples of the fitness evaluation and 40 for each of the descriptor models, recombining them to make 1600 predicted distances

$$\hat{\delta}(x) = \sqrt{\sum [\mathbb{P} - \hat{g}(x)]^2} \tag{15}$$

and calculate the utility sample as:

$$\hat{u} = \hat{y} - \alpha\hat{\delta}(x) \tag{16}$$

and the expected improvement in utility as:

$$EUI = \frac{1}{n}\sum_{i} \max\{\hat{u} - u^*, 0\} \tag{17}$$

where u^* is the utility of the best solution found so far.

This quantity can now be optimized with standard continuous optimizers. As each optimization run will have different z's, we run with multiple restarts and compare the final points of each run with a final set of z's to choose the best next point to sample. Once observed, the point is added to the model and the process continues until the budget has been spent.

5 EXPERIMENTAL RESULTS

5.1 Experimental design

We provide a usage scenario for BOIL over a 10 dimensional search space with fitness and descriptor functions generated from 3 independent randomly generated Gaussian process functions with randomised lengthscales in the range $[0, 1]$. This forms a synthetic function problem with 10 input dimensions over the unit hypercube, 1 fitness function and 2 descriptor functions.

Following the general wisdom on initial designs, $10d = 100$ points are selected via Sobol sampling. We define a [5x5] solution archive with boundaries uniformly distributed in descriptor space. The best points in each region become an elite and are added to the archive. We now run BOP-Elites using $EJIE^H$ until the acquisition function valuation drops below 0.1. At this stage we upscale the archive to a [10x10] grid defined over the entire descriptor space. Figures 4 and 5 provide an example of the upscaling process. Initially, BOIL attempts to fill a [5x5] archive, after 165 evaluations (100 initial and 65 BOIL iteration) resulting in Figure 4 (left). The actual distribution of the evaluated points can be better seen in Figure 4 (right), which uses a [40x40] grid. Upscaling to the [10x10] results in the archive depicted in Figure 5 (left), and after an additional 144 evaluations, this archive now looks as in Figure 5 (right).

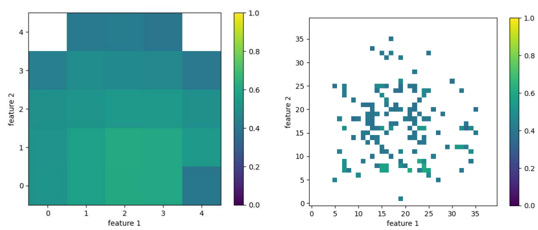


Figure 4: A visualisation of the solution archive at two different resolutions. (left) is the [5x5] solution archive when the acquisition value drops. This happens after 65 steps of the BOIL algorithm, (right) is the [40x40] solution archive which is sparsely populated with the same points.

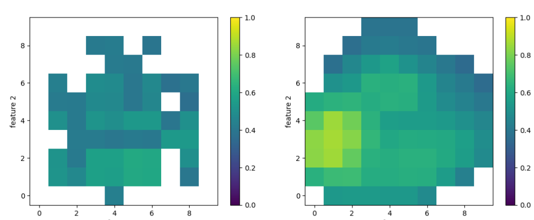


Figure 5: (left) is the [10x10] solution archive at the same stage as Figure 4 after 65 steps of the BOIL algorithm, (right) is the [10x10] solution archive after an additional 144 evaluations. This is now ready to upscale to the [20x20] resolution as indicated by the drop in maximal acquisition function value

We continue the illumination stage until the acquisition drops below 0.5 for the [10x10] grid. At this stage we visualise the prediction map at the [40x40] resolution. We now simulate a DM creating 2 focus-areas in descriptor space and use focused illumination over these areas (see rectangles in Figure 6). Focused illumination occurs in the next available resolution level, given we are currently exiting the [10x10] scale we create focus-areas at the [20x20] resolution. The same stopping criteria apply and we perform this action on both focus-areas.

BOIL produces a good general model from the illumination stage and has benefited from the focused illumination on two areas of interest. Next, we simulate the decision-maker selecting a single point with a continuous descriptor value and a single alpha value. We perform 20 steps with a continuous optimizer over the acquisition function (Equation 17) and return the best solution found.

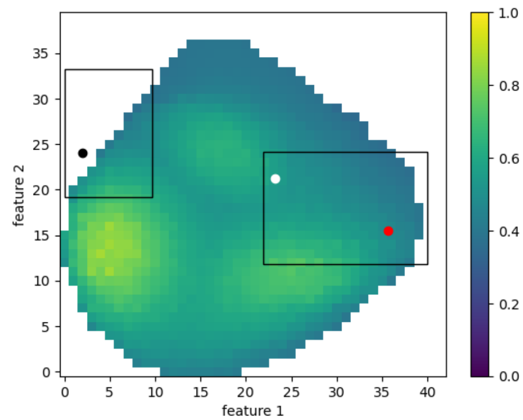


Figure 6: A visualisation of 3 potential target points selected by the DM. As the black point in the left region shows, the DM's target point need not be in reachable space. The α value will define their preference over the trade off between proximity to the point and fitness.

We compare the performance of BOIL's final output against the MAP-Elites algorithm. In our experiments BOIL never exceeded 350 additional observations (450 total) until it has converged to a solution and we give MAP-Elites the same budget over a 20x20 solution archive to aid it in covering the descriptor space. As Figure 7 shows visually, the solution archive created by MAP-Elites using at least as many evaluations as BOIL is less diverse and of lower quality. In order to select a final point, we search the resulting MAP-Elites solution archive for the point which provides the best utility given the DMs preferences, and report on the average utility, fitness and distance from target point obtained.

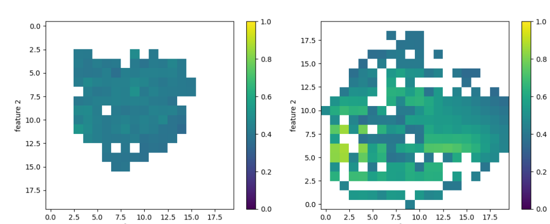


Figure 7: Visualization of MAP-Elites solution archive (left). As MAP-Elites was run at the [20x20] resolution, we compare the observation set of BOIL (Right) at the same resolution after convergence.

5.2 Empirical results

The regions, points and alphas are not known a priori to either algorithm. 2 rectangular regions were generated and 3 target points were randomly chosen from within the focus areas to simulate the behaviour of a DM. In order to produce a clear comparison we run the same experiment with these same regions and points and declare the average performance.

Point	α	**BOIL** Avg.	**MAP-Elites** Avg.
Black	3	0.4790 ±0.0001	0.1562 ±0.0026
White	0.2	0.8407 ±0.0002	0.4116 ±0.0039
Red	1	0.7670 ±0.0007	0.4155 ±0.0124

Table 1: Average utility of the returned points and the standard error (SE) around the mean by BOIL and MAP-Elites for the single target problem, colors of points as in Figure 6, average over 30 runs

As can be seen in Table 1, the BOIL algorithm finds a far better solution in terms of ultility compared to MAP-Elites in every case, targeting an optimal trade-off between the DM's preference for fitness and distance from descriptors. The utilities are a combination of fitness and distance and we report the average performance in each metric in Table 2.

Figure 8 visually displays, for each of the three target point examples (black, white and red), the returned solutions by the two algorithms. For the black example with $\alpha = 3$ (a strong penalty for distance from the target), both algorithms find solutions relatively close, but the one found by BOIL is closer. For the red target point with $\alpha = 1$, the points returned are much further away, but of a much better fitness (lighter green) than potential alternative points closer to the target. The point returned by BOIL is not only closer, but also has higher fitness. Finally, in the example of the white target the DM indicated that distance to the target is not very important by setting $\alpha = 0.2$, and correspondingly solutions are far away once more. BOIL now seems to have returned a point close to the global optimum, whereas the points returned by MAP-Elites are significantly worse. Note that the figure shows several returned solutions for MAP-Elites in the red and white example, from different runs. On the other hand, BOIL consistently returns the same solution.

	Black $\alpha = 3$ Avg.	**White** $\alpha = 0.2$ Avg.	**Red** $\alpha = 1$ Avg.
BOIL fit	0.481 ±0.0004	0.928 ±0.001	0.805 ±0.0003
BOIL Dist.	0.001 ±0.00002	0.437 ±0.003	0.038 ±0.0007
ME Fit.	0.577 ±0.015	0.451 ±0.005	0.777 ±0.010
ME Dist.	0.140 ±0.007	0.197 ±0.013	0.362 ±0.012

Table 2: Average fitnesses and distances with standard errors (SE) from the target for the single target problem, colors of points as in Figure 6

Table 2 further confirms that BOIL, on average, finds both higher fitness points and closer solutions for the higher values of *alpha*.

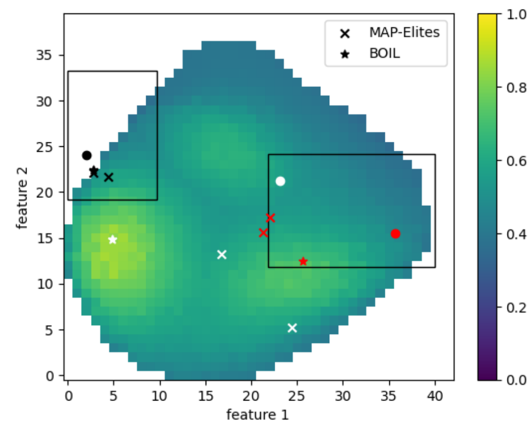


Figure 8: A selection of indicative points from the returned solutions.

When *alpha* = 0.2, for the white target, BOIL finds a point closer to the global maximum and cares less about distance, as per the preference of the DM.

6 CONCLUSION AND FUTURE WORK

We presented a novel, interactive optimization strategy that empowers a decision maker to understand the diversity in behavioural space and the effect of these variations on the maximum achievable fitness. By incorporating human interaction and prioritizing efficient optimization, our approach outperforms MAP-Elites with a subsequent selection from the generated archive and provides the DM with assurance that the behavior space has been thoroughly explored in the most promising regions identified. Additionally, this strategy exhibits excellent efficiency in terms of the number of samples required.

Our method incorporates a powerful entropy based adaptation to the BOP-Elites acquisition function which enables better model building and an interactive scaling method for focusing exploration on regions of interest.

The final point returned to the DM is found by searching using a Monte Carlo acquisition function that returns the best trade-off between fitness and distance from the DM's target descriptor given the DM's preferences.

This novel and effective optimization technique integrates concepts from Bayesian Optimization, Quality-Diversity, and interactive Multi-Objective Optimization. For future work, multi-objective Bayesian optimization (e.g., a variant of [4]) could be used to directly provide the DM with a set of non-dominated solutions regarding fitness and distance to the DM's target descriptor, alleviating the need for the DM to specify a weighting factor. In our paper, we have only considered the use case where the DM will eventually pick a single solution. However, the approach can also be applied in a lifelong learning scenario, where different solution behaviors are required at different points in time, and the BOIL can be used to quickly and precisely generate the specific behavior required each time, accumulating knowledge over time. Finally, further investigation should also consider a wider range of benchmarks and domains, including real-world applications.

ACKNOWLEDGMENTS

The first author gratefully ackowledges support by the Engineering and Physical Sciences Research Council and Medical Research Council through the Mathematics of Systems I Centre for Doctoral Training at the University of Warwick (reference EP/L015374/1).

REFERENCES

[1] Bekir Afsar, Kaisa Miettinen, and Francisco Ruiz. 2021. Assessing the performance of interactive multiobjective optimization methods: A survey. *ACM Computing Surveys (CSUR)* 54, 4 (2021), 1–27.

[2] Alberto Alvarez, Steve Dahlskog, Jose Font, and Julian Togelius. 2019. Empowering quality diversity in dungeon design with interactive constrained map-elites. In *2019 IEEE Conference on Games (CoG)*. IEEE, 1–8.

[3] A. Cully, J. Clune, D. Tarapore, and J.-B. Mouret. 2015. Robots that can adapt like animals. *Nature* 521, 7553 (2015), 503–507.

[4] Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. 2021. Parallel bayesian optimization of multiple noisy objectives with expected hypervolume improvement. *Advances in Neural Information Processing Systems* 34 (2021), 2187–2200.

[5] Matthew C Fontaine, Ruilin Liu, Julian Togelius, Amy K Hoover, and Stefanos Nikolaidis. 2021. Illuminating mario scenes in the latent space of a generative adversarial network. In *AAAI Conference on Artificial Intelligence*.

[6] A. Gaier, A. Asteroth, and J.-B. Mouret. 2018. Data-efficient design exploration through surrogate-assisted illumination. *Evolutionary Computation* 26, 3 (2018), 381–410.

[7] Daniele Gravina, Ahmed Khalifa, Antonios Liapis, Julian Togelius, and Georgios N Yannakakis. 2019. Procedural content generation through quality diversity. In *2019 IEEE Conference on Games (CoG)*. IEEE, 1–8.

[8] Alexander Hagg, Dominik Wilde, Alexander Asteroth, and Thomas Bäck. 2020. Designing air flow with surrogate-assisted phenotypic niching. In *Parallel Problem Solving from Nature–PPSN XVI: 16th International Conference, PPSN 2020, Leiden, The Netherlands, September 5-9, 2020, Proceedings, Part I 16*. Springer, 140–153.

[9] Andrzej Jaszkiewicz and Jürgen Branke. 2008. Interactive multiobjective evolutionary algorithms. *Multiobjective optimization: interactive and evolutionary approaches* (2008), 179–193.

[10] D. R. Jones, M. Schonlau, and W. J. Welch. 1998. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization* 13, 4 (1998).

[11] Paul Kent and Juergen Branke. 2020. Bop-elites, a bayesian optimisation algorithm for quality-diversity search. *arXiv preprint arXiv:2005.04320* (2020).

[12] P. Kent, A. Gaier, J.-B. Mouret, and J. Branke. 2023. BOP-Elites, a Bayesian Optimisation Approach to Quality Diversity Search with Black-Box descriptor functions. (2023). to be published.

[13] Hee-Su Kim and Sung-Bae Cho. 2000. Application of interactive genetic algorithm to fashion design. *Engineering applications of artificial intelligence* 13, 6 (2000), 635–644.

[14] Bryan Lim, Maxime Allard, Luca Grillotti, and Antoine Cully. [n. d.]. Accelerated Quality-Diversity through Massive Parallelism. *Transactions on Machine Learning Research* ([n. d.]).

[15] Jonas Mockus, Vytautas Tiesis, and Antanas Zilinskas. 1978. The application of Bayesian methods for seeking the extremum. *Towards Global Optimization* 2, 117-129 (1978), 2.

[16] Jean-Baptiste Mouret. 2020. Evolving the behavior of machines: from micro to macroevolution. *Iscience* 23, 11 (2020), 101731.

[17] J.-B. Mouret and J. Clune. 2015. Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909* (2015).

[18] Selcen Phelps and Murat Köksalan. 2003. An interactive evolutionary metaheuristic for multiobjective combinatorial optimization. *Management Science* 49, 12 (2003), 1726–1738.

[19] Justin K Pugh, Lisa B Soros, and Kenneth O Stanley. 2016. Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI* (2016), 40.

[20] Hideyuki Takagi. 2001. Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation. *Proc. IEEE* 89, 9 (2001), 1275–1296.

[21] Jonas Verhellen and Jeriek Van den Abeele. 2020. Illuminating elite patches of chemical space. *Chemical Science* 11, 42 (2020), 11485–11491.