



Modelling human active search in optimizing black-box functions

Antonio Candelieri¹ · Riccardo Perego¹ · Ilaria Giordani¹ · Andrea Ponti¹ · Francesco Archetti¹

Published online: 24 October 2020
© The Author(s) 2020

Abstract

Modelling human function learning has been the subject of intense research in cognitive sciences. The topic is relevant in black-box optimization where information about the objective and/or constraints is not available and must be learned through function evaluations. In this paper, we focus on the relation between the behaviour of humans searching for the maximum and the probabilistic model used in Bayesian optimization. As surrogate models of the unknown function, both Gaussian processes and random forest have been considered: the Bayesian learning paradigm is central in the development of active learning approaches balancing exploration/exploitation in uncertain conditions towards effective generalization in large decision spaces. In this paper, we analyse experimentally how Bayesian optimization compares to humans searching for the maximum of an unknown 2D function. A set of controlled experiments with 60 subjects, using both surrogate models, confirm that Bayesian optimization provides a general model to represent individual patterns of active learning in humans.

Keywords Bayesian optimization · Cognitive models · Active learning · Search strategy

1 Introduction

We consider as reference problem the black-box optimization: the objective function and/or constraints are analytically unknown and evaluating them might be very expensive and noisy. In black-box situations, as we cannot assume any prior knowledge about the objective function $f(x)$, any functional form is a priori admissible and the value of the function at a point says nothing about the value at other points: the only way to develop a problem specific algorithm is to assume a model of $f(x)$ and to learn through a sample of function values.

Such an algorithm must be sample efficient because the cost of function evaluations is the dominating cost. This problem has been addressed in several fields under different names, including active learning (Kruschke et al. 2008; Griffiths et al. 2008; Wilson et al. 2015), Bayesian optimization (BO) (Zhigljavsky and Zilinskas 2007; Candelieri et al. 2018; Archetti and Candelieri 2019), hyperparameter

optimization (Eggenesperger et al. 2019), Lipschitz global optimization (Sergeyev et al. 2020a, b; Lera et al. 2021), and others.

In BO, a probabilistic surrogate model of the objective function is built to sum up our a priori beliefs about the objective function and the informative value of new observations. Two probabilistic frameworks are usually considered: the Gaussian processes (GPs) and random forests (RF) which offer alternative ways to update the beliefs as new data arrives and to provide an estimate of the expected value of the objective function and the uncertainty in this estimate.

Both GP and RF provide surrogate models: the major driver of the choice is the type of design variables: continuous ones are better dealt with GP, while integer/categorical and conditional ones with RF.

In both cases the next sampled point is chosen on the basis of its informative value through the maximization of an acquisition function: this choice brings up the so-called exploration vs exploitation dilemma, where exploration means devoting resources to know more about possible solutions while exploitation devotes resources to improve on solutions already identified in previous phases. The search for the new point must strike an effective balance between the needs of exploration and exploitation.

Communicated by Yaroslav D. Sergeyev.

✉ Antonio Candelieri
antonio.candelieri@unimib.it

¹ University of Milano-Bicocca, 20126 Milan, Italy

Psychologists have extensively studied how humans balance exploration and exploitation (Krusche et al. 2008; Mehlhorn et al. 2015), with a recent attention to the links between modern machine learning algorithms and psychological processes. (Gershman 2018; Schulz et al. 2016; Gopnik et al. 2017). Psychological research has so far mostly focused on how people learn functions according to a protocol in which an input is presented to participants and they are asked to predict the corresponding output value. Then, they observe the true output value (usually noisy) in order to update their own “predictive model” that is to adjust their internal representation of the underlying function. Psychologists have largely focused on GP: the issues of GP regression, kernel composition for different degrees of smoothness and safe optimization in their relation to cognition is studied in a recent survey by Schulz et al. (2018b). Exploration is realized by adding the so-called *uncertainty bonus* to estimated values obtaining the *upper confidence bound* (UCB) algorithm (Srinivas et al. 2010). In Wu et al. (2018) the human search strategy is analysed for rewards under limited search horizons, concluding that GP offers the best model for generalization and UCB the best solution of the exploration/exploitation dilemma.

A significant application of RF is given in Plonsky et al. (2019) as a hybrid model of machine learning and decision mechanisms. A key driver in the above research activities is that human learners are increasingly fast at adapting to unfamiliar environments. Psychologists are investigating the intriguing gap between the capabilities of human and machine learning.

Most previous research findings in human learning refer to function learning because it is related to a probabilistic perspective on predictability and provides a proxy to generalization capability. Contrary to function learning, optimization is not yet widely considered in the literature; in Borji and Itti (2013) a simple 1-D optimization problem has been considered.

The approach presented in this paper has been sketched in Candelieri et al. (2019). The present paper has been significantly augmented and rewritten. A set of new computational results are related to the use, along with the GP, of the RF as surrogate model. The set of references has been also enlarged, and the whole perspective has been widened to reflect that: learning and optimization of black-box functions are two faces of the same coin. Moreover, other global optimization strategies have been also considered in this study, both deterministic and stochastic, leading to the main result that BO is the one offering a reasonable unifying framework of human function learning, active sampling and optimization.

The structure of the paper is as follows: Sect. 2 outlines the methodological background of BO including the basis of the two main surrogate models, GP and RF, and the

management of the exploration/exploitation dilemma. Section 3 is devoted to the experimental set-up, and Sect. 4 reports the experimental results about the behavioural patterns of humans in optimizing black-box functions, compared to both BO and other five global optimization strategies. Section 5 outlines the conclusions of this study and the perspectives of future works.

2 Methodological background

This section provides the underlying methodological framework of the study. The global optimization problem we consider is defined as:

$$\max_{x \in \chi \subset \mathbb{R}^d} f(x)$$

where the search space χ is generally box-bounded and $f(x)$ is black-box meaning that no gradient information is available and that we have only access to noisy observations of $f(x)$ which are computationally expensive.

2.1 Gaussian processes

GPs are a powerful nonparametric model for implementing both regression and classification. One way to interpret a GP is as a distribution over functions, with inference taking place directly in the space of functions (Williams and Rasmussen 2006). A GP, therefore, is a collection of correlated random variables, any finite number of which have a joint Gaussian distribution. A GP is completely specified by its mean function $\mu(x)$ and covariance function $\text{cov}(f(x), f(x')) = k(x, x')$:

$$\begin{aligned} \mu(x) &= \mathbb{E}[f(x)] \\ \text{cov}(f(x), f(x')) &= k(x, x') = \mathbb{E}[(f(x) - \mu(x))(f(x') - \mu(x')))] \end{aligned}$$

and will be denoted by: $f(x) \sim GP(\mu(x), k(x, x'))$. This means that the behaviour of the model can be controlled entirely through the mean and covariance.

Usually, for notational simplicity we will take the prior of the mean function to be zero, although this is not necessary. The covariance function assumes a critical role in the GP modelling, as it specifies the distribution over functions, depending on a sample $X_{1:n} = \{x_1, \dots, x_n\}$. More precisely, $f(X_{1:n}) \sim \mathcal{N}(0, K(X_{1:n}, X_{1:n}))$ with $K(X_{1:n}, X_{1:n})$ an $n \times n$ matrix whose entry $K_{ij} = k(x_i, x_j)$ with k a kernel function specifying a prior about the smoothness of the function to be approximated and optimized.

We usually have access only to noisy function values, denoted by $y = f(x) + \varepsilon$, with ε an additive independent identically distributed Gaussian noise with variance λ^2 .

Let $y = (y_1, \dots, y_n)$ denote a set of n noisy function evaluations, then the resulting covariance matrix becomes $K(X_{1:n}, X_{1:n}) + \lambda^2 I$.

Let $D_{1:n} = \{(x_i, y_i)\}_{i=1, \dots, n}$ denote the set containing both the locations and the associated function evaluations, then the predictive equations for GP regression are updated by conditioning the joint Gaussian prior distribution on $D_{1:n}$:

$$\mu(x) = \mathbb{E}[f(x)|D_{1:n}, x] = k(x, X_{1:n})[K(X_{1:n}, X_{1:n}) + \lambda^2 I]^{-1} y$$

$$\sigma^2(x) = k(x, x) - k(x, X_{1:n})[K(X_{1:n}, X_{1:n}) + \lambda^2 I]^{-1} k(X_{1:n}, x)$$

where $k(x, X_{1:n}) = [k(x, x_1), \dots, k(x, x_n)]$.

The covariance function is the crucial ingredient in a GP predictor, as it encodes assumptions about the function to approximate: function evaluations that are near to a given point should be informative about the prediction at that point. Under the GP view, it is the covariance function that defines nearness or similarity. Once the prior mean and the kernel are chosen, they are updated with the observation of $f(x)$ to find a-posterior distribution $f(x|D_{1:n})$ and this allows us to find the expected value of the function at any point and to calculate its uncertainty through its predicted variance.

Examples of covariance (aka kernel) functions:

Squared Exponential (SE) kernel: $k_{SE}(x, x') = e^{-\frac{x-x'}{\ell}} e^{-\frac{(x-x')^2}{2\ell^2}}$.

Exponential kernel: $k_{Exp}(x, x') = e^{-\frac{|x-x'|}{\ell}}$.

Power exponential kernel: $k_{PowExp}(x, x') = e^{-\left(\frac{|x-x'|}{\ell}\right)^p}$.

Matérn kernels: $k_{Mat}(x, x') = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{|x-x'|\sqrt{2\nu}}{\ell}\right)^\nu K_\nu\left(\frac{|x-x'|\sqrt{2\nu}}{\ell}\right)$.

The hyperparameter ℓ , in all the reported kernels, is the characteristic length-scale (updated by maximum likelihood estimation). The hyperparameter $\nu > 0$ in the Matérn kernel governates the smoothness of the GP samples, which are $\nu - 1$ differentiable, and where $\Gamma(\nu)$ is the gamma function and K_ν is the modified Bessel function of the second kind. The most widely adopted versions, specifically in the machine learning community and considered in this paper, are $\nu = 3/2$ and $\nu = 5/2$. The Matérn kernel encodes the expected smoothness of the target function explicitly.

2.2 Random forest

Random forest (RF) is an ensemble learning method, based on decision trees, for both classification and regression problems (Ho 1995). According to the originally proposed implementation, RF aims at generating a multitude of decision trees, at training time, and providing as output the mode of the classes (classification) or the mean/median of the predictions (regressions) of the individual trees.

Although originally designed and presented as a machine learning algorithm, RF is also an effective and efficient alternative to GP for implementing BO. RF consists of an ensemble of different regressors (i.e. decision trees); it is possible to compute—as for GP—both $\mu(x)$ and $\sigma(x)$, simply as mean and variance of the samples of the individual outputs provided by the regressors. Due to the different nature of RF and GP, they result in significantly different surrogate models. While GP is well suited to model smooth functions in search space spanned by continuous variables, RF can also deal with discrete and conditional variables.

Moreover, fitting an RF is computationally more efficient than fitting a GP; indeed, the kernel matrix inversion required to fit a GP is a well-known computational issue.

2.3 The acquisition functions

The acquisition function is the mechanism to implement the trade-off between exploration and exploitation in BO. More precisely, any acquisition function aims to guide the search of the optimum towards points with potentially high values of objective function either because the prediction of $f(x)$, based on the probabilistic surrogate model, is high or the uncertainty, also based on the same model, is high (or both). Indeed, exploitation means to consider the area providing more chance to improve over current solution, while exploration means to move towards less explored regions of the search space where predictions based on the surrogate model are more uncertain, with higher variance. There are many acquisition functions, we quote only those used in this study. *Probability of improvement* (PI) (Kushner 1964) and *expected improvement* (EI) (Mockus 1975) measure, respectively, the probability and the expectation of the improvement over the best observed value of $f(x)$ given the predictive distribution of the probabilistic surrogate model. More precisely, they are defined as follows:

$$EI(x) = \max_{x \in X} \{0, \mu(x) - y^+\}$$

$$PI(x) = \Phi\left(\frac{\mu(x) - y^+}{\sigma(x)}\right)$$

with y^+ the best value observed so far and Φ is the cumulative density function of a normal distribution. The next evaluation point, x_{n+1} , is obtained by maximizing $EI(x)$ or $PI(x)$. More recently, *Upper/Lower Confidence Bound*, (Srinivas et al. 2010), denoted with UCB/LCB, is widely used. It is an acquisition function that manages exploration–exploitation by being optimistic in the face of uncertainty where UCB and LCB are used, respectively, for maximization and minimization problems:

$$UCB(x) = \mu(x) + \beta\sigma(x) \text{ and } LCB(x) = \mu(x) - \beta\sigma(x)$$

where $\beta \geq 0$ is the parameter to manage the trade-off between exploration and exploitation ($\beta = 0$ is for pure exploitation; on the contrary, higher values of β emphasizes exploration by inflating the weight of model uncertainty). In Srinivas et al. (2010), a policy is provided for updating the value of β along function evaluations, with also a proof of convergence of such a policy. In the case of a maximization problem, the next point is chosen as $x_{n+1} = \operatorname{argmax}_{x \in X} UCB(x)$ while in the case of a minimization problem the next point is selected as $x_{n+1} = \operatorname{argmin}_{x \in X} LCB(x)$.

2.4 Bayesian optimization

Algorithm 1 summarizes a general BO schema where the acquisition function, whichever it is, is denoted by $\alpha(x, D_{1:m})$. In order to maintain the schema as general as possible we do not specify the probabilistic surrogate model, as well as the kernel in the case of a GP.

In this study we have used both GP (considering the kernel types presented in the previous section) and RF as surrogate probabilistic models. The three different acquisition functions previously described have been used for both the two surrogates.

functions. Each subject was informed about the goal of the experiment and the available number of clicks for the play, before to start. Subjects started by clicking at any point in the screen and getting the corresponding value; previously clicked points and their values remained on the screen until the end of the trial game. More specifically, points are coloured and resized according to the associated score (i.e. the value of f at that point), providing a visual feedback about the distribution of the scores collected so far.

Moreover, the software has been developed to manage three different game modalities. At the start of each game, along with the test function, also the modality is randomly chosen among the following.

Find a point whose associated evaluation is close to $f^* = f(x^*)$, but without knowing f^* a-priori. From an optimization point of view, this means that the human player searches for $y_n^+ = \max_{i=1, \dots, n} y_i$, with $y_i = f(x_i)$ the value observed for the i th location sequentially chosen by the player. This is equivalent to optimize the *simple regret* (without knowing f^*):

$$\min_{n=1, \dots, N} f^* - y_n^+$$

As in the previous game modality, the goal is to find a point whose associated evaluation is close to f^* but, in this case, the human player knows f^* before making his/her choices. Thus, also in this case the player will optimize the *simple regret*: we are interested in understanding if and

Algorithm 1. General Bayesian Optimization Algorithm

Generate an initial set of m points $X_{1:m}$ randomly sampled (e.g., via Latin Hypercube Sampling)

Evaluate the function in the initial set of points and obtain $D_{1:m}$

Define a further budget N

1 **for** $n = m + 1, \dots, m + N$ **do**

2 update the surrogate model obtaining the new estimates of $\mu(x)$ and $\sigma(x)$

3 select a new x_{n+1} by optimizing an acquisition function α , such that

$$x_{n+1} = \operatorname{argmax}_x \alpha(x|D_{1:n})$$

4 evaluate the objective function to obtain $y_{n+1} = f(x_{n+1})$

5 update the dataset of observations $D_{1:n+1} = D_{1:n} \cup \{(x_{n+1}, y_{n+1})\}$

6 **endfor**

7 Output: the best y value observed over the entire optimization process

3 Experimental set-up

3.1 User interface

The software related to the game playing has been developed by implementing 15 stimuli, that are 15 different 2D

how knowledge about f^* can modify humans' strategies with respect to the first game modality.

Maximize the cumulative value of all the evaluations at the selected points, without knowing a-priori f^* . In this case human players tries to maximize $\sum_{i=1}^N y_i$, which is equivalent to optimize the *cumulative regret*.

$$\min \sum_{i=1}^N (f^* - y_i)$$

Our software collects all the choices made by each player along every game play and stores data into a database whose structure is summarized in Fig. 1. Into the “games” table, each row (i.e. record) represents a single point, and the associated function value, in a game, along with information about the player’s identifier, the test function, and the game modality. The games are univocally identified by including the timestamp relative to the end of the game.

Figure 2 shows a frame of the game.

Moreover, for the GP, five kernels have been used to consider different possibility of smoothness approximation of the objective function.

The choices of the human player and those of the global optimization algorithm are considered compliant, pointwise, if the distance between the point chosen by the human player and the “algorithmic player” is less than a given “threshold”, namely τ . Finally, the strategy of the human player is assimilated to the acquisition function most frequently compliant, pointwise, along a trial game. The procedures for GP-based and RF-based BO are summarized in the following pseudo-codes:

Algorithm 2. Algorithm for compliance analysis with respect to GP-based BO

Let denote by:

- p a participant
- k a kernel
- α an acquisition function
- n a generic iteration
- $s^{p,k,n}$ a search strategy (i.e., an acquisition function, under a GP with a given kernel, at a specific iteration)

```

1 foreach  $p, k$  and  $n$ 
2   fit a GP  $GP^{p,k,n}$ 
3 foreach  $\alpha$ 
4   select  $x_{n+1}^{GP^{p,k,n}}$ 
5   compute  $d^{p,k,n,\alpha} = \left\| x_{n+1}^{GP^{p,k,n}} - x_{n+1}^p \right\|$ , where
       $x_{n+1}^{GP^{p,k,n}} = \operatorname{argmax}_x \alpha(x) | GP^{p,k,n}$  is the next point according to acquisition function  $\alpha$ 
      under  $GP^{p,k,n}$ 
      and  $x_{n+1}^p$  is the next point chosen by participant  $p$  at iteration  $n + 1$ 
6 endforeach
7  $\bar{d}^{p,k,n} = \min_{\alpha} \{d^{p,k,n,\alpha}\}$ 
8 if  $\bar{d}^{p,k,n} \leq \tau$  then
9    $s^{p,k,n} = \operatorname{argmin}_{\alpha} \{d^{p,k,n,\alpha}\}$ 
10 else
11    $s^{p,k,n} = \emptyset$ 

```

3.2 Procedure

For each player, at each iteration, GP and RF models are fitted on the observed points, and then the three acquisition functions, previously mentioned in Sect. 2.3, are used to select the next point to query. All these points are compared, via Euclidean distance, to the corresponding choice made by the human player.

Finally, for a given kernel \bar{k} , the search strategy of the participant \bar{p} is compliant to the most frequent acquisition function in the series $s^{p,k} = \{s^{p,k,n}\}_{n=m:m+N}$.

Algorithm 3. Algorithm for compliance analysis with respect to RF-based BO

```

Let denote by:
-  $p$  a participant
-  $\alpha$  an acquisition function
-  $n$  a generic iteration
-  $s^{p,n}$  a search strategy (i.e., an acquisition function)

1 foreach  $p$  and  $n$ 
2   fit a Random Forest  $RF^{p,n}$ 
3 foreach  $\alpha$ 
4   select  $x_{n+1}^{RF^{p,n}}$ 
5   compute  $d^{p,n,\alpha} = \|x_{n+1}^{RF^{p,n}} - x_{n+1}^p\|$ , where
       $x_{n+1}^{RF^{p,n}} = \operatorname{argmax}_x \alpha(x) | RF^{p,n}$  is the next point according to acquisition function  $\alpha$  under  $RF^{p,n}$ 
      and  $x_{n+1}^p$  is the next point chosen by participant  $p$  at iteration  $n + 1$ 
6 endforeach
7  $\bar{d}^{p,n} = \min_{\alpha} \{d^{p,n,\alpha}\}$ 
8 if  $\bar{d}^{p,n} \leq \tau$  then
9    $s^{p,n} = \operatorname{argmin}_{\alpha} \{d^{p,n,\alpha}\}$ 
10 else
11    $s^{p,n} = \emptyset$ 

```

Finally, the search strategy of the participant \bar{p} is compliant to the most frequent acquisition function in the series $s^{\bar{p}} = \{s^{p,n}\}_{n=m:m+N}$.

3.3 Software resources and analysis

Software resources consist of two components developed in R. The first component is the procedure to compute the distance between the humans' and global optimization choices during each game, while the second aggregates all

the calculated distances and generates the related statistics. More precisely, the procedure operated by the first component is summarized in previous Algorithm 2 and Algorithm 3 for GP-based and RF-based BO, respectively. Since this study considers also other global optimization strategies—detailed in the following—the first component has been specialized for each one of the other strategies. (For ease of reading, we do not report every single implementation, but they can be easily obtained starting from Algorithm 3.)

Fig. 1 Entity–relation (ER) diagram of database

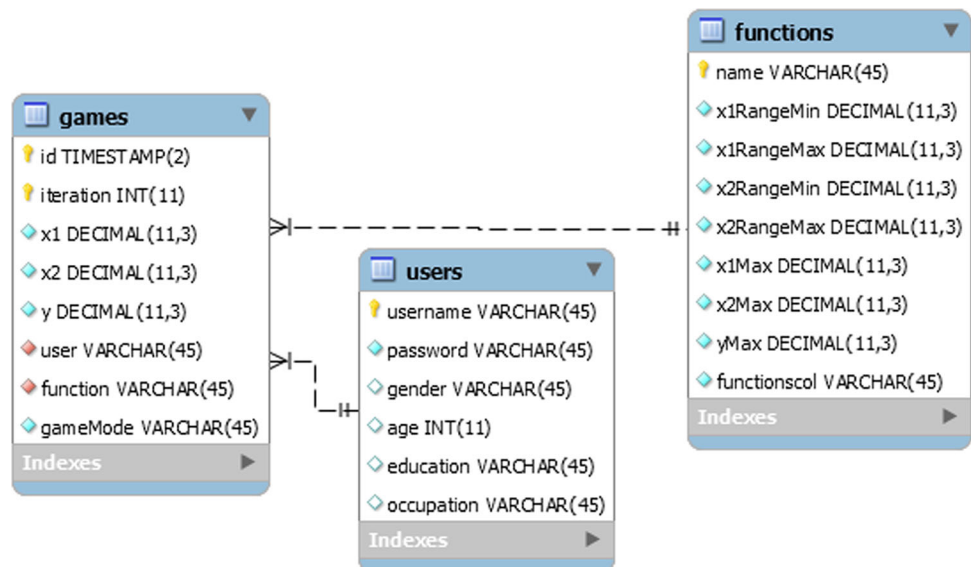
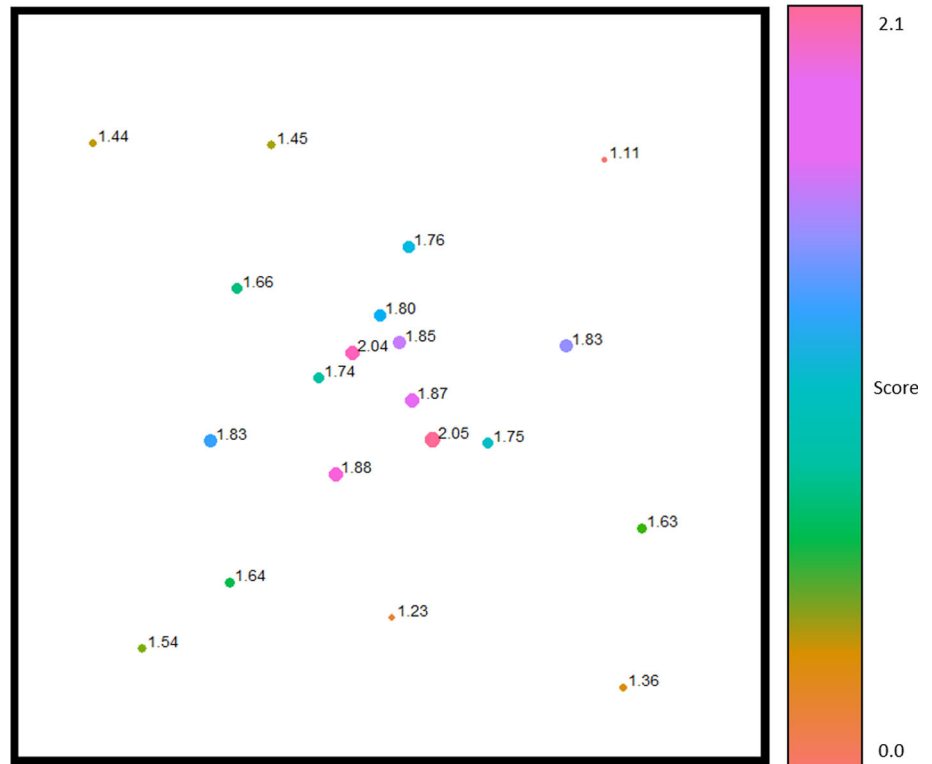


Fig. 2 An example of a game play: each dot is a location selected by the player. As a feedback, the associated numerical value is visualized, and size and colour of the dot are set accordingly



The implementations of Algorithm 2 and Algorithm 3 are based on the R package named “mlrMBO”, which offers both GP and RF as surrogate model along with the acquisition functions adopted in this study. L-BFGS algorithm is used to optimize the acquisition function based on GP on a continuous search space; for RF the “focus-search” algorithm (Bischl et al. 2017) is adopted: it can handle with numeric, discrete and mixed search spaces, also involving conditional variables. Focus-search starts with a large set of random points where the acquisition function is evaluated. Then, the search space is shrunk around the current best point and a new random sampling is performed within the “focused space”. Shrinkage is iteratively performed until a maximum number of iterations, and the entire procedure can be restarted multiple times to mitigate the risk of convergence to a local optimum. Finally, the best point over all restarts and iterations is returned as the solution.

As already mentioned, we have also compared the humans’ strategies with other deterministic and stochastic global optimization techniques, more specifically:

- *Random search (RS)* we used the R package “random-search” which implements a simple RS function.
- *DIRECT* we used the R package “nlopt” for nonlinear optimization, offering, among others, DIRECT as an optimization algorithm based on Jones et al. (1993).

- *Genetic algorithm (GA)* we used the R package “GA”, a toolbox implementing GA for stochastic optimization (Scrucca 2013, 2016).
- *Particle swarm optimization (PSO)* we used the R package “pso” which offers a standard PSO implementation based on Banks et al. (2007).
- *Simulated annealing (SA)* we used the R package “optim_sa” which offers an implementation of SA based on Laarhoven and Aarts (1987).

We have then computed the distances, in terms of x , between the best location found by the humans at the end of each game with those provided by each one of the other five global optimization techniques and BO. Let denote with x^+ the location found by the player (human or algorithmic) such that $y^+ = f(x^+)$ is the best value observed over a specific trial game by that user.

3.4 Participants and analytical settings

In this section, we further detail the features of the data collected and used for this study. Initially, around 70 volunteers were enrolled for the study; while they used the software to play and collect data, their feedbacks about the user-friendliness of the interface were also used to improve our software. Due to this *work-in-progress* nature of the game playing software, not all the games collected were

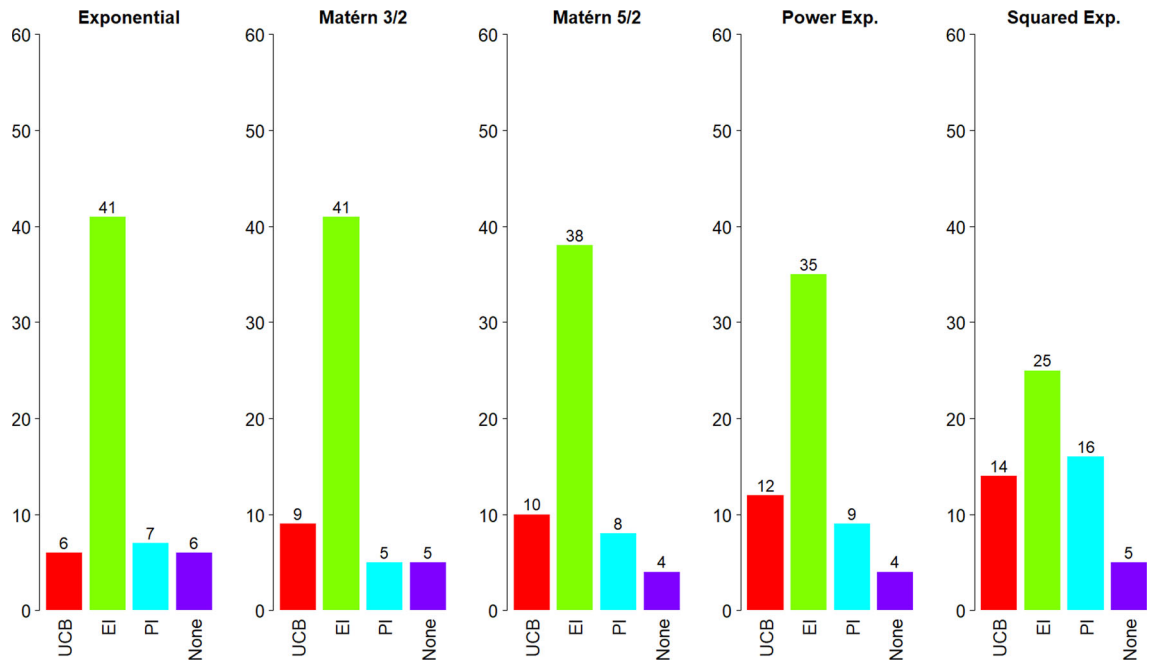


Fig. 3 Number of human players whose strategy is compliant with respect to kernel type and acquisition functions, with the threshold τ set to 0.10, and $\beta = 1$ in UCB. Last bar represents the number of non-compliant

Table 1 Fisher’s test p values for each pair related to the most frequently compliant acquisition function in Fig. 3 (i.e. EI) and the others

| | EI versus UCB | EI versus PI | EI versus none |
|---------------------|---------------|--------------|----------------|
| Exponential | < 0.001 | < 0.001 | < 0.001 |
| Matérn 3/2 | < 0.001 | < 0.001 | < 0.001 |
| Matérn 5/2 | < 0.001 | < 0.001 | < 0.001 |
| Power exponential | 0.758 | < 0.001 | < 0.001 |
| Squared exponential | 0.359 | 0.375 | 0.010 |

The hypothesis that one acquisition function is more significantly compliant than the other is accepted with a confidence 0.05

well suited for the analysis. Thus, we have decided to restrict the analysis to the largest homogeneous set, that is:

- 60 subjects
- One test function (i.e. Styblinski-Tang).
- One modality (i.e. finding a point whose evaluation is close to $f^* = f(x^*)$, but without knowing f^* a-priori.)

With respect to Algorithm 2 and Algorithm 3, we declare here the values of β in the UCB equation, that is $\beta = \{0.0, 0.5, 1.0\}$, and of the threshold $\tau = \{0.10, 0.15\}$.

4 Experimental results

4.1 Experiment 1: Gaussian process

The following figures summarize the main results related to the comparison between human players strategies and GP-based BO. We start with the case $\beta = 1.0$ in UCB.

From Fig. 3, EI is preferred, indicating a dominant exploitative behaviour among humans. Indeed, $\beta = 1.0$ in the UCB gives some chance to exploration, depending on the value of $\sigma(x)$. Statistical significance of this result was evaluated via a Fisher’s test for each pair consisting in EI versus another acquisition function. Given a pair of acquisition functions, the test hypothesis is that there is not a more compliant one. As summarized in Table 1, the hypothesis is rejected (p value < 0.05) for exponential and Matérn kernels, while it is accepted for power and squared exponential kernels.

Results reported in Fig. 4 are coherent with those previously depicted in Fig. 3. Moreover, the number of human players whose strategy is not compliant with BO is reduced, according to the less restrictive value of the threshold τ . We have applied a Fisher’s test to evaluate the statistical significance of the obtained outcomes; results are summarized in Table 2. According to the Fisher’s test, EI is

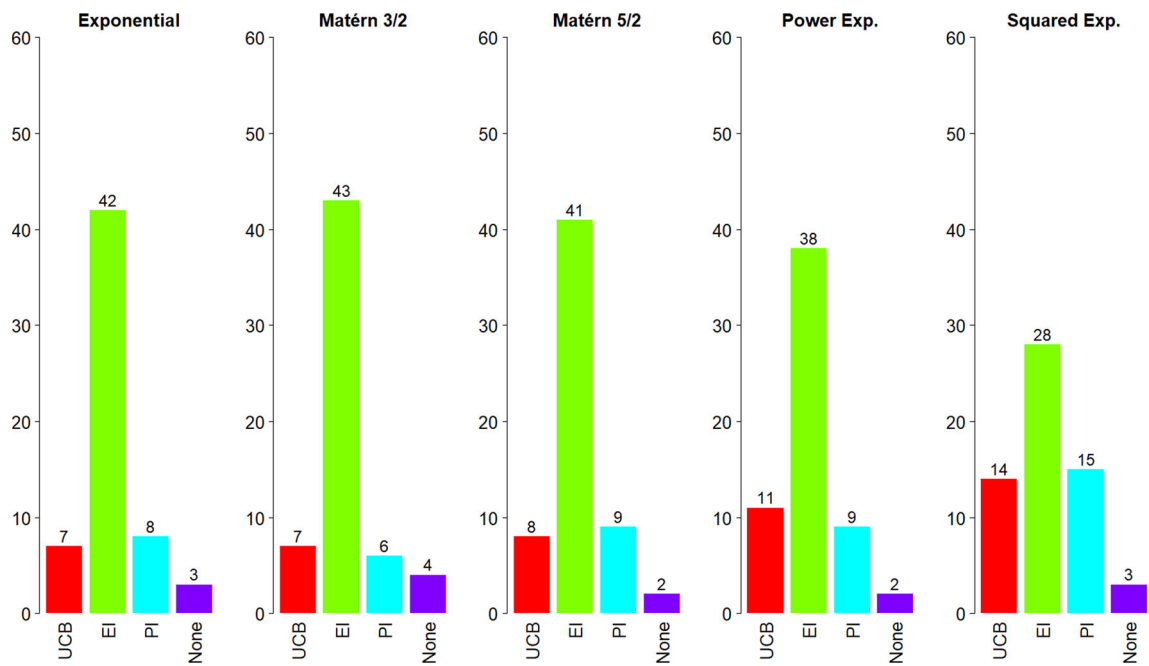


Fig. 4 Number of human players whose strategy is compliant with respect to kernel type and acquisition functions, with threshold τ set to 0.15, and $\beta = 1$ in UCB. Last bar represents the number of non-compliant

Table 2 Fisher’s test p values for each pair related to the most frequently compliant acquisition function in Fig. 4 (i.e. EI) and the others. The hypothesis that one acquisition function is more significantly compliant than the other is accepted with a confidence 0.05

| | EI versus UCB | EI versus PI | EI versus none |
|---------------------|---------------|--------------|----------------|
| Exponential | < 0.001 | < 0.001 | < 0.001 |
| Matérn 3/2 | < 0.001 | < 0.001 | < 0.001 |
| Matérn 5/2 | < 0.001 | < 0.001 | < 0.001 |
| Power exponential | 0.010 | < 0.001 | < 0.001 |
| Squared exponential | 0.184 | 0.191 | < 0.001 |

always the most compliant acquisition function, except for the squared exponential kernel.

According to the charts reported in Fig. 5, one can see that with the reduction in β , UCB gains a larger share of participants. Again, we have applied a Fisher’s test to evaluate the statistical significance of the obtained outcomes; the results are summarized in Table 3. According to the Fisher’s test, the EI is still the most compliant acquisition function, for most of the kernels, even if the p values increases with respect to the previous cases.

Figure 6 is again a confirmation that, for human players, “greed is good”: $\beta = 0$ means no-exploration in UCB whose fully exploitative version gets the largest share of participants. In this case we have performed a Fisher’s test for each pair related to UCB (instead of EI) versus the other acquisition functions, since UCB is now the most frequently compliant option. Results of the Fisher’s test are summarized in Table 4. Although UCB with $\beta = 0$ is the

most frequently compliant acquisition function, the number of shares does not result significantly different from that of EI. In any case this result confirms the exploitative nature of humans, who largely adopt strategies close to EI and “pure exploitative” UCB acquisition functions.

Although maximizing EI or UCB with $\beta = 0$ should analytically lead to the same next location x to evaluate, the numerical results can be quite different, due to the different shapes of these two acquisition functions. Indeed, when $\mu(x)$ is lower than the best value observed so far, y^+ , over a large portion of the search space, then EI is almost completely flat, as depicted in the example of Fig. 7. Consequently, whichever optimization strategy is used to optimize EI could easily fail, leading to no improvement of y^+ over new function evaluations. On the contrary, UCB with $\beta = 0$ coincides with $\mu(x)$ (Fig. 7): this allows to converge towards a local optimum of $\mu(x)$ and, in case, to improve over y^+ .

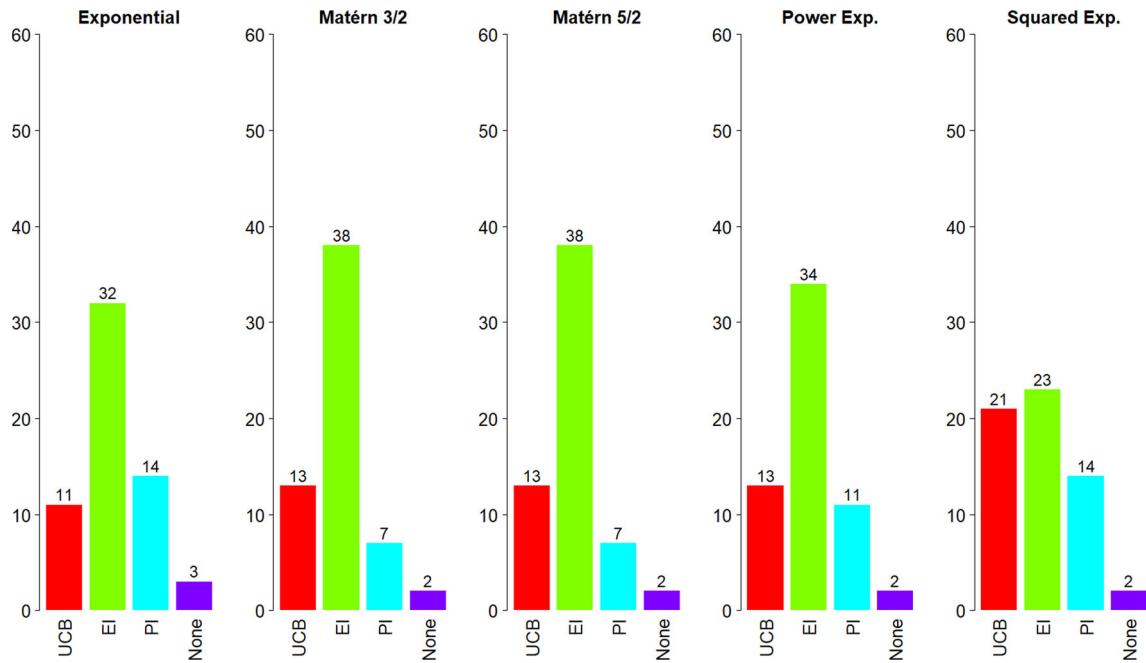


Fig. 5 Number of human players whose strategy is compliant with respect to kernel type and acquisition functions, with threshold τ set to 0.15, and $\beta = 0.5$ in UCB. Last bar represents the number of non-compliant

Table 3 Fisher’s test p values for each pair related to the most frequently compliant acquisition function in Fig. 5 (i.e. EI) and the others

| | EI versus UCB | EI versus PI | EI versus None |
|---------------------|---------------|--------------|----------------|
| Exponential | 0.044 | 0.088 | < 0.001 |
| Matérn 3/2 | 0.024 | < 0.001 | < 0.001 |
| Matérn 5/2 | 0.024 | < 0.001 | < 0.001 |
| Power exponential | 0.055 | 0.016 | < 0.001 |
| Squared exponential | 1.000 | 0.350 | 0.001 |

The hypothesis that one acquisition function is more significantly compliant than the other is accepted with a confidence 0.05

4.2 Experiment 2: random forest

The same analysis has been also performed by using RF as surrogate model instead of GP. Figure 8 shows that PI, a notoriously exploitative acquisition function, gets the larger share of compliant when UCB accounts for exploration ($\beta = 1$).

Although the exploitative nature of the human choice is confirmed also in this case, we observed a shift from EI to PI (that is even more exploitative than EI). This is basically due to the combination between the type of acquisition function and the shape of the surrogate model, that is piecewise constant for RF and continuous for GP.

The situation changes by reducing the exploration component in UCB, which with $\beta = 0.5$ and $\beta = 0.0$ dominates choices among the compliant. Summarizing, also in the case of RF, the human’s nature appears to be exploitative. We have performed a Fisher’s test to evaluate,

for each combination of τ and β , if the BO algorithm associated to the larger share of compliant participants can be statistically considered the most compliant one. Results of the Fisher’s test are reported in Table 5.

4.3 Comparison against other global optimization techniques

In this section we consider, among all the possibilities, the BO approach that proved to be more compliant with the humans’ strategies, overall. Let BO** denote this specific approach, that is: GP surrogate with Matérn 3/2 kernel, EI as acquisition function and $\tau = 0.15$. Indeed, as reported in Fig. 4, this BO configuration is associated to the largest share (i.e. 43) of participants compliant to an optimization strategy.

We computed all the distances between the optimal solutions (i.e. x^+) found by humans with those provided by

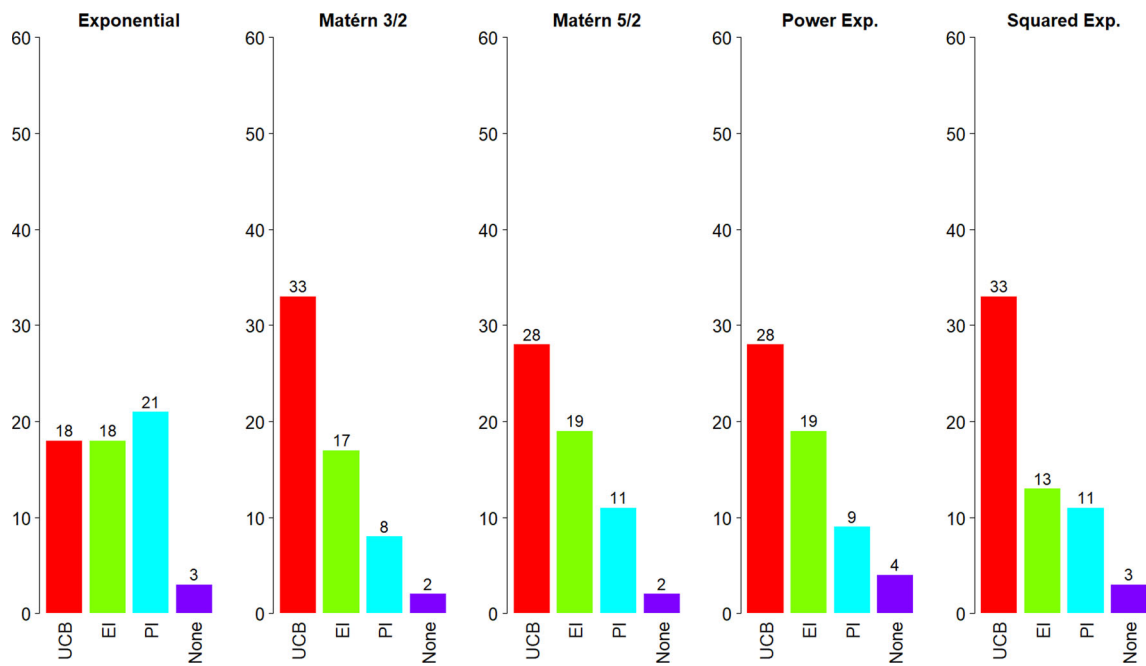


Fig. 6 Number of human players whose strategy is compliant with respect to kernel type and acquisition functions, with threshold τ set to 0.15, and $\beta = 0$ in UCB. Last bar represents the number of non-compliant

Table 4 Fisher’s test p values for each pair related to the most frequently compliant acquisition function in Fig. 6 (i.e. UCB) and the others

| | UCB versus EI | UCB versus PI | UCB versus none |
|---------------------|---------------|---------------|-----------------|
| Exponential | 1.000 | 1.000 | 0.020 |
| Matérn 3/2 | 0.156 | 0.005 | < 0.001 |
| Matérn 5/2 | 0.408 | 0.063 | < 0.001 |
| Power exponential | 0.408 | 0.030 | < 0.001 |
| Squared exponential | 0.054 | 0.027 | < 0.001 |

The hypothesis that one acquisition function is more significantly compliant than the other is accepted with a confidence 0.05

BO** and the other 5 global optimization techniques cited in Sect. 3.3 (i.e. RS, DIRECT, GA, PSO and SA).

According to the boxplot reported in Fig. 9, optimal solutions provided by BO** are, on average, closer to the humans’ ones compared to the other techniques.

More interestingly, the density plot in Fig. 10 shows that the distributions of the distances are basically bimodal, for all the global optimization techniques but BO**. Indeed, for BO** the density associated with small values of the distance is around twice that related to other global optimization techniques, proving that BO** is closer to the strategies applied by the human players.

Beyond the visual representation in Fig. 10, a further confirmation that BO is closer to human search comes from statistical tests. Since distances do not show a normal distribution, we could not apply ANOVA and, therefore, we used nonparametric tests. According to the Friedman’s test, distance values result significantly different among all

the techniques (p value < 0.001). To better investigate these differences, we have then performed a pairwise Wilcoxon test obtaining that:

- distance values related to RS, GA, PSO and SA are, pairwise, significantly similar (p value > 0.192).
- distance values related to DIRECT are, pairwise, significantly different from those related to all the other techniques (p value < 0.05) except for RS (p value = 0.106). This occurs because the peaks in the distributions of both DIRECT and RS correspond to distance values slightly greater than those of the other techniques.
- distance values related to BO** are, pairwise, significantly different from DIRECT and RS (p value < 0.001). Although BO**, in Fig. 10, is clearly different from every other optimization techniques, Wilcoxon test suggests that the hypothesis that they are similar must be accepted for GA, SA and PSO,

Fig. 7 An example of the reason why maximizing EI and UCB with $\beta = 0$ might numerically lead to different solutions. Since EI is almost completely flat over the search space, whichever optimization algorithm could easily fail in maximizing it, contrary to UCB with $\beta = 0$

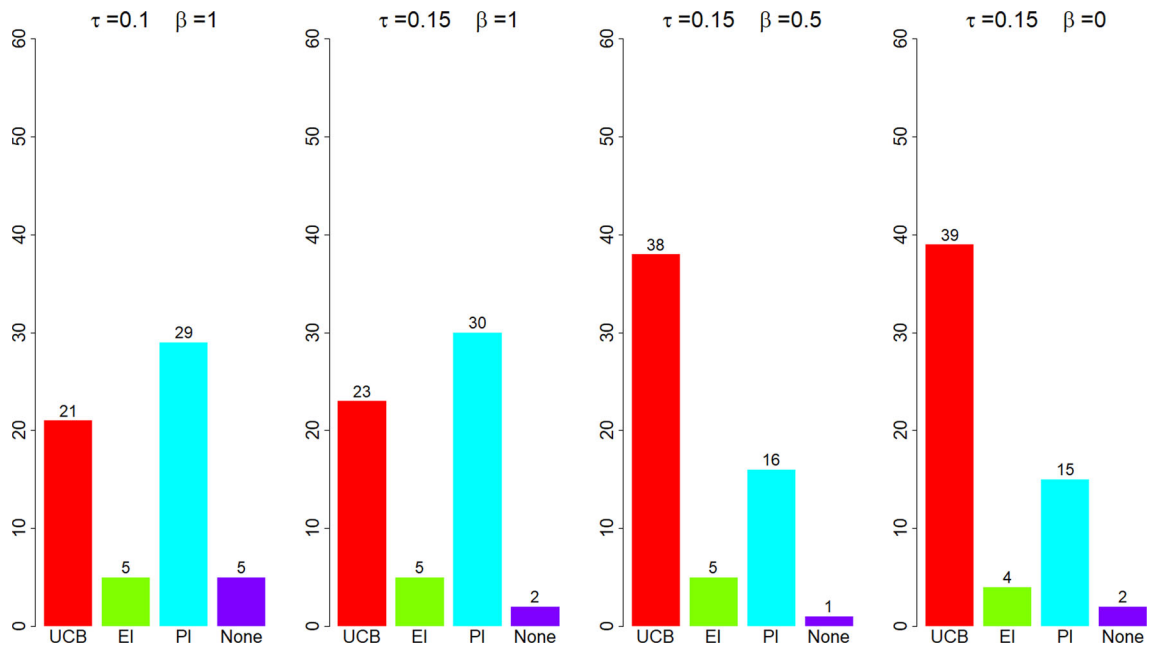
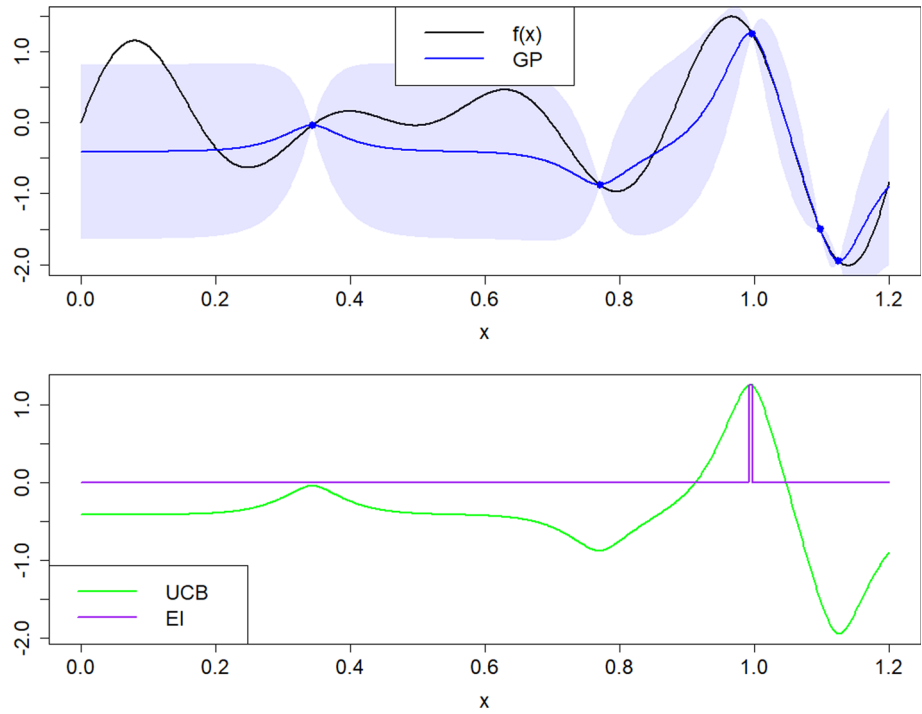


Fig. 8 Number of human players whose strategy is compliant with respect to different acquisition functions, with RF as surrogate model, in different settings of threshold τ and β values of UCB

even if with a very poor statistical significance level: p value = 0.062 for BO** versus GA and SA, and p value = 0.073 for BO** versus PSO. This is mainly because BO**, GA, SA and PSO have their highest peak around the same value of distance, even if the height of the BO**'s peak is higher than the others.

5 Conclusions

It is important to remark that the aim of the analysis performed in this study is not to compare the effectiveness of different techniques in solving the global optimization problem but to compare human search strategies with those

Table 5 Fisher’s test p values for each pair related to the most frequently compliant acquisition function in Fig. 7, for the four cases related to the values of the threshold τ and UCB’s β

| τ | β | Most compliant acquisition function | p values |
|--------|---------|-------------------------------------|---|
| 0.1 | 1 | PI | 0.547 (PI vs UCB) 0.004 (PI vs EI) 0.004 (PI vs None) |
| 0.15 | 1 | PI | 0.697 (PI vs UCB) 0.004 (PI vs EI) < 0.001 (PI vs None) |
| 0.15 | 0.5 | UCB | < 0.001 (UCB vs EI) 0.005 (UCB vs PI) 0.006 (UCB vs None) |
| 0.15 | 0 | UCB | < 0.001 (UCB vs EI) 0.029 (UCB vs PI) 0.025 (UCB vs None) |

The hypothesis that one acquisition function is more significantly compliant than the other is accepted with a confidence 0.05

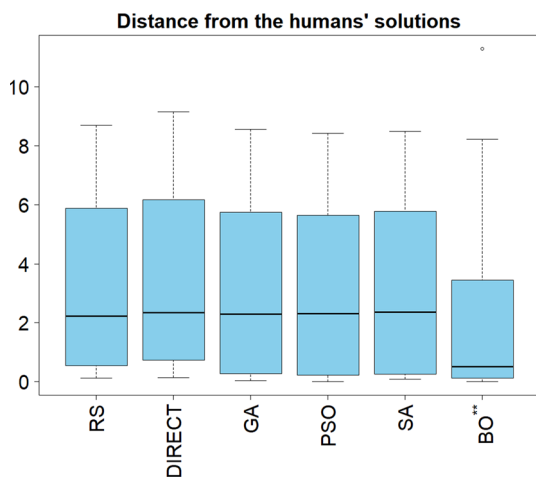


Fig. 9 Distances between the best solutions found by the humans and those provided by each global optimization strategy. BO** denotes the overall most compliant BO algorithm (i.e. GP surrogate with Matérn 3/2 kernel, EI acquisition function, from Fig. 4)

embodied in global optimization algorithms. Results proved that BO is closer to human search strategy and therefore provides a valuable insight about how humans deal with the trade-off between immediate rewards (exploitation) and risk propension (exploration). This can be explained by two major factors: first, the most coherent model for learning is provided by Bayesian learning, as recognized in the cognitive sciences community. Therefore, BO as a model-based active learning is revealed by the humans’ reward driven sampling behaviour and embodies underlying decision-making strategies. Second,

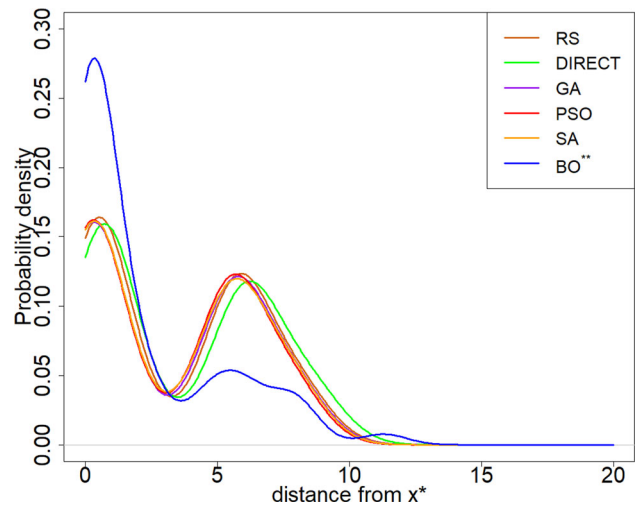


Fig. 10 Density plot of the distances between the best solutions found by the humans and those provided by each global optimization strategy

the flexibility of BO, as recognized also in the machine learning community, is larger than in other optimization techniques and is given by the freedom to choose the probabilistic surrogate model (GP or RF)—and the kernel in the case of GP—and the acquisition function.

Since the flexibility in BO is structural, while in the other optimization techniques considered it is just parametric, BO can naturally model better the internal representation of the function that humans maintain when they search for the optimum. In this paper this conclusion is supported by a wide set of statistical tests: the number of BO compliant participants is very high (less than 10% of participants resulted non-compliant to all models and acquisition functions). This result is also visually captured by the results depicted in Figs. 9 and 10. Moreover, non-parametric statistical testing provides meaningful outcomes. The authors admit that their statistical significance should be improved: indeed, the analysis is being extended, by the use of the Amazon Mechanical Turk, to a larger sample of participants, each one performing all the stimuli and game modalities developed in the game playing software.

An interesting result comes from the analysis of which space model, that is GP’s kernel, is implied by human search: based on the limited set of results presented in this paper, kernel is not a major factor in determining compliance to GP-based BO. On the other hand, the main drivers for compliance is the acquisition function, and its parametrization for UCB: the value of β significantly affects the relative importance of UCB over the search processes performed by the compliant participants. Then, we can conclude that exploitation-oriented acquisition functions get a consistently larger share, meaning that

“greed is good”, that is the human behaviour when looking for rewards in black-box situations is dominantly exploitative.

Acknowledgements We gratefully acknowledge the hard work and extremely helpful suggestions of the unknown reviewers.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Archetti F, Candelieri A (2019) Bayesian optimization and data science. Springer, New York
- Banks A, Vincent J, Anyakoha C (2007) A review of particle swarm optimization. Part I: background and development. *Nat Comput* 6(4):467–484
- Bischl B, Richter J, Bossek J, Horn D, Thomas J, Lang M (2017) mlrMBO: a modular framework for model-based optimization of expensive black-box functions. *arXiv preprint arXiv:1703.03373*
- Borji A, Itti L (2013) Bayesian optimization explains human active search. In: *Advances in neural information processing systems*, pp 55–63
- Candelieri A, Perego R, Archetti F (2018) Bayesian optimization of pump operations in water distribution systems. *J Glob Optim* 71(1):213–235
- Candelieri A, Perego R, Giordani I, Archetti F (2019) Are humans Bayesian in the optimization of black-box functions? In: *International conference on numerical computations: theory and algorithms*. Springer, Cham, pp 32–42
- Eggenesperger K, Lindauer M, Hutter F (2019) Pitfalls and best practices in algorithm configuration. *J Artif Intell Res* 64:861–893
- Gershman SJ (2019) Uncertainty and exploration. *Decision* 6(3):277
- Gopnik A, O’Grady S, Lucas CG, Griffiths TL, Wente A, Bridgers S, Dahl RE (2017) Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proc Natl Acad Sci* 114(30):7892–7899
- Griffiths TL, Kemp C, Tenenbaum JB (2008) Bayesian models of cognition. In: Sun R (ed) *Cambridge handbook of computational cognitive modelling*. Cambridge University Press, Cambridge
- Hartford JS, Wright JR, Leyton-Brown K (2016) Deep learning for predicting human strategic behaviour. In: *Advances in neural information processing systems*, pp 2424–2432
- Ho TK (1995) Random decision forests. In: *Proceedings of 3rd international conference on document analysis and recognition*. IEEE, vol 1, pp 278–282
- Jones DR, Perttunen CD, Stuckman BE (1993) Lipschitzian optimization without the Lipschitz constant. *J Optim Theory Appl* 79(1):157–181
- Kruschke JK (2008) Bayesian approaches to associative learning: From passive to active learning. *Learn Behav* 36(3):210–226
- Kushner HJ (1964) A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise
- Van Laarhoven PJ, Aarts EH (1987) Simulated annealing. In: van Laarhoven PJ, Aarts EH (eds) *Simulated annealing: theory and applications*. Springer, Dordrecht, pp 7–15
- Lera D, Posypkin M, Sergeyev YD (2021) Space-filling curves for numerical approximation and visualization of solutions to systems of nonlinear inequalities with applications in robotics. *Appl Math Comput* 94:390
- Mehlhorn K, Newell BR, Todd PM, Lee MD, Morgan K, Braithwaite VA, Gonzalez C (2015) Unpacking the exploration–exploitation tradeoff: a synthesis of human and animal literatures. *Decision* 2(3):191
- Moćkus J (1975) On Bayesian methods for seeking the extremum. In: *Optimization techniques IFIP technical conference*. Springer, Berlin, pp 400–404
- Plonsky O, Apel R, Ert E, Tennenholtz M, Bourgin D, Peterson JC, Cavanagh JF (2019) Predicting human decisions with behavioral theories and machine learning. *arXiv preprint arXiv:1904.06866*
- Schulz E, Konstantinidis E, Speekenbrink M (2018a) Putting bandits into context: How function learning supports decision making. *J Exp Psychol Learn Mem Cogn* 44(6):927
- Schulz E, Speekenbrink M, Krause A (2018b) A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions. *J Math Psychol* 85:1–16
- Schulz E, Tenenbaum JB, Reshef DN, Speekenbrink M, Gershman S (2015) Assessing the perceived predictability of functions. In: *CogSci*
- Schulz E, Speekenbrink M, Hernández-Lobato JM, Ghahramani Z, Gershman SJ (2016a) Quantifying mismatch in Bayesian optimization. In: *Nips workshop on Bayesian optimization: black-box optimization and beyond*
- Schulz E, Tenenbaum J, Duvenaud DK, Speekenbrink M, Gershman SJ (2016b) Probing the compositionality of intuitive functions. In: *Advances in neural information processing systems*, pp 3729–3737
- Scrucca L (2013) GA: a package for genetic algorithms in R. *J Stat Softw Found Open Access Stat* 53(4):1–37
- Scrucca L (2016) On some extensions to GA package: hybrid optimisation, parallelisation and islands evolution. *arXiv preprint arXiv:1605.01931*
- Sergeyev YD, Nasso MC, Mukhametzhonov MS, Kvasov DE (2020a) Novel local tuning techniques for speeding up one-dimensional algorithms in expensive global optimization using Lipschitz derivatives. *J Comput Appl Math* 383:113134
- Sergeyev YD, Candelieri A, Kvasov DE, Perego R (2020b) Safe global optimization of expensive noisy black-box functions in the δ -Lipschitz framework. *Soft Comput* 2020:1–21

- Srinivas N, Krause A, Kakade SM, Seeger M (2009) Gaussian process optimization in the bandit setting: No regret and experimental design. arXiv preprint [arXiv:0912.3995](https://arxiv.org/abs/0912.3995)
- Williams CK, Rasmussen CE (2006) Gaussian processes for machine learning, vol 2, no 3. MIT Press, Cambridge, p 4
- Wilson AG, Dann C, Lucas C, Xing EP (2015) The human kernel. In: Advances in neural information processing systems, pp 2854–2862
- Wu CM, Schulz E, Speekenbrink M, Nelson JD, Meder B (2018) Generalization guides human exploration in vast decision spaces. *Nat Hum Behav* 2(12):915–924
- Zhigljavsky A, Zilinskas A (2007) Stochastic global optimization, vol 9. Springer, New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.