Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Novelty and Interestingness Measures for Design-space Exploration

Edgar Reehuis, Markus Olhofer, Bernhard Sendhoff, Thomas Bäck

2013

Preprint:

This is an accepted article published in 15th Annual Conference on Genetic and Evolutionary Computation, GECCO'13. The final authenticated version is available online at: https://doi.org/[DOI not available]

Novelty and Interestingness Measures for Design-space Exploration

Edgar Reehuis
Natural Computing Group
LIACS, Leiden University
Niels Bohrweg 1, 2333 CA
Leiden, The Netherlands
ereehuis@liacs.nl

Markus Olhofer
Honda Research Institute Europe
Carl-Legien-Straße 30
63073 Offenbach/M, Germany
markus.olhofer@honda-ri.de

Michael Emmerich
Natural Computing Group
LIACS, Leiden University
Niels Bohrweg 1, 2333 CA
Leiden, The Netherlands
emmerich@liacs.nl

Bernhard Sendhoff
Honda Research Institute Europe
Carl-Legien-Straße 30
63073 Offenbach/M, Germany
bernhard.sendhoff@honda-ri.de

Thomas Bäck
Natural Computing Group
LIACS, Leiden University
Niels Bohrweg 1, 2333 CA
Leiden, The Netherlands
baeck@liacs.nl

ABSTRACT

Measures of novelty and interestingness are frequently encountered in the context of developmental robotics, being derived from human psychology. This work addresses these measures from the viewpoint of enhancing design-space exploration in black-box optimization. We provide a unifying notational and naming scheme with the intent of facilitating comparison, implementation, and application in the domain of design optimization. Initial analysis shows a promising interestingness measure for being tried on real-world design problems.

Categories and Subject Descriptors

I.2.8 [**Computing Methodologies**]: ARTIFICIAL INTELLIGENCE—*Problem Solving, Control Methods, and Search*

Keywords

Novelty, interestingness, surrogate model, exploration measure

1. INTRODUCTION

The idea of using *novelty* and *interestingness* of possible actions or solutions as an indication of what to try or where to explore next is widespread in *neuro-control* and *reinforcement learning* [19, 20, 3, 12, 6, 14, 13, 1, 7, 10]. In lesser extent, it is being proposed for design-space exploration in *black-box optimization* [5, 18, 2, 17].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
GECCO'13, July 6–10, 2013, Amsterdam, The Netherlands.
Copyright 2013 ACM 978-1-4503-1963-8/13/07 ...$15.00.

Novelty of an observation, action, or solution, is a subjective notion in the sense that it depends on *similarity* to what was encountered before, and on the *state* of the agent, approach, or model used in *determining* this similarity. Easily modeled and frequently occurring solutions quickly lose their novelty, while solutions that are subject to *noise* always remain novel [6]. For efficient learning or modeling of a solution space, those solutions are selected at each time step that improve the understanding or model of the solution space most: An *interesting* solution has sufficient (not too simple, boring) but not too extreme (beyond understandable, for instance, again, noise) novelty [20]. Thus in learning, not the most novel but the most interesting solutions are to be selected (see Fig. 1).

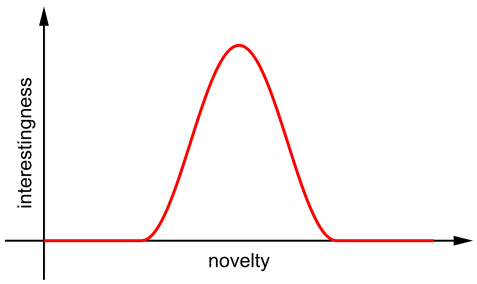


Figure 1: *Wundt's Curve.* Interestingness shown as a function of novelty. Image courtesy of Graziano et al. [6].

This paper provides an overview of novelty and interestingness expressions from the perspective of design-space exploration by unifying the expressions in notation, and proposing adjustments with this application in mind where necessary. A division is made between similarity-derived, distance-based novelty, and learning-based novelty related to fitting a surrogate model (in [6] referred to as *coherence* and *compression*, respectively). Note that the surrogate model is *not* intended here for reducing the number of used quality

function evaluations, but for expressing the understanding, locally, of the quality landscape during optimization.

We aim for an exploration measure to focus on irregular areas in the search space, but eventually steer clear from regions with excessive noise. The first aspect is tested for in a preliminary experimental setup involving a low-dimensional artificial test function. A comparison is made in selecting on the exploration measures only, quality is not involved, while the reviewed measures are intended, in a later stage, to be employed as additional selection criterion in quality-based search, either in parallel to quality (e.g., [5, 18, 17]) or in an alternating fashion (e.g., [2]). Learning-based novelty should inherently focus on areas with higher variance in quality as these are more difficult to model, while distance-based novelty will only do this if supplied with a distance formulation that is derived from quality.

Following, Section 2 continues with an overview of novelty measures, and Section 3 with interestingness measures, after which the experiments are described in Section 4, of which the outcome and outlook are discussed in Section 5.

2. NOVELTY

To determine a *candidate solution*'s distinctiveness or *novelty* from the knowledge gathered on the design space during design-space exploration, an appropriate representation of the available knowledge is required such that new solutions can be compared to it. Here the notion of a candidate solution is based on an iterative search method that generates multiple candidates in each iteration, of which it selects a certain number to proceed into the next iteration. Possible representations of the available knowledge include an archive containing all sampled solutions, and a surrogate model that learns a mapping between different characteristics of a solution, for instance from the solution vector in the search space to the quality value that becomes available after evaluation.

2.1 Distance-based Novelty

In distance-based novelty a candidate solution is compared to an archive of earlier sampled solutions, and its difference to the most resembling solutions from this archive serves as novelty expression, determined with respect to a *domain-specific distance measure*. For the search spaces considered the solutions are real-valued vectors, therefore a generally-applicable distance formulation is the Euclidean distance between the solution vectors.

Sparseness

Lehman and Stanley [10] propose *sparseness* (Sp) as distance-based novelty measure: A candidate solution $\mathbf{x}$ is compared to the solutions $\mathbf{q}$ from the union of a *partial* archive S of earlier sampled solutions (that is, solutions $\mathbf{x}$ for which the quality value $f(\mathbf{x})$ was determined) and the set X of current candidate solutions excluding $\mathbf{x}$, to determine its closest k neighbors based on a domain-specific distance measure $\mathrm{d}(\mathbf{x}, \mathbf{q})$,

$$\underset{\{\mathbf{q}_1,\ldots,\mathbf{q}_k\} \subseteq Q}{\arg\min} \sum_{i=1}^{k} \mathrm{d}(\mathbf{x}, \mathbf{q}_i), \tag{1}$$

$$Q = \{\mathbf{q} \mid (\mathbf{q}, f(\mathbf{q})) \in S\} \cup X \setminus \{\mathbf{x}\}. \tag{2}$$

The sparseness of $\mathbf{x}$ is then taken as the average distance to these closest k solutions,

$$\mathrm{Sp}(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^{k} \mathrm{d}(\mathbf{x}, \mathbf{q}_i). \tag{3}$$

Settings for the following parameters are required [10]:

- Number of closest solutions k: a value of 15 is used by Lehman and Stanley for varying problems;
- Dynamic archive-inclusion threshold $\mathrm{Sp}_{\min}$: only solutions with sparseness that exceeds this threshold are added to the archive. After some initial setting (we use 1.0), $\mathrm{Sp}_{\min}$ is lowered by 5% if $\frac{1}{100}$-part of the sampling budget was used and no solutions were added; it is increased by 20% if $\frac{1}{100}$-part of the sampling budget was used and more than 4 solutions were added;
- Each sampled solution has an unconditional 0.1% chance to be added to the archive.

Uniqueness

Seeking to reduce the complexity and the number of parameters involved, we reformulate sparseness as the *uniqueness* (Un) of a solution, that is, the minimum distance occurring to a solution $\mathbf{q}$ from an archive S containing *all* sampled solutions (also used by Hester and Stone [7]),

$$\mathrm{Un}(\mathbf{x}) = \min_{\mathbf{q}, (\mathbf{q}, f(\mathbf{q})) \in S} \mathrm{d}(\mathbf{x}, \mathbf{q}). \tag{4}$$

In expressing uniqueness there is no difference in novelty score in case the *closest solution* $\mathbf{q}$ has no similar counterparts in the archive, compared to when the closest solution is part of a group of similar archived solutions. However, as all sampled solutions are added to the archive, this effect is of less consequence than it would be in a sparseness scheme with $k = 1$ as the chance for a solution of being included in the archive there depends heavily on the moment during the search that it is generated. Furthermore, the unconditional archive leads to a higher computational load than the partial archive in sparseness, but the archive size is of course bounded by the number of allowed to be sampled solutions.

2.2 Learning-based Novelty

In learning-based novelty the error that a surrogate model makes or is likely to make in its prediction for a candidate solution serves as novelty expression [5, 14, 3]. This way the distinctiveness from available knowledge is represented as the *generalization performance* that the surrogate model trained on the available knowledge has in approximating the to-be-learned mapping for an unseen solution.

Prediction Error

For a surrogate model that approximates a mapping from solution vectors $\mathbf{x}$ to quality-function values $f(\mathbf{x})$, the *prediction error* (PE) is calculated as the absolute difference between the model's predicted quality value $\hat{f}(\mathbf{x})$ and the actual outcome $f(\mathbf{x})$ of evaluating $\mathbf{x}$ by the quality function,

$$\mathrm{PE}(\mathbf{x}) = \left| \hat{f}(\mathbf{x}) - f(\mathbf{x}) \right|. \tag{5}$$

Dispersion in Predictions

Since PE depends on the local variation in the outcome of the quality function (i.e., its ruggedness) and on the spatial

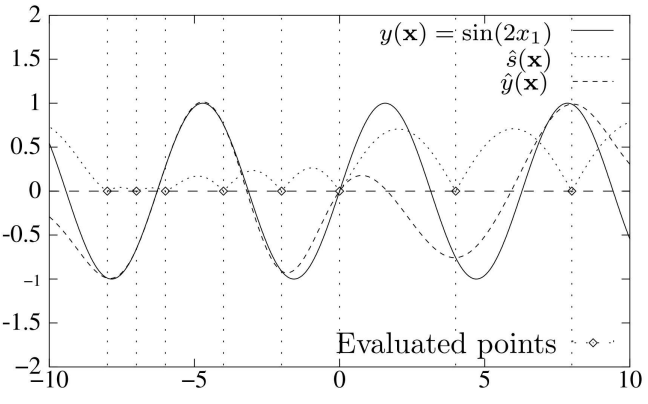


Figure 2: *Predictive Variance.* Quality function $y(\mathbf{x})$, predictive mean $\hat{y}(\mathbf{x})$, and predictive variance $\mathrm{PV}(\mathbf{x}) = \hat{s}^2(\mathbf{x})$ for the one-dimensional function $y(\mathbf{x}) = \sin(2x_1)$. Predictive variance depends on the spatial distribution of the sample points. Image courtesy of Emmerich [4].

distribution of the sampled solutions used in training the surrogate model, PE can be estimated by training an ensemble of multiple surrogate models [11] and measuring the *dispersion in predictions* (DP) for a candidate solution [17, 7]. DP is assumed to provide a similar ordering between candidate solutions as ranking them on their actual PE would, without the requirement of evaluating candidate solutions on quality beforehand; formally, for any two solutions $\mathbf{x}_1, \mathbf{x}_2$ it is assumed that

$$\mathrm{PE}(\mathbf{x}_1) < \mathrm{PE}(\mathbf{x}_2) \Rightarrow \mathrm{DP}(\mathbf{x}_1) < \mathrm{DP}(\mathbf{x}_2). \quad (6)$$

For calculating DP we make use of the *interquartile range* (IQR, the difference between the third and first quartile of the values, which is a robust statistic in the sense that it is not prone to outliers) between the quality value predictions of the different models,

$$\mathrm{DP}(\mathbf{x}) = \mathrm{IQR}(\hat{f}_1(\mathbf{x}), \ldots, \hat{f}_{\dim(\hat{\mathbf{f}})}(\mathbf{x})), \quad (7)$$

where $\dim(\hat{\mathbf{f}})$ equals the number of surrogate models used in parallel. DP requires surrogate models with potentially differing outcomes, either a combination of different modeling techniques or multiple instances of a technique that may vary in prediction outcome when differently initialized in presence of limited training data.

Predictive Variance

In using *Gaussian Process modeling* (GP) and *Kriging* methods for surrogate modeling, next to the predicted value $\hat{f}(\mathbf{x})$ (*predictive mean*) an estimate of the associated prediction error formulated as *predictive variance* (PV) is readily available for a point $\mathbf{x}$ [15, 9],

$$\mathrm{PV}(\mathbf{x}) = \sigma^2(\mathbf{x}) \text{ [GP]} \hat{=} s^2(\mathbf{x}) = \mathrm{MSE}(\mathbf{x}) \text{ [Kriging]}, \quad (8)$$

assuming the actual quality value $f(\mathbf{x})$ to be a realization of $\mathcal{N}\left(\hat{f}(\mathbf{x}), \sqrt{\mathrm{PV}(\mathbf{x})}\right)$.

Schaul et al. [18] present PV as a criterion for design-space exploration[1], indirectly, through showing a direct relationship from PV to what they term *expected information gain*,

$$\Psi(\mathbf{x}) = \frac{1}{2}\log \mathrm{PV}(\mathbf{x}) - \frac{1}{2}\log \sigma^2_{\mathrm{noise}}, \quad (9)$$

where the noise variance $\sigma^2_{\mathrm{noise}}$ [15] is constant, and thus holds that $\mathrm{PV}(\mathbf{x}_1) < \mathrm{PV}(\mathbf{x}_2) \Leftrightarrow \Psi(\mathbf{x}_1) < \Psi(\mathbf{x}_2)$ for any two solutions $\mathbf{x}_1, \mathbf{x}_2$.

Although PV depends on the ruggedness of the quality landscape, *relatively* PV depends only on the distribution of the sample points used in training because of the assumption that this ruggedness is uniform in each region of the input space (*stationarity assumption*), see Fig. 2. PV could be adjusted to depend on local ruggedness, like PE and DP, by using a *non-stationary* variant of GP or Kriging [22], or by re-training the surrogate for each solution $\mathbf{x}$ for which PV is to be assessed using its closest neighbors only from the available training data.

3. INTERESTINGNESS

The learning-based novelty of a solution is expressed as the prediction error in that point made by a surrogate model approximating the mapping of solution vectors to, e.g., the quality-function landscape. In selecting on learning-based novelty, solutions with maximum error are promoted. However, should notoriously hard regions exist in the design space that are effectively impossible to model, e.g., because of non-stationary noise or high non-stationary uncertainty, these are ideally detected and avoided.

Instead of directly selecting on the prediction error that the latest state of the surrogate model gave rise to, we can compare the modeling error at different moments in time during the learning process. The *interestingness* of a solution is formulated as the *learning progress* in the region that the solution lies in, serving as a way of omitting regions that cannot be modeled, that is, in which the surrogate model's performance does not improve over time. In selecting on interestingness those solutions are picked that improve or are expected to improve the surrogate model's approximation of the to-be-learned mapping most by including them in model training, after sampling them (i.e., evaluating them on quality).

3.1 Interestingness Measures

The interestingness measures are all based on *determining* or *estimating* the learning progress that a candidate solution will induce *after* including it in model training. In defining the interestingness measures, the following notation will be used:

- t is the current position in time; time is viewed as passing *discretely*, each iteration of the iterative search method defining a single time step and initializing at $t = 0$;
- $\mathbf{x}_{i,t}$ is a candidate solution from the set X_t of all candidate solution vectors at time t, with $1 \leq i \leq |X_t|$; X_t contains the solutions that are candidate *at* time t, *not* all previous candidates up to time t;

[1]We view PV as an expression of learning-based novelty, and not like Schaul et al. of interestingness (see Section 3), as PV does not include a(n) (implicit) mechanism for dealing with regions of the design space that, effectively, cannot be modeled.

- $\mathcal{R}(\mathbf{x}_{i,t})$ is the region in the design space in which $\mathbf{x}_{i,t}$ lies (for a definition, see Section 3.2);
- $\mathbf{s}_{j,t} = (\mathbf{x}, f(\mathbf{x}))$ is a solution sampled at time t, that is, for which quality value $f(\mathbf{x})$ was determined;
- $S_t = \bigcup_j \{\mathbf{s}_{j,t}\}$ is the set of all solutions sampled at time t; not necessarily all candidate solutions at time t are sampled: In interestingness based on PE, S_t equals $\{(\mathbf{x}_{i,t}, f(\mathbf{x}_{i,t})) \mid 1 \leq i \leq |X_t|\}$, but in using DP and PV, S_t is based on a subset of X_t;
- $\mathcal{M}(T)$ is a surrogate model that approximates the mapping from solution vectors $\mathbf{x}$ to quality values $f(\mathbf{x})$, and that is trained on tuples of solutions with quality values in the supplied training set T;
- $T_t = \bigcup_{k=0}^{t-1} S_k$ is the training data available for $\mathcal{M}$ at time t, containing tuples of all earlier sampled solutions with accompanying quality values;
- V_t is the test set for determining the performance of $\mathcal{M}$ at time t; in case of PE, V_t needs to contain tuples of solution vectors with accompanying quality values $(\mathbf{v}, f(\mathbf{v}))$, in case of DP and PV the quality component is not required and can be replaced by 0 for solutions for which it is not available;
- $\mathrm{ME}(\mathbf{v}, \mathcal{M})$ is the modeling error that $\mathcal{M}$ gives rise to for a solution $(\mathbf{v}, f(\mathbf{v})) \in V_t$ at time t; it is either determined as $\mathrm{PE}(\mathbf{v})$, or approximated as $\mathrm{DP}(\mathbf{v})$ or $\mathrm{PV}(\mathbf{v})$;
- $\mathrm{Err}(\mathcal{R}, \mathcal{M}, V)$ is the *regional error value* for region $\mathcal{R}$ that is derived from the modeling errors $\mathrm{ME}(\mathbf{v}, \mathcal{M})$ for solutions $\mathbf{v} \in V_t$; an approach for determining regional error values is given in Section 3.2.

Before we can describe the interestingness measures, a definition of the observed learning progress in a region $\mathcal{R}$ at time u is required. It is calculated as the discrete-time derivative of the regional error $\mathrm{Err}(\mathcal{R}, \mathcal{M}, V)$, taken as the decrease of the error from time $u - \tau$ to time u,

$$\mathrm{LP}(\mathcal{R}, u) = \mathrm{Err}(\mathcal{R}, \mathcal{M}(T_{u-\tau}), V_{u-\tau}) - \mathrm{Err}(\mathcal{R}, \mathcal{M}(T_u), V_u). \tag{10}$$

To determine the performance of surrogate model $\mathcal{M}(T_u)$, at each time step u the set S_u of new solutions sampled *at* time u serves as test set V_u. Now, importantly, Eq. (10) provides the learning progress observed in region $\mathcal{R}$ at time u, depending on the modeling errors observed in the candidate solutions $\mathbf{x}_{i,u}$, since $V_u = S_u \subseteq \{(\mathbf{x}_{i,u}, f(\mathbf{x}_{i,u})) \mid 1 \leq i \leq |X_u|\}$. However, in order to determine the learning progress that a solution that is candidate at time u itself *induces*, we have to be *one time step in the future*! The model has to be re-trained, now including the candidate solution, and then tested again in the same region.

Therefore, if we want to know the learning progress that a candidate solution at time t will have given rise to at time $t + 1$, we perform the following steps:

1. We are at time $t + 1$ and in the region that candidate solution $\mathbf{x}_{i,t}$ defines, we determine the previous regional error based on the modeling errors that the model at time t (taking $\tau = 1$) showed for the set S_t of solutions sampled at time t,

$$\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t), V_t = S_t\right); \tag{11}$$

2. We then re-train the model on the union of the training set used at time t *and* the candidate solution $\mathbf{x}_{i,t}$, which needs to have been sampled, and again determine the error in region $\mathcal{R}(\mathbf{x}_{i,t})$, now defined as

$$\mathrm{Err}(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t \cup \{(\mathbf{x}_{i,t}, f(\mathbf{x}_{i,t}))\}), V_{t+1} = S_{t+1}), \tag{12}$$

thus, the regional error at time $t + 1$ is based on the modeling errors observed in the newly sampled solutions S_{t+1};

3. By subtracting the regional error at time t and $t + 1$ according to Eq. (10), we obtain the learning progress that $\mathbf{x}_{i,t}$ has induced.

Actual Learning Progress

Schmidhuber [19, 13, 14] follows the steps laid out above, but at time t already re-trains the model, for *each* candidate solution $\mathbf{x}_{i,t}$, thus generating $|X_t|$ temporary[2] models to be able to express each candidate's interestingness. The requirement of being one time step in the future rises from the fact that a set of newly sampled solutions S_{t+1} is to be used in evaluating the updated model. This is however circumvented by evaluating the updated model on the *same* test set again, namely the solutions sampled at time t. We term this measure *actual learning progress* (ALP),

$$\mathrm{ALP}(\mathcal{R}(\mathbf{x}_{i,t})) = \underbrace{\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t), V_t = S_t\right)}_{\text{error in region } \mathcal{R}(\mathbf{x}_{i,t}) \text{ before training on } \mathbf{x}_{i,t}} - \underbrace{\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t \cup \{(\mathbf{x}_{i,t}, f(\mathbf{x}_{i,t}))\}), V_{t+1} = S_t\right)}_{\text{error in region } \mathcal{R}(\mathbf{x}_{i,t}) \text{ after training on } \mathbf{x}_{i,t}}. \tag{13}$$

Previous Learning Progress

Oudeyer et al. [14, 13] propose for each region $\mathcal{R}$ to adopt the learning progress $\mathrm{LP}(\mathcal{R}, \theta_{\mathcal{R}})$ that was realized at the last time step $\theta_{\mathcal{R}}$ that a solution was sampled in that region. For the used regional error value definition (see Section 3.2) holds that at time t for all regions $\theta_{\mathcal{R}}$ equals t.

The *previous learning progress* (PLP) measure indicates for a region whether it has been possible to better approximate the to-be-learned mapping there between samplings, and should thereby detect regions in which the surrogate model does not improve. For a candidate $\mathbf{x}_{i,t}$, PLP is taken as the learning progress that *was* realized in the region $\mathcal{R}(\mathbf{x}_{i,t})$ that it defines, at time t (candidate $\mathbf{x}_{i,t}$ itself induces learning progress at time $t + 1$),

$$\mathrm{PLP}(\mathcal{R}(\mathbf{x}_{i,t})) = \mathrm{LP}(\mathcal{R}(\mathbf{x}_{i,t}), t) = \underbrace{\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_{t-\tau}), V_{t-\tau} = S_{t-\tau}\right)}_{\text{error in region } \mathcal{R}(\mathbf{x}_{i,t}) \text{ at time } t-\tau} - \underbrace{\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t), V_t = S_t\right)}_{\text{error in region } \mathcal{R}(\mathbf{x}_{i,t}) \text{ at time } t}. \tag{14}$$

Note that in combination with the used region definition, the discriminative power between candidate solutions $\mathbf{x}_{i,t}$ is limited to the difference in regional error values that arises from the *distance* between the defined regions $\mathcal{R}(\mathbf{x}_{i,t})$.

[2]It could be argued that only $|X_t| - 1$ models are temporary as the model of the best-ranked candidate can be preserved, but this does not hold in case multiple candidates are selected to proceed to the next iteration.

Previous Competence Change

Next to promoting high learning progress, reflecting that the model is capable of improving its mapping in a certain region, Baranes and Oudeyer [1] suggest valuing *negative* learning progress as this could point at deterioration in a region in which improvement was shown earlier. That region should be focused on again for verifying the fact of (not) being able to improve there. *Previous competence change* (PCC) is like PLP based on the most-recent learning progress shown in a region, but taking the absolute value,

$$\begin{aligned}\mathrm{PCC}(\mathcal{R}(\mathbf{x}_{i,t})) \;=\; |\mathrm{PLP}(\mathcal{R}(\mathbf{x}_{i,t}))| \;=\\ |\underbrace{\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_{t-\tau}), V_{t-\tau} = S_{t-\tau}\right)}_{\text{error in region } \mathcal{R}(\mathbf{x}_{i,t}) \text{ at time } \underline{t-\tau}} -\\ \underbrace{\mathrm{Err}\left(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t), V_t = S_t\right)}_{\text{error in region } \mathcal{R}(\mathbf{x}_{i,t}) \text{ at time } \underline{t}}|. \end{aligned} \quad (15)$$

Reducible Error

Yet another approach is valuing *only* negative PLP. Luciw et al. [12] aim for error maximization like in selecting on learning-based novelty directly, but account for regions that cannot be modeled by subtracting the *irreducible* error, that is, the error in the region at an earlier time step, from the novelty score, resulting in the *reducible error* (RE),

$$\begin{aligned}\mathrm{RE}(\mathcal{R}(\mathbf{x}_{i,t})) \;=\; -\,\mathrm{PLP}(\mathcal{R}(\mathbf{x}_{i,t})) \;=\\ \underbrace{\mathrm{Err}(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_t), V_t = S_t)}_{\underline{\text{current}} \text{ error in } \mathcal{R}(\mathbf{x}_{i,t}), \text{ at time } \underline{t}} -\\ \underbrace{\mathrm{Err}(\mathcal{R}(\mathbf{x}_{i,t}), \mathcal{M}(T_{t-\tau}), V_{t-\tau} = S_{t-\tau})}_{\underline{\text{irreducible}} \text{ error in } \mathcal{R}(\mathbf{x}_{i,t}), \text{ at time } \underline{t-\tau}}. \end{aligned} \quad (16)$$

Regions with a high error in the past but a lower error now, *exactly* those that are promoted by PLP, get a *negative* reducible error! They have shown (relatively) high learning progress, meaning that the model was improved there *already*, and therefore the model is now expected to be improved more in regions that featured *low* or *negative* learning progress.

The irreducible error has to be initialized to a low value for *new* regions that have not been visited yet for the principle to work (in the used regional error calculation this is accounted for, as regions that have not been sampled closely yet get errors near zero). A region will then feature *high reducible error* upon first sampling: High *current* error, and low past, *irreducible* error. In *unlearnable* regions the current error will remain high, making that the past error becomes high as well after repeated sampling, resulting in a *low reducible error*.

3.2 Regional Error Value

Most of the interestingness measures in Section 3.1 are intended to be calculated by comparing regional error *averages* $< \mathrm{Err}(\mathcal{R}, \mathcal{M}, V) >$ over two time *periods* [14, 1, 12], instead of regional error values $\mathrm{Err}(\mathcal{R}, \mathcal{M}, V)$ at two singular time steps. First we lay out an approach that derives regional error values from the modeling errors that occurred in the test points, then we give an overview of different ways of obtaining two distinct regional error averages.

Instead of exclusively assigning test points $\mathbf{v}$ to regions, the influence of a test point's modeling error $\mathrm{ME}(\mathbf{v}, \mathcal{M})$ is determined based on the probability density function of the standard normal distribution $\mathcal{N}(\mathbf{v}, \mathbf{I})$ around $\mathbf{v}$. Hence, perhaps a bit contradictory, a region $\mathcal{R}(\mathbf{x})$ is defined merely by the position of $\mathbf{x}$ that indicates the influence of the test point's modeling errors. The regional error value is calculated as the *maximum* of the modeling errors $\mathrm{ME}(\mathbf{v}, \mathcal{M})$ in points $\mathbf{v}$ from the test set V, times each point's influence (see Fig. 3),

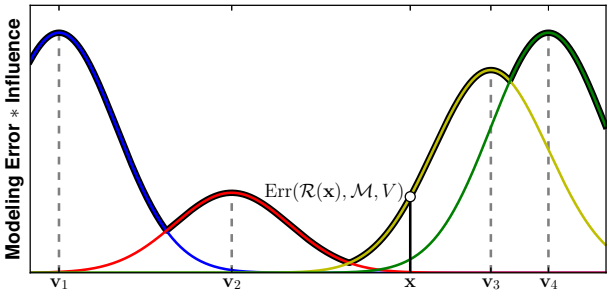


Figure 3: *Regional Error Value.* The regional error $\mathrm{Err}(\mathcal{R}(\mathbf{x}), \mathcal{M}, V)$ in region $\mathcal{R}(\mathbf{x})$, defined by the position of $\mathbf{x}$, is calculated as the maximum of the modeling errors in test points $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}$ times their normally-distributed influence.

$$\mathrm{Err}(\mathcal{R}(\mathbf{x}), \mathcal{M}, V) = \max_{\mathbf{v} \in V} \left(\mathrm{ME}(\mathbf{v}, \mathcal{M}) \cdot \mathrm{pdf}_{\mathcal{N}(\mathbf{v}, \mathbf{I})}(\mathbf{x})\right), \quad (17)$$

where

$$\mathrm{pdf}_{\mathcal{N}(\boldsymbol{\mu}, \mathbf{I})}(\mathbf{x}) = (2\pi)^{-\frac{1}{2}|\boldsymbol{\mu}|} \cdot \exp(-\tfrac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \mathbf{I} (\mathbf{x} - \boldsymbol{\mu})). \quad (18)$$

The result is that *all* regions get an error value at each time step for which training data is available, anywhere in the search space, and if no training data is available at time u, $\mathrm{Err}(\mathcal{R},\ \mathcal{M}(T_u = \emptyset),\ V_u)$ is set to 0.

Singular Errors

The first option for obtaining regional error averages at time steps u and $u - \tau$ is simply adopting the non-averaged, singular error values at those time steps,

$$< \mathrm{Err}(\mathcal{R}, \mathcal{M}(T_u), V_u) >= \mathrm{Err}(\mathcal{R}, \mathcal{M}(T_u), V_u), \quad (19)$$

where the test sets V_u and $V_{u-\tau}$ are defined by the used interestingness measure.

Separated Averages

Oudeyer et al. [14] propose calculating the mean error over two time periods with window size w that lie τ time steps apart. The mean errors, over periods ending in time step u and in $u - \tau$, are calculated as follows,

$$< \mathrm{Err}(\mathcal{R}, \mathcal{M}(T_u), V_u) >= \frac{1}{w} \sum_{k=0}^{w-1} \mathrm{Err}(\mathcal{R}, \mathcal{M}(T_{u-k}), V_{u-k}). \quad (20)$$

Based on [14] we use $w = \min\left(15, \left\lfloor \frac{1}{2} \cdot u \right\rfloor\right)$ and $\tau = \min(25, u - w)$, where w and τ grow until the listed values are possible.

Adjacent Averages

In [1] Baranes and Oudeyer propose a simplification of this scheme by assigning the $2 \cdot w$ most recent time steps to the two error averages, taking $w = \left\lfloor \frac{1}{2} \cdot \min(40, u) \right\rfloor$ and $\tau = w$.

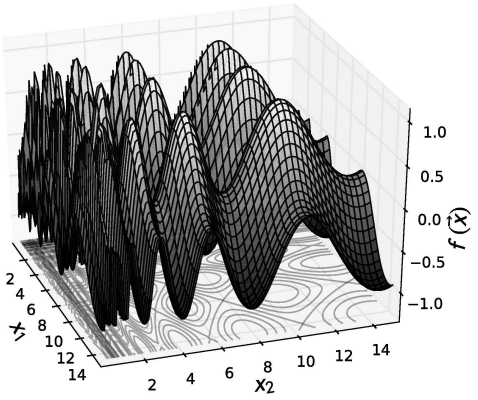


Figure 4: *Sine Non-stationary Variance.* Test function featuring exponentially increasing variance in its function landscape for input variables approaching zero.

Short-term/Long-term Memory

Luciw et al. [12] use *exponential smoothing* to obtain error averages in the form of two error *memories*, one with a slower forgetting rate (Mem_{long}) than the other ($\text{Mem}_{\text{short}}$),

$$\begin{aligned}\text{Mem}_{\text{long}}(\mathcal{R},u) &= (1-\eta_{\text{long}})\cdot\text{Mem}_{\text{long}}(\mathcal{R},u-1)+\\ &\quad \eta_{\text{long}}\cdot\text{Err}(\mathcal{R},\mathcal{M}(T_u),V_u),\\ \text{Mem}_{\text{short}}(\mathcal{R},u) &= (1-\eta_{\text{short}})\cdot\text{Mem}_{\text{short}}(\mathcal{R},u-1)+\\ &\quad \eta_{\text{short}}\cdot\text{Err}(\mathcal{R},\mathcal{M}(T_u),V_u),\end{aligned}\tag{21}$$

using $\eta_{\text{long}} = 0.1$ for forgetting slowly, and $\eta_{\text{short}} = 0.2$ for forgetting quicker. $\text{Mem}_{\text{long}}(\mathcal{R},u)$ is used as the regional error average at time $u-\tau$, and $\text{Mem}_{\text{short}}(\mathcal{R},u)$ is used as regional error average at time u.

4. EXPERIMENTS

We analyze the described novelty and interestingness measures on a two-dimensional instance of the *Sine Non-stationary Variance* (SineNSV) test function [21] (see Fig. 4),

$$f(\mathbf{x}) = -\frac{1}{n}\sum_{i=1}^{n}\sin(10\cdot\ln(x_i)).\tag{22}$$

Within the used search domain $[0.05, 15]^{n=2}$ SineNSV has 181 global extrema, occurring in exponentially increasing frequency when moving toward the $(0.05, 0.05)$-corner.

As we want exploration to focus more on areas featuring higher variance in quality, we determine the *sample Pearson correlation coefficient* between SineNSV's distribution of extrema, and the placement of sample points in the search space that the used exploration measure gives rise to. The search space is per dimension equally divided in 10 intervals, leading to 100 boxes and hence two ordered sets of 100 values, expressing either the number of extrema or sampled points per box, between which the correlation is determined.

In the test runs we select on exploration measure only, the quality function is not (directly) involved. The used iterative search method has to be able to cope with the dynamic search landscape that changes in each iteration, therefore we use a relatively simple isotropic Evolution Strategy that has a single stepsize per individual that is self-adapted through mutation with learning rate $\tau = \frac{1}{\sqrt{2n}}$, population sizes of $(2, 14)$, uses no recombination, and that is allowed to run for 500 generations. We initialize in the $(15, 15)$-corner of the search space with an initial stepsize of $14.95 \cdot 10^{-3}$. For constraint handling we employ a transformation function $\text{T}_{[0.05,15]}(x_i)$ [16] that reflects the object variables x_1, x_2 back at a constraint boundary, or, if the distance to the boundary is larger than 14.95, results in uniform-random placement in feasible space.

For the learning-based novelty measures we use two types of surrogate models: Feed-forward neural networks (FFNN) for PE and DP (using a *vertical* ensemble of 10 instances), and Kriging for PV. We emulate a *global* surrogate model using a *horizontal* ensemble (of, in case of DP, vertical ensembles) of *local* models trained using a sliding window of 5 generations of sample data [17], permanently retaining each 5th model. The novelty score is evaluated on all stored models and the *lowest* of the outcomes is adopted.

The FFNN have linear outputs and a hidden layer of 10 sigmoidally-activated nodes, a bias node, and are trained using improved Rprop [8] for a maximum of 10 epochs from initial connections weights in $[-0.1, 0.1]$. They are fully connected, including direct connections from input to output nodes. All data is used for training, no validation set is used. Ordinary Kriging is employed, fitted through CLF-maximization by tuning the θ_i in $[10^{-10}, 10^2]$ and the p_i in $[1, 2]$ (power-exponential multiparameter kernel), and allowing the λ to vary in $[-10, 0]$. Fitting is done using Differential Evolution with population size 50, for 50 generations, followed by a Hill Climber phase.

All techniques and measures have been implemented using the Shark Machine Learning Library v2.3.3[3].

4.1 Results

The outcome of the experiments is summarized in Fig. 5. A selection of 10 interestingness and error calculation combinations was tested for PE, and 3 combinations for DP and PV. In terms of quality evaluations, the PE-based measures used a count of 6987, DP and PV-based 999, while Sp and Un did not require quality evaluations with their solution-vector-based distance measure.

First of all, judging by the median scores, apparently none of the measures reliably focuses on regions with higher variance in quality. Furthermore, there is not one measure that significantly outperforms the others when it comes to the used performance measure. However, these scores do not fully capture the difference in sampling behavior between the different measures (see Fig. 6). Sp and Un sample everywhere with similar density, while PE, DP, and PV get stuck, even in a relatively simple part of the search space. This is surmised to be due to interaction with stepsize adaptation: Regions with higher novelty become invisible due to the limited search horizon induced by a small step size. ALP and PLP interestingness show uniform sampling as well, while PCC (basically alternation between PLP and RE) and especially RE show clear spatial separation between samples over time. Furthermore, RE based on DP does not perform worse than its PE counterparts, while only using 1/7th of the number of quality evaluations.

[3] `http://shark-project.sourceforge.net`

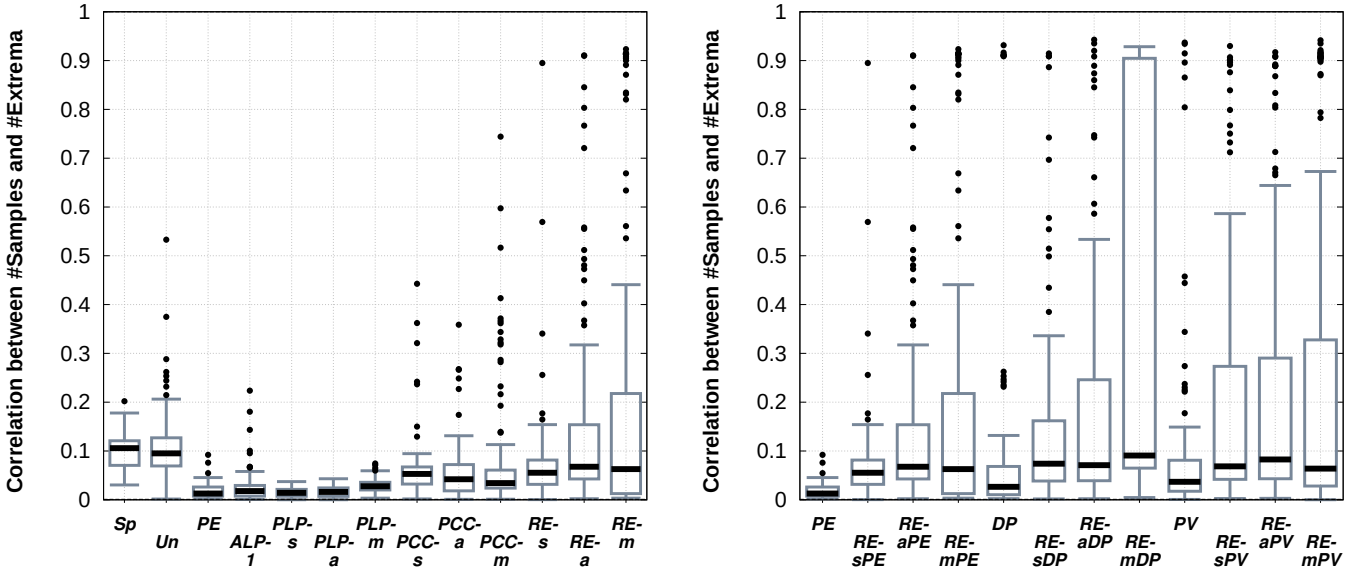
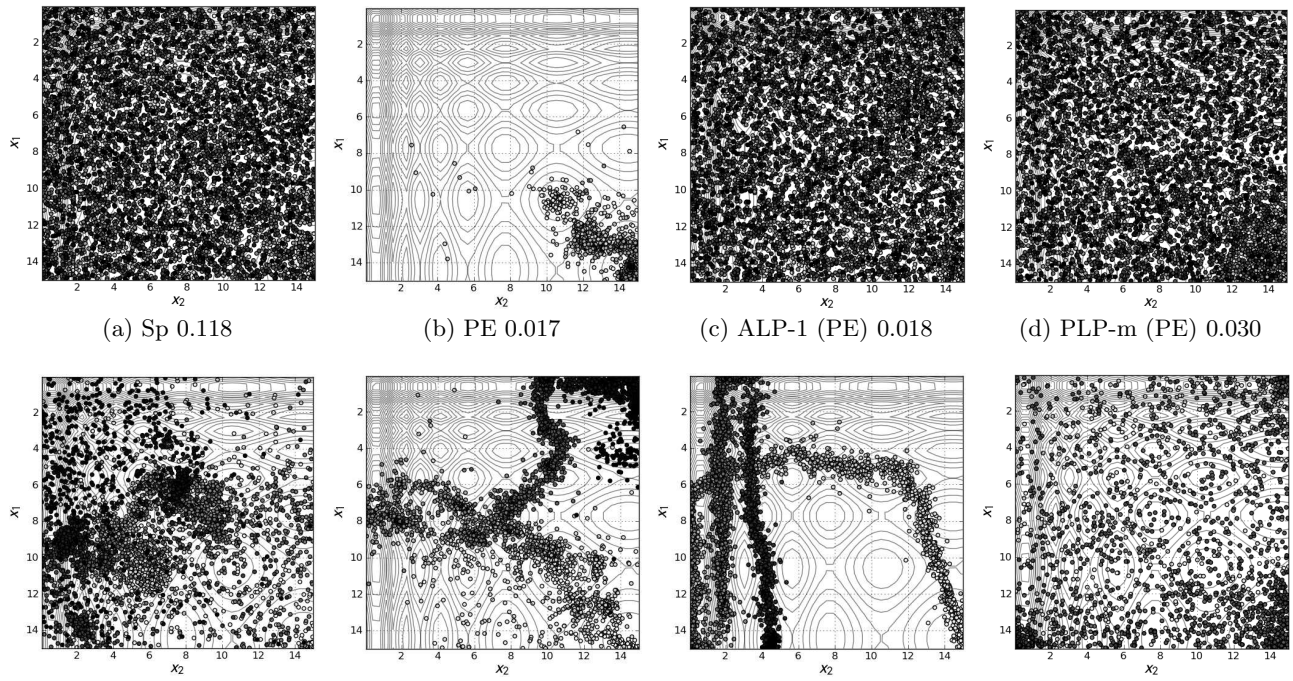


Figure 5: ***Results on Sine Non-stationary Variance.*** **Per exploration measure, the correlation score between the distribution of the solutions sampled during a single run and the distribution of the test function's extrema is determined. The box plots summarize the scores over 100 runs per method, showing the statistics over 100 single-run scores. The *left* figure shows a comparison between distance-based novelty, PE, and interestingness based on PE. The *right* figure shows a comparison between PE, DP, and PV, and interestingness measure RE based on each of these. Used error calculation is indicated by 1 (singular), s (separated), a (adjacent), m (memory).**



(a) Sp 0.118 (b) PE 0.017 (c) ALP-1 (PE) 0.018 (d) PLP-m (PE) 0.030

(e) PCC-s (PE) 0.051 (f) RE-s (PE) 0.056 (g) RE-m (DP) 0.104 (h) RE-a (PV) 0.092

Figure 6: ***Sampling Behavior on Sine Non-stationary Variance.*** **Per selected exploration measure, a temporal plot is shown of all solutions generated by a single run with median *correlation between samples and extrema*-score (dark: lastly sampled solutions, bright: firstly sampled solutions).**

5. DISCUSSION

This paper primarily provides an overview of novelty and interestingness measures, grouped together, described in unifying notation, and adapted to be applied in design-space exploration by proposing a regional error calculation scheme. While the view on how to maximize learning progress differs strongly between authors, it turns out that the mathematical formulations of their interestingness measures are nevertheless highly related.

Secondarily, for verification and initial analysis, the described measures are tested on a low-dimensional test case of a quality landscape with exponentially increasing non-stationary variance. As performance measure the focus of the methods on areas with higher number of extrema is evaluated. In the tested setup, none of the methods is sufficient for focusing on locally higher variance. Even though areas with high variance are more difficult to model, the interestingness measures do not generate more samples there. This is either inherent to their formulation, or could be due to insufficient complexity in the used test case.

On the other hand, interestingness expressed as reducible error (RE) shows clearly directed sampling in the search space. Furthermore, RE based on dispersion in predictions (DP) performs similarly to its prediction error (PE) counterpart at much lower evaluation count.

Further study is required to determine the behavior of the different measures in the presence of non-stationary noise and in higher-dimensional test cases, including real-world design problems. They are to be integrated in quality-based search, requiring appropriate schemes of combining quality and exploration measure. Moreover, in real-world scenarios the strength of distance-based novelty is better captured through the use of a problem-specific distance measure, on which the regional error calculation could be based as well.

6. ACKNOWLEDGMENTS

The authors thank Giles Endicott for providing the used Kriging implementation.

7. REFERENCES

[1] A. Baranes and P.-Y. Oudeyer. Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots. *Robotics and Autonomous Systems*, 61(1):49–73, 2013.

[2] G. Cuccu, F. Gomez, and T. Glasmachers. Novelty-Based Restarts for Evolution Strategies. In *CEC 2011*, pages 158–163. IEEE, 2011.

[3] G. Cuccu, M. Luciw, J. Schmidhuber, and F. Gomez. Intrinsically Motivated NeuroEvolution for Vision-Based Reinforcement Learning. In *ICDL 2011*. IEEE, 2011.

[4] M. Emmerich. *Single- and Multi-objective Evolutionary Design Optimization Assisted by Gaussian Random Field Metamodels*. PhD thesis, Department of Computer Science, TU Dortmund, Germany, 2005.

[5] L. Graening, N. Aulig, and M. Olhofer. Towards Directed Open-ended Search by a Novelty Guided Evolution Strategy. In R. Schaefer et al., editor, *PPSN XI, LNCS, vol. 6239*, pages 71–80. Springer, 2010.

[6] V. Graziano, T. Glasmachers, T. Schaul, L. Pape, G. Cuccu, J. Leitner, and J. Schmidhuber. Artificial curiosity for autonomous space exploration. *Acta Futura*, 4:41–51, 2011.

[7] T. Hester and P. Stone. Intrinsically Motivated Model Learning for a Developing Curious Agent. In *ICDL 2012*. IEEE, 2012.

[8] C. Igel and M. Hüsken. Empirical evaluation of the improved Rprop learning algorithms. *Neurocomputing*, 50:105–123, 2003.

[9] D. Jones, M. Schonlau, and W. Welch. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4):455–492, 1998.

[10] J. Lehman and K. Stanley. Abandoning Objectives: Evolution through the Search for Novelty Alone. *Evolutionary Computation*, 19(2):189–223, 2011.

[11] D. Lim, Y. Jin, Y.-S. Ong, and B. Sendhoff. Generalizing Surrogate-Assisted Evolutionary Computation. *IEEE Transactions on Evolutionary Computation*, 14(3):329–355, 2010.

[12] M. Luciw, V. Graziano, M. Ring, and J. Schmidhuber. Artificial Curiosity with Planning for Autonomous Perceptual and Cognitive Development. In *ICDL 2011*. IEEE, 2011.

[13] P.-Y. Oudeyer and F. Kaplan. What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1(6):1–14, 2007.

[14] P.-Y. Oudeyer, F. Kaplan, and V. Hafner. Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286, 2007.

[15] C. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT, 2006.

[16] E. Reehuis and T. Bäck. Mixed-Integer Evolution Strategy Using Multiobjective Selection Applied to Warehouse Design Optimization. In J. Branke et al., editor, *GECCO'10*, pages 1187–1194. ACM, 2010.

[17] E. Reehuis, J. Kruisselbrink, M. Olhofer, L. Graening, B. Sendhoff, and T. Bäck. Model-guided Evolution Strategies for Dynamically Balancing Exploration and Exploitation. In J.-K. Hao et al., editor, *Artificial Evolution 2011, EA 2011*, pages 306–317, 2011.

[18] T. Schaul, Y. Sun, D. Wierstra, F. Gomez, and J. Schmidhuber. Curiosity-driven Optimization. In *CEC 2011*, pages 1343–1349. IEEE, 2011.

[19] J. Schmidhuber. Curious Model-Building Control Systems. In *1991 International Joint Conference on Neural Networks, IJCNN'91*, pages 1458–1463. IEEE, 1991.

[20] J. Schmidhuber. What's interesting? Technical Report IDSIA-35-97, IDSIA, Lugano, Switzerland, 1997.

[21] O. Shir, M. Emmerich, and T. Bäck. Adaptive Niche Radii and Niche Shapes Approaches for Niching with the CMA-ES. *Evolutionary Computation*, 18(1):97–126, 2010.

[22] Y. Xiong, W. Chen, D. Apley, and X. Ding. A non-stationary covariance-based Kriging method for metamodelling in engineering design. *Numerical Methods in Engineering*, 71(6):733–756, 2007.