Algorithm 1 generating time series with RSGP

Require: $0 \leq \alpha \leq 1$

Ensure: $K(t,t') = K_{RSGP} = \alpha K_{GP} + (1-\alpha) K_{RS}$

for $t := 1, ..., T$ **do**

1. Given the previous value and reward history, infer the distribution from the conditional probability:

$$[x^1, ..., x^{t-1}, x^t] \sim \mathcal{N}\big(0, \big[K_{t-1,t-1}, K_{t-1,1}; K_{1,t-1}, K_{1,1}\big]\big)$$

2. Sample x^t from the distribution obtained in step 1.
3. Update the reward, K_{RS} and K based on new sample x^t

end for

$$K_{RSGP}(t, t-s) = \sigma_{SE}^2 K_{SE}(t, t-s) + \sigma_{RS}^2 K_{RS}(t, t-s) + \sigma_0^2 I$$

$$K_{SE}(t, t-s) = \exp(-\frac{s^2}{2l_{SE}^2})$$

$$K_{RS}(t, t-s) = \begin{cases} \exp(-\frac{s^2}{2l_{RS}^2}) & \textbf{if trial t-s was rewarded} \\ 0 & \textbf{otherwise} \end{cases}$$

Algorithm 2 using reward gradient search

for $t := 1, ..., T$ **do**

1. Gradient derived from the previous performance and reward

$$\Delta x = \alpha(x^{t-1} - x^{t-2})(r^{t-1} - r^{t-2})$$

2. Generate x^t with the added scalar production noise $\epsilon \sim \mathcal{N}(0, \sigma_p)$

$$x_n = x^{t-1} + \Delta x + \epsilon$$

3. Compute the amplitude of reward r^t based on x^t and reward profile.

end for

Algorithm 3 Markov chain Monte Carlo sampling

for $t := 1, ..., T$ **do**

1. Keep in memory a target variable x^* that is currently highest scored according to the value estimate $V(x)$

2. One this trial, sample the target variable x^+ from a Gaussian distribution in the vicinity of x^* and standard deviation σ_e

3. Generate a new x by sampling from a Gaussian distribution with mean x^+ and standard deviation σ_p (scalar noise). Assign the reward as the value of new sample $V(x^+)$.

4. Use a probabilistic Metropolis-Hastings rule to accept or reject x^+ as the new x^*. The acceptance probability is

$$\mathbf{P}(\textbf{accept}\, x^+) = \frac{e^{\beta V(x^+)}}{e^{\beta V(x^+)} + e^{\beta V(x^*)}}$$

where β is a free parameter controlling the volatility, known as the inverse temperature in Reinforcement learning.

end for