Simulation of foraging behavior using a decision-making agent with Bayesian and inverse Bayesian inference: Temporal correlations and power laws in displacement patterns

Shuji Shinohara[a*], Hiroshi Okamoto[a], Nobuhito Manome[a,b], Pegio-Yukio Gunji[c], Yoshihiro Nakajima[d], Toru Moriyama[e], and Ung-il Chung[a]

[a] Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

[b] Department of Research and Development, SoftBank Robotics Group Corp., Tokyo, Japan

[c] Department of Intermedia Art and Science, School of Fundamental Science and Technology, Waseda University, Tokyo, Japan

[d] Graduate School of Economics, Osaka City University, Osaka, Japan

[e] Faculty of Textile Science, Shinshu University, Ueda, Japan

*** Corresponding author**

Email: shinohara@bioeng.t.u-tokyo.ac.jp

Postal address: Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Keywords

Bayesian and inverse Bayesian inference, power-law distribution, exponential distribution, foraging behavior, decision-making agent

Author contributions

Shuji Shinohara: Conceptualization, Formal analysis, Methodology, Software, Writing—Original draft preparation, Funding acquisition. **Hiroshi Okamoto:** Conceptualization, Writing—Review and editing. **Nobuhito Manome:** Writing—Review and editing. **Pegio-Yukio Gunji:** Writing—Review and editing, Supervision. **Yoshihiro Nakajima:** Writing—Review and editing. **Toru Moriyama:** Writing—Review and editing. **Ung-il Chung:** Writing—Review and editing, Supervision, Project Administration.

Abstract

It has been stated that in human migratory behavior, the step length series may have temporal correlation and that there is some relationship between this time dependency and the fact that the frequency distribution of step length follows a power-law distribution. Furthermore, the frequency of occurrence of the step length in some large marine organisms has been found to switch between power-law and exponential distributions, depending on the difficulty of prey acquisition. However, to date it has not been clarified how the aforementioned three phenomena arise: the positive correlation created in the step length series, the relation between the positive correlation of the step length series and the form of an individual's step length distribution, and the switching between power-law and exponential distributions depending on the abundance of prey. This study simulated foraging behavior using the Bayesian decision-making agent simultaneously performing both knowledge learning and knowledge-based inference to analyze how the aforementioned three phenomena arise. In the agent with learning and inference, past experiences were stored as hypotheses (knowledge) and they were used in current foraging behavior; at the same time, the hypothesis continued to be updated based on new experiences. The simulation results show that the agent with both learning and inference has a mechanism that simultaneously causes all the phenomena.

1. Introduction

Lévy walks are found in the migratory behavior of organisms at various levels, from bacteria and T cells to humans [1–6]. The Lévy walk is a random walk where the frequency of occurrence of a linear step length l follows a power-law distribution $P(l) \sim l^{-\eta}, 1 < \eta \leq 3$. Compared with the Brownian walk, which is also a random walk (the frequency of occurrence of the step length l is characterized by an exponential distribution (EP) $P(l) \sim e^{-\lambda l}$), the Lévy walk is characterized by the appearance of linear movements over extremely long distances.

The reason behind the formation of such patterns in the movement of organisms has been investigated [7], and it has been found that they can be explained by the Lévy-flight foraging hypothesis [8,9]. This hypothesis states that when prey is sparse and randomly distributed and predators have no information (memory) about the prey, the Lévy walk becomes the optimal foraging behavior, whereas Brownian motion is theoretically optimal when prey is abundant [10,11]. In fact, it has been indicated that in a few marine organisms, when prey is abundant, a Brownian walk is observed, whereas when prey is scarce, a Lévy walk is observed [4,12]. Mechanisms for switching between Lévy and Brownian behavior have been the focus of considerable attention [13].

Humans have advanced learning and estimation capabilities and can acquire useful information through systematic search even in dynamic and uncertain nonstationary environments [14]. Ross et al. [15] demonstrated that in human exploration behavior, the mode of exploration changes depending on the encounter with prey. Particularly, they indicated that in response to encounters, hunters produce a more tortuous search of patches of higher prey density and spend a large part of their search time in patches; however, they adopt more efficient unidirectional, inter-patch movements after failing to encounter prey over a sufficiently long period. This type of search behavior is called area-restricted search (ARS) [15,16].

Wang et al. [17], Rhee et al. [18], and Zhao et al. [19] demonstrated that the temporal variation of step length is autocorrelated; this means that there is a trend in the time variation of step length, such that short

(long) steps are followed by short (long) steps. Such a behavioral pattern, even if the frequency of occurrence of step length follows a power-law distribution, is not a pure random walk and therefore may not be called a Lévy flight. In addition, it has been emphasized that the dependence of the step length on time is related to the frequency distribution of the step length following a power-law distribution [17]. It is necessary to clarify the mechanisms that create these relationships.

Perhaps, in organisms with advanced cognitive functions, such as humans, various movement patterns may appear as a result of learning and inference through interaction with the environment, rather than as a result of random search [14,20]. In this study, we simulated the foraging behavior of predators based on the following two assumptions: 1) prey exist in two-dimensional space according to a density distribution like a normal distribution, and 2) predators have the ability to learn and infer the prey density distribution based on their own experience in acquiring prey.

Consider a situation where a predator estimates the distribution of prey using Bayesian inference—a machine learning method. In Bayesian inference, several candidate distributions are first prepared as hypotheses about how the prey are distributed. Next, confidence in each hypothesis based on observed data (experience of where and how many of the prey are actually acquired) is updated. It would be quite possible for predators to memorize locations where they have caught much food in the past as candidates (hypotheses) for foraging locations and use them in current and future foraging behavior. However, it is not possible to prepare an appropriate hypothesis at the initial stage when there is no information on prey. In addition, if the prey is an animal, it may move around, so there is no guarantee that the prey will be in the same place in the future. Furthermore, the consumption of food by predators could deplete the food supply in the area. In such a case, the hypothesis itself needs to be reformulated or revised based on observed data.

Gunji et al. [21-23] proposed a novel mathematical model called Bayesian and inverse Bayesian (BIB) inference that forms hypotheses based on observed data and subsequently performs inference using the hypotheses and observed data; in other words, learning and inference are performed simultaneously and seamlessly. The process of simultaneous learning and inference, similar to that in BIB inference, has been

suggested in human-related experiments. Zaki and Nosofsky conducted a behavioral experiment that demonstrated the influence of unlabeled test data on learning [24]. Their experiment demonstrated that humans do not fix their hypothesis after training; instead, unlabeled test data influence the learned hypotheses of humans [25].

In this study, we propose a novel BIB inference model based on Bayesian inference with discount rate [26,27] as an algorithm for inference and the expectation-maximization (EM) algorithm [28,29], an unsupervised clustering method, as an algorithm for learning (inverse Bayesian inference). The model is used as a decision-making model for foraging agents. On the one hand, Bayesian inference is an algorithm that identifies the hypothesis that best explains the observed data, from one of several hypotheses (generating distributions) prepared in advance. On the other hand, the EM algorithm is an algorithm that, by modifying the initial distribution based on the observed data, approaches the distribution that best explains the observed data. Bayesian inference and the EM algorithm have the same goal of estimating the most appropriate generative model based on observed data. However, the direction of the estimation is opposite. If the estimated generative model is considered as knowledge, then the EM algorithm can be assumed to be a type of learning because it forms and modifies knowledge based on observed data. Further, Bayesian inference can be considered to be knowledge-based inference because it decides the best among the several hypotheses (generative models) prepared in advance.

In this study, we simulated the foraging behavior of BIB agents and analyzed the following phenomena: 1) the positive correlation in the step length time series, 2) the relationship between the positive correlation and the shape of the frequency distribution of the step length, and 3) the switching between the power law and the exponential distribution depending on the difficulty of food acquisition. The results of the foraging simulation showed that the mechanism of the BIB agent simultaneously causes the aforementioned phenomena.

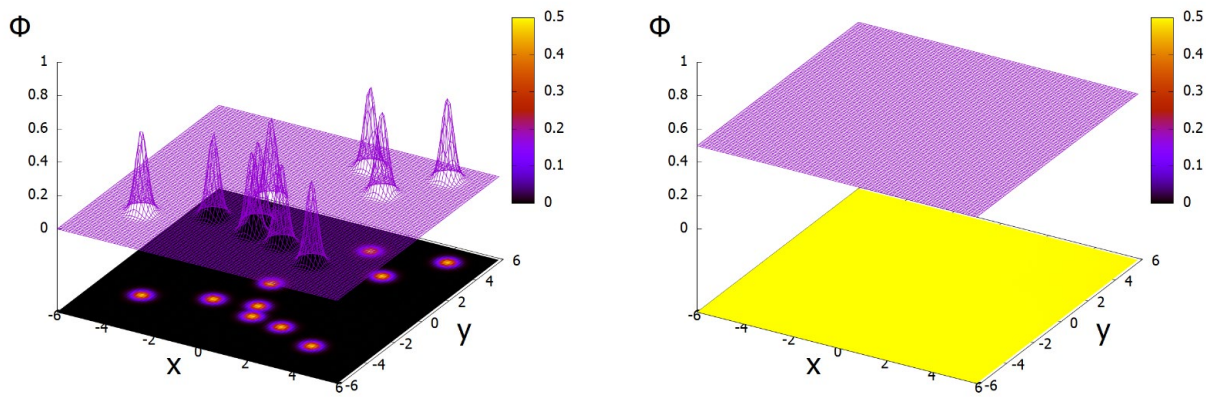
2. Materials and Methods

2.1. Prey density distribution

In this study, we consider a situation wherein the prey are distributed in a two-dimensional space. We will deal with the case where prey are unevenly distributed in patches, and the case where prey are randomly distributed with uniform probability throughout the space (Figure 1). As an example of the former, we may consider the case where the prey is a plant or an animal with a nest. The plants can be unevenly distributed depending on the amount of sunlight in the area, the nutrients in the land, and the ground conditions (soil or rock). Furthermore, considering that animals roam around the nest, the closer to the nest, the higher the probability that animals are present. For the purpose of simplicity, we assume that the density (probability of existence) of prey in each patch is normally distributed. In other words, the prey density distribution of the patch *p* at a certain point $d = (d_x, d_y)$ at time *t* is defined as follows:

$$\Phi_p^t(d|\mu_f^p, \Sigma_f^p) = w_p^t exp\left(-\frac{1}{2}(d-\mu_f^p)^T(\Sigma_f^p)^{-1}(d-\mu_f^p)\right) \quad (1)$$

where μ_f^p and Σ_f^p are the center and variance of patch *p*, respectively. w_p^t is a parameter that controls the total amount of prey of patch *p*, at time *t*. The distribution takes a maximum value when the condition $d = \mu_f^p$ is satisfied, and the value is $\Phi_p^t(d|\mu_f^p, \Sigma_f^p) = \Phi(\mu_f^p|\mu_f^p, \Sigma_f^p) = w_p^t$. Moving away from the center μ_f^p of the distribution, the prey density of decreases. Σ_f^p is a variance parameter, and the larger $|\Sigma_f^p|$ is, the more widely the prey are distributed and the larger is the total number of prey.



(a) **(b)**

Figure 1. Examples of prey density distribution. Predator probabilistically acquires prey in proportion to the prey density at the foraging point. (a) Distribution in the case of $w_p^t = 0.5, \Sigma_f^p = 0.05$. (b) Distribution in the case of $w_p^t = 0.5, \Sigma_f^p = 5000.0$. This represents an approximate uniform distribution.

2.2. Predator agent with Bayesian and inverse Bayesian (BIB) inference

2.2.1 Overview

The main objective of the agent is to acquire as many preys as possible within a certain period. The agent has the ability to learn and estimate the prey density distribution in order to acquire food efficiently. The agent has several candidate of prey density distributions as hypotheses, and the correct hypotheses are estimated based on the food acquisition information. In other words, the agent updates its confidence in each hypothesis. At the same time, the agent modifies the contents of each hypothesis, that is, the parameters (model) of the prey density distribution corresponding to each hypothesis. The Bayesian inference is used to update the confidence level. The EM algorithm is used to modify the hypothesis models. These are described in detail in the next subsection.

The agent decides probabilistically at which point $d^t = \left(d_x^t, d_y^t\right)$ to forage according to the model of the

hypothesis h_{max}^t, with the highest confidence at time *t*, as follows:

$$d^t \sim N(d|\mu_{max}^t, \Sigma_{max}^t) \qquad (2)$$

where $N(d|\mu_{max}^t, \Sigma_{max}^t)$ is the model of the hypothesis wherein the agent is most confident, and is represented by a normal distribution with μ_{max}^t as the centre and Σ_{max}^t as the variance. Next, whether a prey is acquired is probabilistically determined based on the prey density at point d^t.

$$success^t = \begin{cases} 1 & if\ rnd \leq \Phi_{\hat{p}}^t\left(d^t|\mu_f^{\hat{p}}, \Sigma_f^{\hat{p}}\right) \\ 0 & else \end{cases} \quad (3)$$

where *rnd* represents a uniform random number in the [0,1] interval; $success^t = 1$ represents successful prey acquisition, and $success^t = 0$ represents unsuccessful prey acquisition. The foraging patch $\hat{p}$ is defined as the patch with the highest prey density at the foraging site d^t. In other words,

$$\hat{p} = argmax_p\, \Phi_p^t\left(d^t|\mu_f^p, \Sigma_f^p\right) \qquad (4)$$

It is assumed that the agent consumes food every time it acquires, and the prey amount is set to decrease by Δ. If the agent acquires prey at the location d^t at time *t*, then $w_{\hat{p}}^t$ in the patch $\hat{p}$ decreases as follows:

$$w_{\hat{p}}^{t+1} \leftarrow w_{\hat{p}}^t - \Delta \qquad (5)$$

The following points should be noted here. Φ, defined in Equation (1), is the actual prey density distribution, which varies depending on the consumption of food. On the other hand, the *N* used in Equation (2) is the prey distribution estimated by the agent. Therefore, this *N* changes according to the agent's learning and estimation based on the information of food acquisition.

2.2.2. Sequential discounting expectation-maximization (SDEM)

In this subsection, we describe how to modify the hypothetical model. The EM algorithm is used to modify the model. The EM algorithm [28] is a machine learning technique used for unsupervised clustering. The classification of a given set of *n* data $\{d^1, d^2, \cdots, d^n\}$ into *M* clusters is considered here. First, a mixture normal distribution comprising *M* normal distributions is considered for generating data d.

$$P(d) = \sum_{m=1}^{M} \pi_m N(d|\mu_m, \Sigma_m) \qquad (6)$$

where d is a D-dimensional data vector and $N(d|\mu_m, \Sigma_m)$ is a D-dimensional normal distribution with mean μ_m and variance-covariance matrix Σ_m, as shown below.

$$N(d|\mu_m, \Sigma_m) = \frac{1}{\sqrt{(2\pi)^D |\Sigma_m|}} exp\left(-\frac{1}{2}(d-\mu_m)^T {\Sigma_m}^{-1}(d-\mu_m)\right) \qquad (7)$$

Here, $\sum_{m=1}^{M} \pi_m = 1$ is assumed. π_m is called the "mixing weights" and represents the weight of each normal distribution. The aim of the EM algorithm is to estimate the mixing weights π_m, mean μ_m, and variance Σ_m of *M* normal distributions from *n* data $\{d^1, d^2, \cdots, d^n\}$. First, in Step E, the responsibility is calculated. The responsibility for the normal distribution *m* of data d^i is calculated as follows:

$$q_m^i = \frac{\pi_m N(d^i|\mu_m, \Sigma_m)}{\sum_{j=1}^{M} \pi_j N(d^i|\mu_j, \Sigma_j)} \qquad (8)$$

Next, the weights of data d^i with respect to the normal distribution *m* are defined as follows:

$$w_m^i = \frac{q_m^i}{\sum_{j=1}^{n} q_m^j} \qquad (9)$$

Next, in Step M, the mixing weights, mean, and variance of each normal distribution are updated.

$$\pi_m \leftarrow \frac{\sum_{i=1}^{n} q_m^i}{n} \qquad (10)$$

$$\mu_m \leftarrow \sum_{i=1}^{n} w_m^i d^i \qquad (11)$$

$$\Sigma_m \leftarrow \sum_{i=1}^{n} w_m^i (d^i - \mu_m)(d^i - \mu_m)^T \qquad (12)$$

In the EM algorithm, the values converge by alternately repeating Steps E and M. However, in the EM algorithm, it is necessary to provide all observational data simultaneously. In practice, there are cases where data processing must be performed sequentially, each time after the data are observed. Various online algorithms have been proposed to address this situation [30–37]. For example, Yamanishi et al. [29] proposed an SDEM algorithm. In the SDEM algorithm, Steps E and M are executed once for each data observed sequentially. In addition, the SDEM algorithm introduces the effect of forgetting to deal with unsteady situations where the inferred target changes. Specifically, weighting is performed to weaken the

influence of the older observation data by introducing a discount rate β (0<β<1). When data d^t are observed at time t, the mixing weights, mean, and variance of each normal distribution are updated, as shown below.

$$\tilde{\pi}_m^{t+1} \leftarrow (1-\beta)\tilde{\pi}_m^t + \beta q_m^t = \beta \sum_{i=1}^{t} (1-\beta)^{t-i} q_m^i \quad (13)$$

$$\tilde{\mu}_m^{t+1} \leftarrow (1-\beta)\tilde{\mu}_m^t + \beta q_m^t d^t = \beta \sum_{i=1}^{t} (1-\beta)^{t-i} q_m^t d^i \quad (14)$$

$$\tilde{\Sigma}_m^{t+1} \leftarrow (1-\beta)\tilde{\Sigma}_m^t + \beta q_m^t (d^t - \mu_m^t)(d^t - \mu_m^t)^T = \beta \sum_{i=1}^{t} (1-\beta)^{t-i} q_m^t \left(d^i - \mu_m^t\right)(d^t - \mu_m^t)^T \quad (15)$$

$$\pi_m^{t+1} \leftarrow \frac{\tilde{\pi}_m^{t+1}}{\sum_{j=1}^{M} \tilde{\pi}_j^{t+1}} \quad (16)$$

$$\mu_m^{t+1} \leftarrow \frac{\tilde{\mu}_m^{t+1}}{\tilde{\pi}_m^{t+1}} \quad (17)$$

$$\Sigma_m^{t+1} \leftarrow \frac{\tilde{\Sigma}_m^{t+1}}{\tilde{\pi}_m^{t+1}} \quad (18)$$

In this study, for simplicity, the case of a single normal distribution, that is, M=1, is treated as the generating distribution of the data. In the case of M=1, q_1^t and $\tilde{\pi}_1^t$ are always one. For simplicity, μ_1^t and Σ_1^t are denoted as μ^t and Σ^t, respectively. At this time, Equations (17) and (18) show that the new estimated value is obtained as a convex combination of the current estimated value and current observed data, as follows:

$$\mu^{t+1} \leftarrow (1-\beta)\mu^t + \beta d^t \quad (19)$$

$$\Sigma^{t+1} \leftarrow (1-\beta)\Sigma^t + \beta (d^t - \mu^t)(d^t - \mu^t)^T \quad (20)$$

Equations (19) and (20) represent the exponential moving averages of μ^t and Σ^t, respectively.

2.2.3. Bayesian inference with discount rate (BID)

In this subsection, we describe how to update the confidence level for each hypothesis. The Bayesian inference is used to update the confidence level. In Bayesian inference, first, several hypotheses (generative models) are prepared for an inferred target. Next, the evidence (the data generated by the target) is observed

and the confidence in each hypothesis is updated by quantitatively evaluating the fit between each hypothesis and the observed data. Finally, based on the confidence level, the single best hypothesis is determined.

In general, it is better to obtain as much observational data (information) as possible for the accurate estimation of the target. However, this is true only in a stationary environment. When the target varies dynamically, it is necessary to discard the distant past observation data for a more accurate estimation. To address this situation, BID is introduced here [26,27].

First, a discrete version of Bayesian inference is described. In Bayesian inference, first a number of hypotheses h_k are defined and a model (the generating distribution of the data) for each hypothesis is prepared in the form of a conditional probability $P(d|h_k)$. Specifically, d is assumed to be a D-dimensional vector, as in the case of SDEM, and the model $P(d|h_k)$ is assumed to be a D-dimensional normal distribution, as follows:

$$P(d|h_k) = N(d|\mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^D|\Sigma_k|}} exp\left(-\frac{1}{2}(d-\mu_k)^T {\Sigma_k}^{-1}(d-\mu_k)\right) \quad (21)$$

This conditional probability is called the likelihood if the data are fixed and considered as a function of the hypothesis h_k; that is, the parameter $\theta_k = (\mu_k, \Sigma_k)$ of the normal distribution. An initial confidence $P^1(h_k)$ value for each hypothesis is provided as a prior probability. If the confidence in hypothesis h_k at time t is $P^t(h_k)$ and data d^t are observed, the posterior probability $P^t(h_k|d^t)$ is calculated using Bayes' theorem, as follows:

$$P^t(h_k|d^t) = \frac{P^t(h_k)P(d^t|h_k)}{P^t(d^t)} \quad (22)$$

where $P^t(d^t)$ is the marginal probability of the data at time t and is defined as follows:

$$P^t(d^t) = \sum_k P^t(h_k)P(d^t|h_k) \quad (23)$$

The following Bayesian update accordingly replaces the posterior probability with the confidence the next time.

$$P^{t+1}(h_k) \leftarrow P^t(h_k|d^t) \qquad (24)$$

Combining Equations (22) and (24), we obtain Equation (25):

$$P^{t+1}(h_k) \leftarrow \frac{P^t(h_k)P(d^t|h_k)}{P^t(d^t)} \qquad (25)$$

The estimation proceeds by updating the confidence in each hypothesis by Equation (25), each time the data are observed. Here, the confidence in each hypothesis satisfies $\sum_k P^{t+1}(h_k) = 1$. Focusing on the recursive nature of $P^t(h_k)$, Equation (25) can be rewritten as follows:

$$P^{t+1}(h_k) \leftarrow P^1(h_k) \prod_{i=1}^{t} \frac{P(d^i|h_k)}{P^i(d^i)} \qquad (26)$$

Here, $P^i(d^i)$ is common to all hypotheses and can be regarded as a constant. Omitting the normalization process, Equation (26) can be written as follows:

$$P^{t+1}(h_k) \leftarrow P^1(h_k) \prod_{i=1}^{t} P(d^i|h_k) \qquad (27)$$

That is, the current confidence of each hypothesis is proportional to the prior probability multiplied by the likelihood of the data observed each time.

Next, the forgetting function to deal with unsteady situations is incorporated by introducing a discount factor into the Bayesian inference, as follows [26,27]. However, in the following, *C* is used instead of *P*, to distinguish between Bayesian inference and BIB inference, by introducing a discount factor. If the normalization process is omitted, the updating formula for the confidence of each hypothesis is defined as follows:

$$C^{t+1}(h_k) \leftarrow \left[\frac{C^t(h_k)}{C^t(d^t)}\right]^{1-\beta} C(d^t|h_k) \qquad (28)$$

where β is the discount factor and Equation (28) agrees with Equation (25) when $\beta = 0$. After Equation (28), normalization is conducted to satisfy $\sum_k C^{t+1}(h_k) = 1$, as follows:

$$C^{t+1}(h_k) \leftarrow \frac{C^{t+1}(h_k)}{\sum_j C^{t+1}(h_j)} \qquad (29)$$

If we substitute *t*-1 for *t* in equation (28), we get

$$C^t(h_k) \leftarrow \left[\frac{C^{t-1}(h_k)}{C^{t-1}(d^{t-1})}\right]^{1-\beta} C(d^{t-1}|h_k) \qquad (30)$$

Therefore, Equation (28) can be transformed using Equation (30) as follows:

$$\begin{aligned} C^{t+1}(h_k) &\leftarrow \left[\frac{\left[\frac{C^{t-1}(h_k)}{C^{t-1}(d^{t-1})}\right]^{1-\beta} C(d^{t-1}|h_k)}{C^t(d^t)}\right]^{1-\beta} C(d^t|h_k) \\ &= \left[\frac{[C^{t-1}(h_k)]^{1-\beta} C(d^{t-1}|h_k)}{C^t(d^t)[C^{t-1}(d^{t-1})]^{1-\beta}}\right]^{1-\beta} C(d^t|h_k) \qquad (31) \\ &= \frac{[C^{t-1}(h_k)]^{(1-\beta)^2} C(d^t|h_k)[C(d^{t-1}|h_k)]^{1-\beta}}{[C^t(d^t)]^{1-\beta}[C^{t-1}(d^{t-1})]^{(1-\beta)^2}} \end{aligned}$$

By repeating the above operation recursively, we finally obtain the following equation.

$$C^{t+1}(h_k) \leftarrow [C^1(h_k)]^{(1-\beta)^t} \prod_{i=1}^{t} \frac{[C(d^i|h_k)]^{(1-\beta)^{t-i}}}{[C^i(d^i)]^{(1-\beta)^{t+1-i}}} \quad (32)$$

In Equation (32), the denominator $C^i(d^i)$ on the right-hand side is common to each hypothesis and can be regarded as a constant; therefore, omitting the normalization process, the following equation is obtained:

$$C^{t+1}(h_k) \leftarrow [C^1(h_k)]^{(1-\beta)^t} \prod_{i=1}^{t} \left[C(d^i|h_k)\right]^{(1-\beta)^{t-i}} \qquad (33)$$

In other words, it is understood that the current confidence is multiplied by the likelihood in the distant past with less weight. When β = 0, using $C^{t+1}(h_k) \leftarrow C^1(h_k)C(d^1|h_k)C(d^2|h_k)\cdots C(d^t|h_k)$ in the Bayesian inference, the current likelihood and the likelihood in the distant past are evaluated with equal weight. When β = 1, $C^{t+1}(h_k) \leftarrow C(d^t|h_k)$ is used to calculate the next confidence using only the likelihood of the current observation. In the BID, the confidence for each hypothesis is updated according to Equation (28), each time data are observed, and the best hypothesis h_{max}^t is determined based on the confidence level.

$$h_{max}^t = argmax_{h_i} C^t(h_i) \qquad (34)$$

2.2.4. BIB inference

SDEM and BID are already explained. If the generative model is regarded as knowledge, SDEM can be

regarded as a type of learning because knowledge is formed and modified based on the observed data. BID, however, can be called knowledge-based inference because it decides which of the several hypotheses (generative models) prepared in advance is the best.

In this section, a model of BIB inference is proposed that employs SDEM and BID to perform learning and inference simultaneously. Here, BID and SDEM are the models for Bayesian inference and inverse Bayesian inference, respectively. This is a model where the hypotheses are revised based on inverse Bayesian inference, while Bayesian inference is used to assess which hypothesis is correct. Because the model $C(d|h_k)$ of the hypothesis h_k changes with time through inverse Bayesian inference, it is denoted by $C^t(d|h_k)$, with time t as a subscript.

It is assumed that food is obtained by foraging at point d^t at time t. At this time, the confidence of each hypothesis is updated using the BID equation:

$$C^{t+1}(h_k) \leftarrow \left[\frac{C^t(h_k)}{C^t(d^t)}\right]^{1-\beta} C^t(d^t|h_k) = C^t(h_k) + \left[\left[\frac{C^t(h_k)}{C^t(d^t)}\right]^{1-\beta} C^t(d^t|h_k) - C^t(h_k)\right] \quad (35)$$

On the other hand, when it fails to acquire food at d^t, the confidence level is updated as follows:

$$C^{t+1}(h_k) \leftarrow C^t(h_k) - \left[\left[\frac{C^t(h_k)}{C^t(d^t)}\right]^{1-\beta} C^t(d^t|h_k) - C^t(h_k)\right] \quad (36)$$

In other words, contrary to the successful case, the higher the likelihood of the hypothesis for d^t, the lower the confidence level in the hypothesis.

When the food is successfully acquired, the hypothesis h_{max}^t with the highest confidence is selected using Equation (34). SDEM is thereafter used to update the model $C^{t+1}(d|h_{max}^t) = N(d|\mu_{max}^{t+1}, \Sigma_{max}^{t+1})$, for the following hypothesis h_{max}^t, as follows.

$$\mu_{max}^{t+1} \leftarrow (1-\beta)\mu_{max}^t + \beta d^t \quad (37)$$

$$\Sigma_{max}^{t+1} \leftarrow (1-\beta)\Sigma_{max}^t + \beta(d^t - \mu_{max}^t)(d^t - \mu_{max}^t)^T \quad (38)$$

This is repeated every time data are observed.

The group of processes described in this section is summarized as the following algorithm:

1. Determine the parameters $\left(\mu_f^p, \Sigma_f^p\right)$ of the prey density distribution.
2. Set the values for parameters β, K.
3. Establish the initial values for $\theta_k^1 = (\mu_k^1,\ \Sigma_k^1),\ C^1(h_k) \quad (k = 1,2,\cdots K)$.
4. Repeat the following.
   - 4.1 The agent chooses the hypothesis h_{max}^t it believes in the most, according to Equation (34).
   - 4.2 Based on the model of the hypothesis h_{max}^t, the agent decides the foraging point d^t according to Equation (2).
   - 4.3 Equation (3) determines the success or failure of foraging.
   - 4.4 The agent executes the following processes simultaneously:
     - 4.4.1 Bayesian inference (BID): The confidence $C^t(h_k)$ of each hypothesis is updated according to Equation (35) or (36).
     - 4.4.2 Inverse Bayesian inference (SDEM): The model $C^t(d|h_{max}^t) = N(d|\mu_{max}^t, \Sigma_{max}^t)$ of the hypothetical h_{max}^t is modified according to Equations (37) and (38) only if the prey is successfully acquired.

2.3. Parameter setting

The dimensionality of the data was set to D=2 because we are simulating foraging behavior in a two-dimensional space. The number of hypotheses was set as K=15. For the initial values of the model parameters for each hypothesis, the mean μ_k^1 was determined randomly in the range [-5,5] for both the x and y coordinates. The variance was set as $\Sigma_k^1 = \begin{pmatrix} \Sigma_{xx} & 0 \\ 0 & \Sigma_{yy} \end{pmatrix}$, $\Sigma_{xx} = \Sigma_{yy} = 0.2 \times rnd_k$, where rnd_k represents a uniform random number in the [0,1] interval. The initial value of confidence (that is, prior probability) for each hypothesis was set as $C^1(h_k) = 1/K$, such that they would be equal. For all agents,

the discount rate was set as $\beta = 0.03$. We set the target two-dimensional space in the range [-5,5] for both x and y coordinates. If a larger range is set, a larger number of hypotheses K and the variance Σ_k^1 need to be set in order to cover the entire space. We also set the discount factor $\beta = 0.03$. If the food consumption rate Δ becomes large, the prey density distribution will fluctuate greatly, and thus, it is necessary to set a larger discount for efficient foraging behavior.

For prey density distributions, five cases were simulated: from scarce prey to abundant prey. First, the central coordinates μ_f^p of prey distribution were randomly determined in the range [-5,5] for both the x and y coordinates. The variances of the prey distribution were set as $\Sigma_f^p = \begin{pmatrix} 0.05 & 0 \\ 0 & 0.05 \end{pmatrix}$, $\begin{pmatrix} 0.25 & 0 \\ 0 & 0.25 \end{pmatrix}$, $\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$, $\begin{pmatrix} 1.0 & 0 \\ 0 & 1.0 \end{pmatrix}$, $\begin{pmatrix} 5000 & 0 \\ 0 & 5000 \end{pmatrix}$. A distribution with $\Sigma_f^p = \begin{pmatrix} 5000 & 0 \\ 0 & 5000 \end{pmatrix}$ represents an approximate uniform distribution. w_p^t is a parameter that controls the total amount of prey of patch p, at time t. Here, $w_p^1 = 0.5$ was set as the initial value. The number of patches was set to 10. The food consumption rate Δ was set to 0.0005. The simulation was conducted up to the 30000th step.

The simulation program was developed based on the C++ language. The compiler was MinGW 8.1.0 64-bit for C++ [38]. The Qt library (Qt version Qt 5.15.2 MinGW 64-bit) was also used for development [39].

2.4. Autocorrelation coefficient and correlation time

In this study, autocorrelation coefficient and correlation time are used to measure the time dependence of step-length time series. For a given T time-series data $(l_1, l_2, \cdots l_T)$, the autocorrelation coefficient r(τ) with time lag τ is expressed as follows:

$$r(\tau) = \frac{\sum_{t=1}^{T-\tau}(l_t - \bar{l}_{1,T-\tau})(l_{t+\tau} - \bar{l}_{1+\tau,T})}{\sqrt{\sum_{t=1}^{T-\tau}(l_t - \bar{l}_{1,T-\tau})^2}\sqrt{\sum_{t=1}^{T-\tau}(l_{t+\tau} - \bar{l}_{1+\tau,T})^2}}$$
$$\bar{l}_{1,T-\tau} = \frac{\sum_{t=1}^{T-\tau} l_t}{T-\tau} \qquad (39)$$
$$\bar{l}_{1+\tau,T} = \frac{\sum_{t=1}^{T-\tau} l_{t+\tau}}{T-\tau}$$

In this study, the autocorrelation coefficient was obtained in the range τ = 1–2500. The autocorrelation coefficients for every time-series data are averaged from $r(1)$ to $r(2500)$ and are expressed as $\bar{r}$.

$$\bar{r} = \frac{\sum_{\tau=1}^{2500} r(\tau)}{2500} \qquad (40)$$

Thereafter, $\bar{r}$ is referred to as correlation time.

2.5. Statistical analyses

The following analyses were performed using the statistical software R, version 3.6.1 (2019-07-05) [40], unless otherwise specified. We used the R packages of RVAideMemoire version 0.9.81 for the pairwise comparisons using Fisher's exact test and multcomp version 1.4.16 for the Tukey–Kramer test. The operating system used was Windows 10 (Microsoft Corp., Redmond, WA, USA). For all analyses statistical significance was set at $p < 0.05$.

3. Results

We performed foraging simulations for five cases from scarce prey to abundant prey. The agent probabilistically decided where to forage according to its own estimated model that was represented by a normal distribution in this study. Prey were acquired stochastically in proportion to the prey density at the foraging point. The objective of the agent was to acquire as many prey as possible within a certain period.

Figure 2 shows an example of the foraging behavior of an agent in the case of $w_p^t = 0.5, \Sigma_f^p = 0.25$. Figure 2(a) plots the foraging points at each time point. The z axis and color represent time steps. In the figure, two points d^{t-1} and d^t at adjacent times are connected by straight lines. Figures 2(b) and 2(c) show examples of the prey distribution at the initial stage (t = 1) and t = 3000, respectively. It can be seen that some patches are on the verge of disappearing due to the consumption of food. Figure 2(d) shows the time series of patch numbers explored by the agent. Figure 2(e) shows the time evolution of the expected

reward, expressed in terms of prey density $\Phi_{\hat{p}}^{t}\left(d^{t}|\mu_{f}^{\hat{p}},\Sigma_{f}^{\hat{p}}\right)$ at the foraging point d^{t}.

As can be seen in these figures, when an agent finds a patch, it forages intensively in that patch. When the food is consumed and the amount of food in the patch becomes low, the agent repeats the process of searching and moving to a new patch.

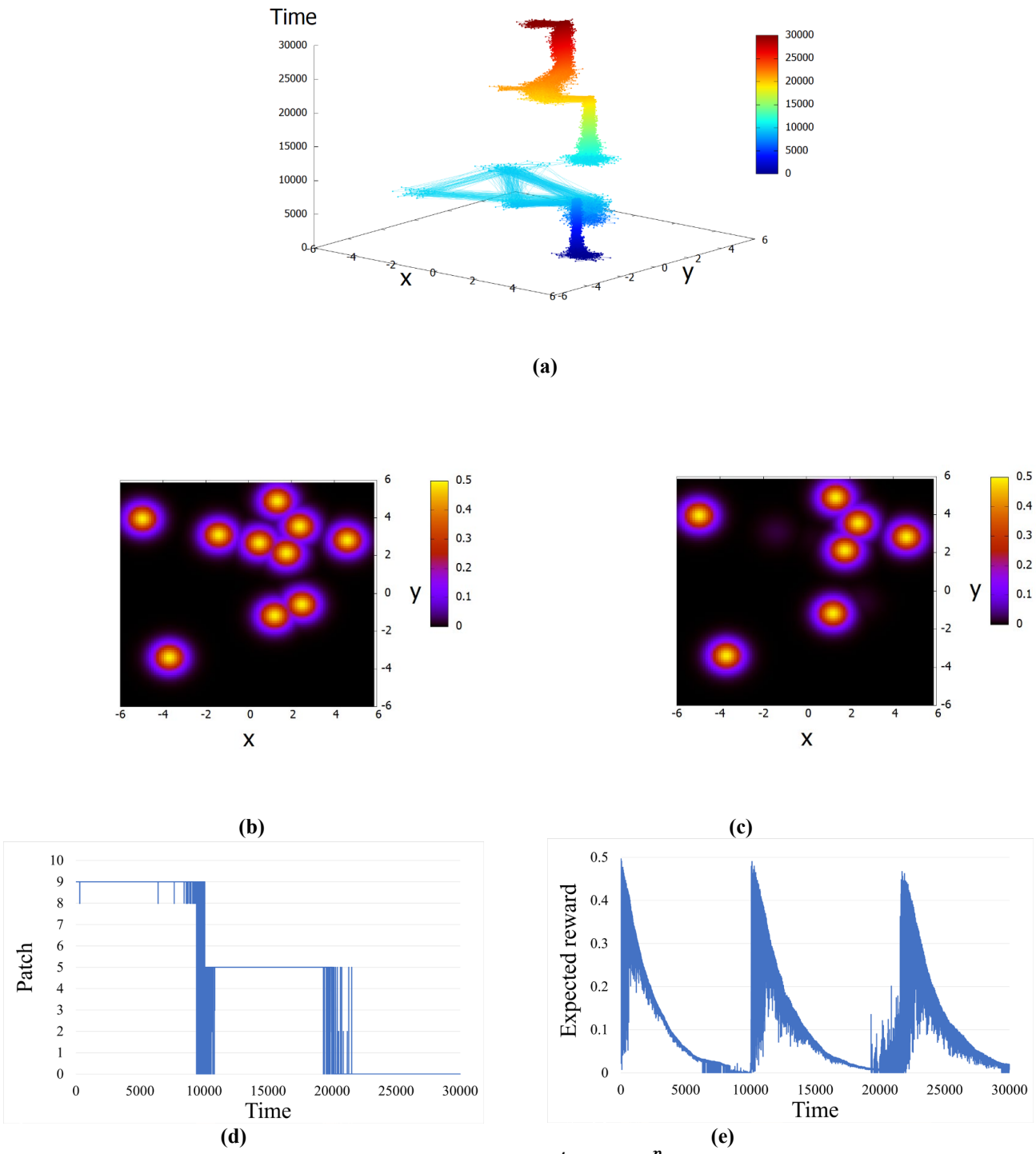
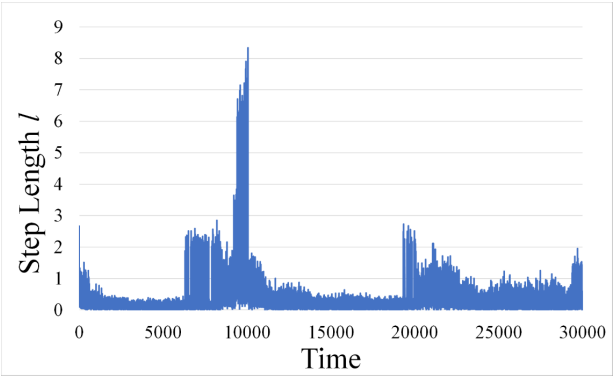


Figure 2 Example of foraging behavior in the case of $w_p^t = 0.5, \Sigma_f^p = 0.25$. (a) Movement pattern from time 1 to 30000. Two foraging points at adjacent time points are connected by a straight line. Z axis and color represent time steps. (b) Prey density distribution at the initial stage (t = 1) (c) Prey density distribution at t = 30000. (d) Time evolution of patch numbers explored by an agent. (e) Time evolution of the expected reward.

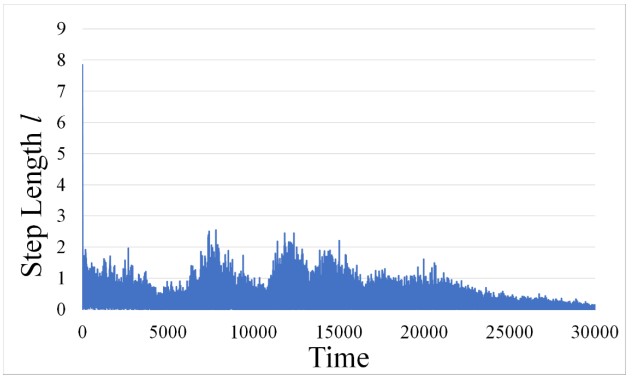
373

Figures 3(a) and 3(b) show examples of the time evolution of the step length $l_t = \|d_t - d_{t-1}\|_2$. Figures 3(c) and 3(d) show the complementary cumulative distribution function (CCDF) of the step length l_t when Figures 3(a) and 3(b), respectively. These figures also show the results of fitting a truncated power-law distribution (TP) model (green) and an EP model (red) to the simulation data. Of the step length data, only the data within the range $\hat{l}_{min} \leq l_t \leq \hat{l}_{max}$ are used to improve the fitting as much as possible. See the Supplementary Information (SI) for details on how the fitting range $[\hat{l}_{min}, \hat{l}_{max}]$ and fitting method are determined. The CCDF graph is such that the CCDF value when $l = \hat{l}_{min}$, is set to one. Figures 3(b) and 3(d) are double logarithmic graphs with both axes displayed in logarithmic form.

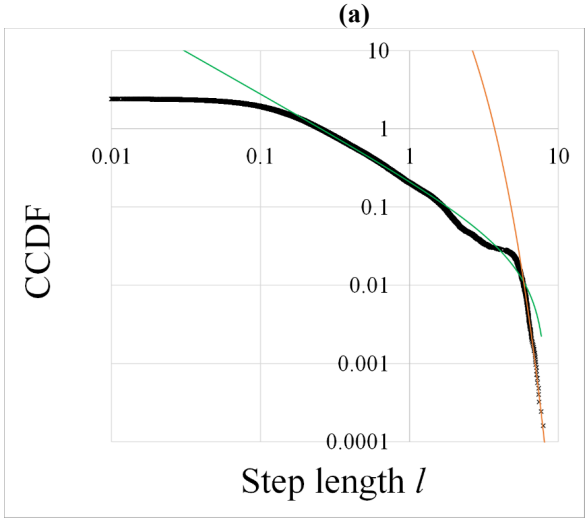
As seen from these figures, a Lévy walk-like gait pattern, wherein the frequency distribution of the step length shows a power distribution, and a Brownian walk-like gait pattern, which shows an exponential distribution, appear. Figures 3(e) and 3(f) show the correlogram of the step-length time series shown in Figures 3(a) and 3(b), respectively. The horizontal and vertical axes represent the time lag τ and the autocorrelation coefficient for each τ, respectively. From these figures, we can see that the temporal autocorrelation is stronger for Lévy walk-like gait pattern.


Step Length l
Time

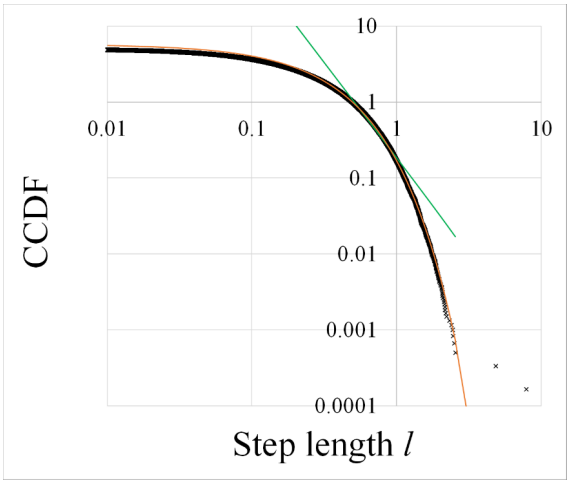
(a)


Step Length l
Time

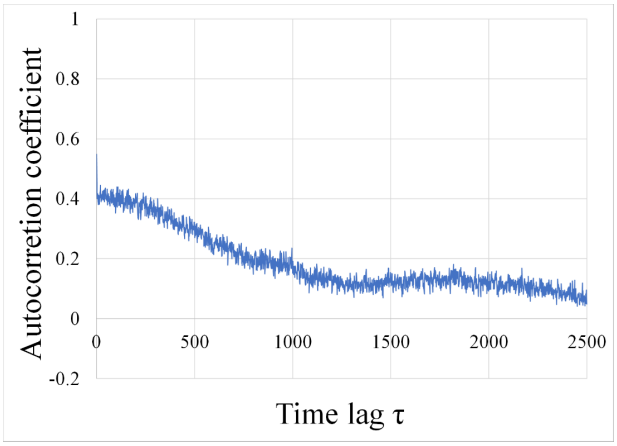
(b)


CCDF
Step length l

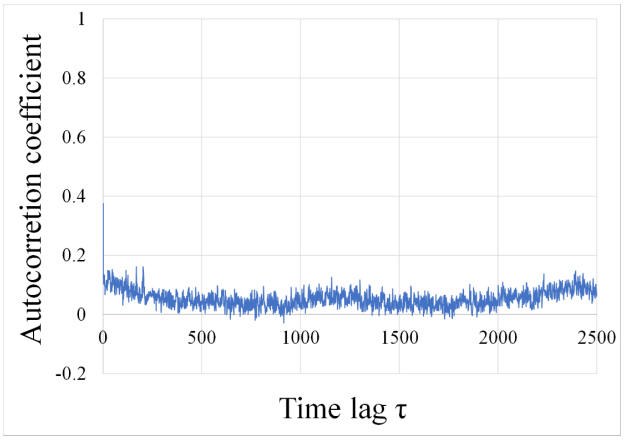
(c)


CCDF
Step length l

(d)


Autocorretion coefficient
Time lag τ

(e)


Autocorretion coefficient
Time lag τ

(f)

388

Figure 3. Examples of time evolution of step length, Complementary cumulative distribution function (CCDF) of step length, and autocorrelation of step length time series. (a) Time evolution of step length in the case where a Lévy walk-like gait pattern is observed. (b) Time evolution of step length in the case where a Brownian walk-like gait pattern is observed. (c) CCDF of the step length in the case of (a). This figure also shows the results of fitting a TP model (green) and an EP model (red) to the simulation data. Exponent of the TP is η = 2.08. Fitting range is [0.25, 8.34]. (d) CCDF of the step length in the case of (b). This figure also shows the results of fitting a TP model (green) and an EP model (red) to the simulation data. Exponent of the EP is λ = 3.53. Fitting range is [050, 7.84]. (e) Correlogram for the step-length time series shown in (a). (f) Correlogram for the step-length time series shown in (b).

Hereafter, we will refer to the group of agents whose frequency distribution of step length is classified as EP as the EP group, and the group of agents whose frequency distribution of step length is classified as TP as the TP group. Furthermore, we will divide the TP group into those whose exponent η is in the range $1.0 \leq \eta \leq 3.0$ (TP1) like the Levy walk and those whose exponent η is not in the range of $1.0 \leq \eta \leq 3.0$ (TP2). Table 1 shows the number of times the frequency distribution of the step length has been classified as TP1, TP2, and EP, as the results of 100 simulation trials with different random seeds for each of the cases of the five prey density distributions (Σ_f^p =0.05, 0.25, 0.5, 1.0, 5000.0).

Fisher's exact test with Bonferroni correction was performed for the ratio of TP1 to EP. The results showed that there were significant differences in the ratio at 0.1% level for all combinations except $\Sigma_f^p = 0.05$–0.25 and $\Sigma_f^p = 1.0$–5000.0. In other words, the scarcity of food tends to increase the percentage of TP1 compared to EP.

Table 1. Number of trials for which the frequency distribution of the step lengths was classified as TP1, TP2, and EP.

Group	Variance of prey density distribution				
	$\Sigma_f^p = 0.05$	$\Sigma_f^p = 0.25$	$\Sigma_f^p = 0.5$	$\Sigma_f^p = 1.0$	$\Sigma_f^p = 5000.0$
TP1	62	62	30	4	7
TP2	12	14	15	9	2
EP	26	24	55	87	91
Total	100	100	100	100	100

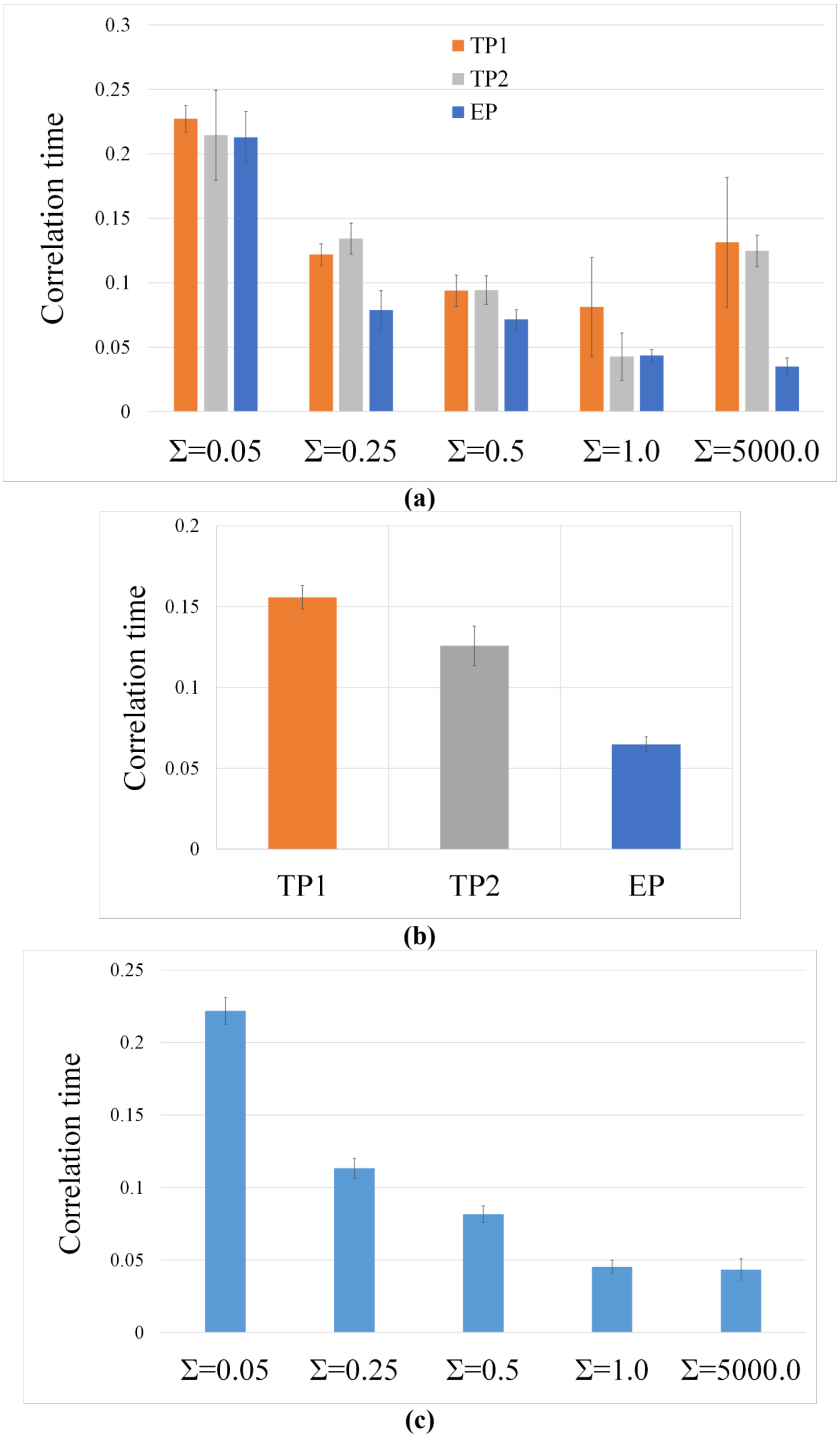


Figure 4. Comparison of correlation time. (a) Average correlation time for each group shown in Table 1 for each prey density distribution. (b) Average correlation time for each group. (c) Average correlation time for each prey density distribution. In all figures, the error bars represent the standard error.

Figures 4 shows the time dependence of the step length. Figure 4(a) shows the mean values of correlation time for each group shown in Table 1 for each prey density distribution. Two-way ANOVA of correlation time was conducted for two factors, groups (i.e., EP, TP1, or TP2) and variances of prey density distribution (i.e., Σ_f^p =0.05, 0.25, 0.5, 1.0, 5000.0). Significant differences were noted in both group and variance factors ($F(2, 485) = 8.83$, $p < 0.001$, $F(4,485) = 73.16$, $p < 0.001$). However, there was no interaction between them ($F(8, 485) = 1.50$, $p = 0.154$).

Figure 4(b) shows the average correlation time for each group. The Tukey–Kramer test revealed significant differences between the EP and the TP1 groups ($p < 0.001$) and between the EP and the TP2 groups ($p < 0.001$). However, there was no significant difference between TP1 and TP2 groups ($p = 0.070$). The correlation time of TP1 is larger than that of EP, i.e., there is a stronger time dependence in the time variation of the step length for the agents belonged in TP1 group.

Figure 4(c) shows the average correlation time for each prey density distribution. The results of the Tukey–Kramer test showed that there are significant differences in correlation time among all variances except between Σ_f^p=1.0 and 5000.0 at the level of 5%. Thus, the more abundant the food, the smaller the correlation time. Conversely, the sparser the food, the stronger the time dependence of the step length.

4. Discussion

In this study, a novel model of BIB inference was proposed wherein knowledge learning and knowledge-based inference are performed simultaneously. The learning was modeled using SDEM, an online EM algorithm, and the inference was modeled using BID, a Bayesian inference with discount factor. Next, foraging simulations were conducted using a decision-making agent incorporating the BIB inference model. In BIB agents, past experiences are stored as hypotheses (knowledge) and are used in the current foraging

behavior; at the same time, the hypothesis continues to be updated based on new experiences.

The following are the three results of the simulations. First, when prey are sparse, the frequency distribution of the step length (that is, the distance between the foraging point at one time and the foraging point the next time) tends to follow a TP, whereas when prey are dense, the distribution tends to become an EP. This result is consistent with the finding that the frequency of occurrence of the step length in some marine organisms switch between power-law and exponential distributions depending on the amount of prey [4,12]. Second, the time variation of the step length tends to show a strong time dependence when the frequency distribution of the step length follows a power-law distribution. This result is consistent with the finding [17] that there is a relationship between the dependence of the step length on time and the fact that the frequency distribution of the step length follows a power-law distribution. Third, the more sparsely distributed the prey, the stronger the time dependence of the step length.

When the frequency distribution of the step length follows a TP, the search behavior of BIB agent is divided into two modes: a rough search to select the patch and a detailed search to detect the center of the selected patch. This switch may be considered a model of ARS. The agents show extremely long steps in the patch search mode. However, after fixing the patch, gradually, the agent will choose a location near the center of the patch as its foraging point in the detailed search mode. Accordingly, the step length approaches zero. Because the step length tends to decrease gradually, there is a stronger time dependence of the step length in the TP group. It is interesting to note here that although the foraging points are sampled from a normal distribution as shown in Equation (2), the frequency distribution of the distances between foraging points, that is, the step lengths, results in a TP.

In the case of prey abundance, the frequency distribution of the step length tends to become an EP. In the case of prey abundance, it is not necessary to cause the center of the estimated model to coincide with the center of the prey distribution and to converge the variance of the estimated model to zero, because prey can be obtained with a high probability without foraging at the center of the patch. Because the foraging point does not gradually move closer to the patch center, the time dependence is expected to be weaker with respect

to the temporal variation of the step length. Thus, it can be noted that depending on the difficulty of prey acquisition, various migratory behavioral patterns emerge.

In general, the EM algorithm or Bayesian inference attempts to accurately estimate the parameters (mean and variance) of the generating distribution that produces the observed data. It is important to note that the observed data are usually sampled from the generating distribution; therefore, regardless of how the EM algorithm or Bayesian inference estimates the generating distribution, it has no effect on the observed data. However, in the simulation task of this study, foraging points are determined based on the model estimated by the agent; this implies that the observed data depend on the distribution that the agent estimates. This can lead to the variance of the model estimated by the BIB agent approaching zero, even though the variance of the actual prey distribution is not zero.

The exploration and preferential return (EPR) model incorporates the random walk process of exploration and the human propensity to revisit places (preferential return). In EPR, one of these two competing mechanisms is chosen probabilistically each time [41]. Exploration and exploitation are two essential components of any optimization algorithm [42]. Determining an appropriate balance between exploration and exploitation is the most challenging task in the development of any metaheuristic algorithm, owing to the randomness of the optimization process [43]. In the present simulations, two patterns emerged: a rough search using knowledge (hypotheses) and a detailed search where the target was narrowed down by modifying the knowledge. The former may be associated with exploitation and the latter with exploration. It is interesting to note that the emergence of these behavioral patterns in BIB agents is not explicitly controlled by any parameters. As can be seen from the simulation procedure, the BIB agent performs both learning and inference simultaneously at every step.

The limitation of this study is that the simulations addressed only the case of a single cluster with respect to the mixture distribution of the SDEM. In future work, we intend to perform simulations using a generative distribution comprising several normal distributions. In the present study, we regarded the generation distribution and density distribution of prey in each patch as a normal distribution. In future work, we intend

to simulate a case wherein the density distribution of the prey inside the patch is uniform or nonparametric.

Funding

This work was supported by the Center of Innovation Program from the Japan Science and Technology Agency, JST; and JSPS KAKENHI [grant number JP21K12009].

Competing interests

N.M. is employed by SoftBank Robotics Group Corp. This does not alter our adherence to *Chaos, Solitons & Fractals* policies on sharing data and materials. All other authors declare no conflicts of interest.

Data availability statement

The source code and data used to produce the results and analyses presented in this manuscript are available from Bitbucket Git repository:
https://bitbucket.org/ShujiShinohara/levyflightforagingsimulation1/src/master/

References

[1] Harris TH, Banigan EJ, Christian DA, Konradt C, Wojno EDT, Norose K, et al. Generalized Lévy walks and the role of chemokines in the migration of effector CD8+ T cells. Nature. 2012;486:545–8. DOI: 10.1038/nature11098.

[2] Ariel G, Rabani A, Benisty S, Partridge JD, Harshey RM, Be'Er A. Swarming bacteria migrate by Lévy walk. Nat Commun. 2015;6:8396. DOI: 10.1038/ncomms9396.

[3] Shokaku T, Moriyama T, Murakami H, Shinohara S, Manome N, Morioka, K. Development of an automatic turntable-type multiple T-maze device and observation of pill bug behavior. Rev Sci Instrum. 2020;91: 104104. DOI: 10.1063/5.0009531.

[4] Humphries NE, Queiroz N, Dyer JRM, Pade NG, Musyl MK, Schaefer KM, et al. The environmental context explains the Lévy and Brownian movement patterns of marine predators. Nature. 2010;465:1066–9. DOI: 10.1038/nature09116.

[5] Humphries NE, Weimerskirch H, Queiroz N, Southall EJ, Sims DW. Foraging success of biological Lévy flights recorded in situ. Proc Natl Acad Sci U S A. 2012;109:7169–74. DOI: 10.1073/pnas.1121201109.

[6] Raichlen DA, Wood BM, Gordon AD, Mabulla AZP, Marlowe FW, Pontzer H. Evidence of Lévy walk foraging patterns in human hunter–gatherers. Proc Natl Acad Sci U S A. 2014;111:728–33. DOI: 10.1073/pnas.1318616111.

[7] Sims DW, Southall EJ, Humphries NE, Hays GC, Bradshaw CJA, Pitchford JW, et al. Scaling laws of marine predator search behavior Nature. 2008;451:1098–102. DOI: 10.1038/nature06518.

[8] Viswanathan GM, Raposo EP, da Luz MGE. Lévy flights and superdiffusion in the context of biological encounters and random searches Phys Life Rev. 2008;5:133–50. DOI: 10.1016/j.plrev.2008.03.002.

[9] Bartumeus F, da Luz MGE, Viswanathan GM, Catalan J. Animal search strategies: A quantitative random-walk analysis. Ecology. 2005;86:3078–87. DOI: 10.1890/04-1806.

[10] Viswanathan GM, Buldyrev SV, Havlin S, Da Luz MGE, Raposo EP, Stanley HE. Optimization of the success of random searches. Nature, 1999;401:911–4. DOI: 10.1038/44831.

[11] Bartumeus F, Catalan J, Fulco UL, Lyra ML, Viswanathan GM. Optimizing the encounter rate in biological interactions: Lévy versus Brownian strategies. Phys. Rev. Lett. 2002;88(9):097901. DOI: 10.1103/PhysRevLett.88.097901.

[12] Sims DW, Humphries NE, Bradford RW, Bruce BD. Lévy flight and Brownian search patterns of free-ranging predators reflect different prey field characteristics. J Anim Ecol. 2012;81(2):432–42. doi: 10.1111/j.1365-2656.2011.01914.x. Epub 2011 Oct 17.

[13] Ross C, Pacheco-Cobos L, Winterhalder B. A general model of forager search: Adaptive encounter-conditional heuristics outperform Lévy flights in the search for patchily distributed prey. J Theor Biol. 2018;455: 357–69. doi: 10.1016/j.jtbi.2018.07.031.

[14] Namboodiri VM, Levy JM, Mihalas S, Sims DW, Shuler MG. Rationalizing spatial exploration patterns of wild animals and humans through a temporal discounting framework. Proc Natl Acad Sci U S A. 2016;113:8747–52. DOI: 10.1073/pnas.1601664113.

[15] Ross CT, Winterhalder B. Evidence for encounter-conditional, area-restricted search in a preliminary study of Colombian blowgun hunters. PLOS ONE. 2018;13:e0207633. DOI: 10.1371/journal.pone.0207633.

[16] Hills TT, Kalff C, Wiener JM. Adaptive Lévy processes and area-restricted search in human foraging. PLOS ONE. 2013;8(4):e60488. https://doi.org/10.1371/journal.pone.0060488.

[17] Wang XW, Han XP, Wang BH. Correlations and scaling laws for human mobility. PLOS ONE. 2014;9:e84954. DOI: 10.1371/journal.pone.0084954.

[18] Rhee I, Shin M, Hong S, Lee K, Kim SJ, Chong S. Levy-walk nature of human mobility. IEEE/ACM Trans Networking 2011;19:630–43. DOI: 10.1109/TNET.2011.2120618.

[19] Zhao K, Musolesi M, Hui P, Rao W, Tarkoma S. Explaining the power-law distribution of human mobility through transportationmodality decomposition. Sci Rep. 2015;5:9136. https://doi.org/10.1038/srep09136.

[20] Boyer D, Ramos-Fernández G, Miramontes O, Mateos JL, Cocho G, Larralde H, et al. Scale-free foraging by primates emerges from their interactions with complex environments. Proc Biol Sci. 2006;273:1743–50. DOI: 10.1098/rspb.2005.3462.

[21] Gunji YP, Shinohara S, Haruna T, Basios V. Inverse Bayesian inference as a key of consciousness featuring a macroscopic quantum logical structure. Biosystems. 2017;152:44–65. DOI: 10.1016/j.biosystems.2016.12.003.

[22] Gunji YP, Murakami H, Tomaru T, Basios V. Inverse Bayesian inference in swarming behaviour of soldier crabs. Philos Trans A Math Phys Eng Sci. 2018;376. DOI: 10.1098/rsta.2017.0370.

[23] Gunji YP, Kawai T, Murakami H, Tomaru T, Minoura M, Shinohara S. Lévy walk in swarm models based on Bayesian and inverse Bayesian inference. Comput Struct Biotechnol J. 2021;19:247–60. DOI: 10.1016/j.csbj.2020.11.045.

[24] Zaki SR, Nosofsky RM. A high-distortion enhancement effect in the prototype learning paradigm: Dramatic effects of category learning during the test. Mem Cognit. 2007;35:2088–96. DOI: 10.3758/bf03192940.

[25] Zhu X, Goldberg AB. Introduction to semi-supervised learning. San Rafael, CA, USA: Morgan Claypool; 2009.

[26] Shinohara S, Manome N, Suzuki K, Chung UI, Takahashi T, Gunji PY, et al. Extended Bayesian inference incorporating symmetry bias. Biosystems. 2020;190:104104. DOI: 10.1016/j.biosystems.2020.104104.

[27] Shinohara S, Manome N, Suzuki K, Chung UI, Takahashi T, Okamoto H, et al. A new Bayesian causal inference method in nonstationary environments. PLOS ONE. 2020;15:e0233559. DOI: 10.1371/journal.pone.0233559.

[28] Bishop CM. Pattern recognition and machine learning. New York: Springer; 2011.

[29] Yamanishi K, Takeuchi J, Williams G, Milne P. Online unsupervised outlier detection using finite mixtures with discounting learning algorithms. Data Min Knowl Discov. 2004;8:275–300. https://doi.org/10.1023/B:DAMI.0000023676.72185.7c.

[30] Neal RM, Hinton GE. The EM algorithm justifies incremental, sparse, and other variants. In: Jordan MI, editor. Learning in graphical models. Dordrecht: Kluwer Academic Publishers; 1998, p. 355–68. https://doi.org/10.1007/978-94-011-5014-9_12.

[31] Sato M, Ishii S. Online EM algorithm for the normalized Gaussian network. Neural Comput. 2000;12:407–32. DOI: 10.1162/089976600300015853.

[32] Liang P, Klein D. Online EM for unsupervised models. Proceedings of Human Language Technologies 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2009;611–9.

[33] Cappe O, Moulines E. Online expectation–maximization algorithm for latent data models. J R Stat Soc B. 2009;71:593–613. https://doi.org/10.1111/j.1467-9868.2009.00698.x.

[34] Dias JG, Cortinhal MJ. SKM algorithm: k-means algorithm for clustering sequential data. In: Geffner H, Prada R, Machado AI, David N, editors. Lect Notes Comput Sci, New York: Springer; 2008, p. 173–82. https://doi.org/10.1007/978-3-540-88309-8_18.

[35] Zhong S. Efficient online spherical k-means clustering. Proceedings of the IEEE International Joint Conference on Neural Networks 2005;5:3180–5. https://doi.org/10.1109/ijcnn.2005.1556436.

[36] Andrews NO, Fox EA. *Recent developments in document clustering*. Technical Report TR-07. Available at: https://vtechworks.lib.vt.edu/handle/10919/19473-35; 2007.

[37] Tehrani AF, Ahrens D. Modified sequential k-means clustering by utilizing response: A case study for fashion products. Expert Syst. 2017;34:e12226. https://doi.org/10.1111/exsy.12226.

[38] Mingw. http://www.mingw.org/ [Accessed 23 January 2021].

[39] Qt. https://www.qt.io/ [Accessed 23 January 2021].

[40] A Language and Environment for Statistical Computing. Available online: https://www.R-project.org/ [Accessed 5 July 2019].

[41] Pappalardo L, Rinzivillo S, Simini F. Human mobility modeling: Exploration and preferential return meet the gravity model. Procedia Comput Sci. 2016;83:934–9. DOI: 10.1016/j.procs.2016.04.188.

[42] Emary E, Zawbaa HM. Impact of chaos functions on modern swarm optimizers. PLOS ONE. 2016;11:e0158738. DOI: 10.1371/journal.pone.0158738.

[43] Wu X, Zhang S, Xiao W, Yin Y. The exploration/exploitation tradeoff in whale optimization algorithm. IEEE Access. 2019;7: 125919–28. DOI: 10.1109/ACCESS.2019.2938857, 125919.