Chapter 6
Learning Policy for Robot Anomaly Recovery Based on Robot Introspection

Abstract In this chapter, the anomaly recovery would be acted when both of the anomaly monitoring and diagnoses are analysed, which aim to respond the external disturbances from the environmental changes or human intervention in the increasingly human-robot scenarios. To effectively evaluate the exploration, we summarize the anomalies in a robot system include only two catalogues: accidental anomalies and persistent anomalies. In particular, we first diagnose the anomaly as accidental one at the beginning such that the reverse execution is called. If and only if robot reverse many times (not less than twice) and still couldn’t avoid or eliminate the anomaly, the human interaction is called. Our proposed system would synchronously record the multimodal sensory signals during the process of human-assisted demonstration. That is, a new movement primitive is learned once an exploring demonstration acquired. Then, we heuristically generate a set of synthetic demonstrations for augmenting the learning by appending a multivariate Gaussian noise distribution with mean equal to zeros and covariance equal to ones. Such that the corresponding introspective capabilities are learned and updated when another human demonstration is acquired. Consequently, incrementally learning the introspective movement primitives with few human corrective demonstrations when an unseen anomaly occurs. It is essential that, although there are only two different exploring strategies when anomaly occurs, numerous exploring behaviors can be generated according to different anomaly types and movement behaviors under various circumstances.

6.1 Introduction

In recent years, with the widespread promotion of the cooperative robot and human beings’ inclusive operation, the robot’s abnormal recovery behavior cannot be performed by the traditional robot’s own motion planning algorithm, and human expectations for robot motion should be brought into play. The recovery strategy will reflect the human-centered concept of human-machine collaboration. For this reason, the robot anomaly recovery problems considered in this article mainly refer to the recoveries learned from humans by themselves in the case of robots that can recovery anomalies (not considering abnormal situations that cannot be recovered.
autonomously such as power interruptions or system crashes). The strategy makes the robot’s motion change from an abnormal state to a normal state. In other words, how to learn from humans to obtain corresponding recovery strategies for different abnormal types of robots is the key issue of this chapter, and it is required that the proposed recovery strategies have certain expansibility and generalization.

### 6.2 Related Work

To address the unpredictable abnormalities in robot operation tasks in an unstructured dynamic environment, Gutierrez et al. [1] proposed a way of expanding the original operation graph to recovery unforeseen abnormal events. The overall idea of the method is to use finite states machine learns an initial robot operation diagram for limited human demonstration movements, then re-plans the robot’s movement when there is an abnormality in the actual operation application, and adds this recovered movement to the operation diagram, as the operation diagram continues perfect, so that we can always find a feasible recovery behavior for the abnormal event. Paster et al. [2, 3] proposed the establishment of a specific library of motion primitives for the robot’s operation tasks, and taught the different ways to complete the task through human experience. After combining the methods of motion primitive selection, the abnormal premise was detected Next, the next primitive corresponding from the motion primitive library is completed [4, 5], thereby completing the behavior of abnormal recovery. Different from the robot’s own method of motion conversion and selection, Salazar-Gomez et al. [6] proposed to implement robot’s abnormal recovery by introducing human observation and control [7]. This method is a kind of human-in-the-loop (Human-in-the-loop) human-robot interaction. This has the advantage that humans have a more intuitive understanding of the robot’s movements. When abnormal or error occurs, they can think about feasible solutions in time. Such recovery behavior is more Vulnerable to researchers and more conducive to applications in the human-machine environment. Niekum et al. [8] also described the operation tasks of the robot through the finite state machine. In the case of detecting an abnormal event, humans assist the robot to complete the current sub-task through kinematic teaching.

With associating this recovery behavior with the type of anomaly and update it to the original operation diagram, so that the robot can adopt a specific recovery behavior in the face of different anomalies. Mao et al. [9] proposed a method of human-assisted learning to implement robot abnormal recovery, where the robot’s movement from the initial motion primitive will collide with the environment. At this time, the human manual Stop the robot and start the teach and teach mode to re-teach the recovered motion primitives for the robot to complete the abnormal recovery. Karlsson et al. [10, 11] proposed an online method of dynamically modifying robot motion primitives to achieve anomaly recovery. This method also pre-parameterized the robot’s motion by DMP. In the event of an abnormality, it is indicated by human motion Teach ways to learn recovery behavior. In [12], the authors proposed the abort and retry
behavior in grasping by minimizing the overhead time as soon as the task is likely to fail. As an alternative, Johan S. Laursen et al. introduced a system for automatically reversed in robot assembly operations using reverse execution, from which forward execution can be resumed [13]. Similarly, a recovery policy by modelling reversible execution of robotic assembly is proposed in [14]. Another perspectives for anomaly recovery, Arne Muxfeldt et al. proposed a novel approach for recovering from errors during automated assembly in typical mating operations [15–17], which is based on automated error detection with respect to a predefined process model such that a recovery strategy can be selected from an optimized repository.

In summary, the use of human-machine collaboration to demonstrate abnormal recovery behavior has been favored by a large number of researchers, because in daily life or human-machine collaboration environment, the abnormal recovery behavior of robots cannot be achieved by the traditional robot itself. The motion planning algorithm should be carried out, and human expectations for robot motion should be brought into play. The human-assisted robot abnormality recovery strategy will further reflect the “human-centered” human-machine collaboration concept. In view of this, based on anomaly monitoring and classification, this paper proposes two types of recovery strategies for accidental and persistent anomalies, and different strategies will have different recovery behaviors under different types of anomalies.

### 6.3 Statement of Robot Anomaly Recovery

Specifically, the task segmentation, anomaly monitoring and diagnosis can be implemented through the generative methods such as the Bayesian non-parametric time series model that described in Chaps. 3–5 respectively. With the help of the task generalization and introspective capabilities for each movement primitives, two task exploration policies are proposed for responding the anomalies: reverse execution or human interaction. These policies can be learned from the context of manipulation tasks in an increasing manner. An overview of our SPAI framework can be described by a graph $G$ composed of $N_b$ behaviors (or sub-tasks) $B$, which are interconnected by edges $E$ such that $G : (B, E)$, $B = \{B_1, ..., B_{N_b}\}$, $E_{s, t} = \{(s, t) : s, t \in B\}$. Behaviors in turn consist of nodes $N$ and edges $E$, such that: $B = \{N, E\}$. Nodes can be understood as phases of a manipulation task. In our work, we prefer to name them milestones $N_i = (1, ..., N_i)$, as they indicate particularly important achievements in a task. With task exploration behaviors, we may introduce a exploring nodes $N_{ij}$ in-between milestones creating a new branch to the subsequent milestone. It is also possible to introduce further exploring nodes $N_{ijk}$ on already existing branches. The full set of nodes in a task then is described as the union of milestone nodes with all branched nodes $\mathcal{N} = \{N_i \cup N_{ij} \cup \ldots \cup N_{ij...q}\}$. Node Transitions $T$, behave as with behavior transitions $E$, so $T_{s, t} = \{(s, t) : s, t \in \mathcal{N}\}$. In our framework, a node is composed of modules. Modules are dependent processes that play a key role in the execution of a manipulation phase and could include demonstration collection, segmentation, movement learning, introspection, task representation and
exploration, vision, natural language processing, navigation, as well as higher level agents. In this chapter, we restrict a node to segmentation, generalization, introspection and exploration modules.

6.4 Learning and Application of Anomaly Recovery Policy

6.4.1 Learning for Unstructured Demonstrations

As we know, a sentence in language is made up of words according to grammatical rules, and a word is made up of letters according to word formation. Correspondingly, a complex and multi-step robot manipulation task can be represented as a sentence, in which a coupled robot movement primitive is equivalent to a word, and the robot movement primitive of each degree of freedom (DoF) of can be considered as letter such that the robot manipulation task can be represented with a set of movement primitives. To this end, how to effectively learn a multi-functional movement primitive is a critical problem for intelligent robot performing complex tasks in unstructured environment. If so, that is an interesting idea for generalizing the task representation with respect to the external adjustment as well as improving the diversity and adaptability of tasks. As a consequence, the task is consist of sequential movement primitives, which not only take the kinesthetic variables into consideration but also equip with introspective capacities such as the identification of movement, anomaly detection and anomaly diagnoses. We now introduce a Bayesian nonparametric model for robust learning the underlying dynamics from unstructured demonstrations.

• Demonstrations

Generally, capturing the demonstrations by receiving the multimodal input from the end-user, such as a kinesthetic demonstration. The only restriction on the multimodal data is that all the signals must be able to be synchronously recorded at the same frequency, i.e. temporal alignment. Additionally, the multimodal data at each time step should include the Cartesian pose and velocity of the robot end-effector (in case of object-grasping, will along with any other relevant data about the end-effector, e.g. the open or closed status of the gripper and the relative distance between the end-effector and object.) as well as the signals from F/T sensor and tactile sensor. Subsequently, the recorded Cartesian pose and velocity trajectory of the end-effector will be referred to as the kinematic demonstration trajectory for controlling the robot motion, and the recorded signals of F/T sensor and tactile sensor are applied for learning the introspective capacities.

• Learning from Demonstration

Learning from demonstration (LfD) has been extensively used to program the robots, which aiming to provide a natural way to transfer human skills to robot. LfD is
6.4 Learning and Application of Anomaly Recovery Policy

proposed by simply teaching a robot how to perform a task as human-like, in which users can demonstrate new tasks as needed without any prior knowledge about the robot. However, LfD often yields weak interpretation about the environment as well as the task always is single step and lab-level such that lacks of robust generalization capabilities in dynamic scenarios, especially for those complex, multi-step tasks, e.g. human-robot collaborative kitting task designed in this paper.

For this reason, we present the powerful algorithms that draw from recent advances in Bayesian nonparametric HMMs for automatically segment and leverage repeated dynamics at multiple levels of abstraction in unstructured demonstrations. The discovery of repeated dynamics provides discriminating insights for understanding the task invariants, high-level description from scratch, and appropriate features for the task. In this paper, these discoveries could be concatenated using a finite state representation of the task, and consisted of movement primitives that are flexible and reusable. Thus, this implementation provides robust generalization and transfers in complex, multi-step robotic tasks. We now introduce a flowchart which integrates three major modules that critical for implementing the complex task representation from unstructured demonstrations.

- **Segmentation**

The aim of the segmentation is exploring the hidden state representation of the demonstrations, in which a specific hidden state usually denotes the clustering observations would be sampled from a statistical model, e.g. multivariate Gaussian distribution. Hidden state space modeling of multivariate time-series is one of the most important tasks in representation learning by dimensional reduction. In this work, we propose the segmentation approach is a hidden state determined with Bayesian nonparametric hidden Markov model, which leads to tackle the generalization problem in a more natural way to meet the need of real-world applications.

As we discussed above, the HDP-VAR-HMM interpreted each observation $y_t$ by assigning to a single hidden state $z_t$, where the hidden state value is derived from a countably infinite set of consecutive integers $z_t = k \in \{1, 2, 3, \ldots, K\}$. We denote $\Theta_s$ to represent all the parameters of the trained HDP-VAR-HMM from nominal demonstrations, including the hyper-parameters for learning the posterior and the parameters for the observation model definition.

$$z = \arg \max_{1, \ldots, K} p(z_t | y_t, \Theta_s). \quad (6.1)$$

Here, the $z$ would be a variable value, that is, we can concatenate the derived hidden state sequences and group the starting and goal observations for each sequence, respectively.

Assume that we record $N$ multimodal demonstrations via kinesthetic fashion and jointly model them using the HDP-VAR-HMM, result in $N$ hidden state sequences $Z = \{Z_1, Z_2, \ldots, Z_N\}$, where the element of $Z_i$, is integer. Here, the problem is to segment the demonstrations into time intervals associated with individual movement.
primitives by state-specific way. After the complex task segmentation, we achieve the task representation as presented in Chap. 3 using Finite State Machine (FSM) and Dynamical Movement Primitives (DMP).

6.5 Reverse Execution Policy for Accidental Anomalies

Reverse execution allows the robot retry the current movement or several movements that independent with the current state for resolving the accidental faults, such as human collision and mis-grasp. The key question is how far back must we revert in the task? To address this, we are currently evaluating different methodologies to learn reverse policies. Ideally, the performing critic is able to include all task-relevant information: the state of the robot, the state of the environment, the affordances of a task, and the relationship between these elements. However, this is not trivial, we are studying whether we could integrate decision making processes from multiple users. It’s also difficult to measure the motivation of users to select a given node. Human users might have key awareness of the task that may render them select nodes for different reasons. Expected Utility is an area of study in Risk management [18]. An utility probabilistic model is designed to reflect a users intrinsic motivations, not limited to utility, risk propensity, and the influence in learning within a single decision episode, or across episodes. We expect to present preliminary results at the workshop.

An intuitive illustration of reverse execution is presented in Fig. 6.1. We assume the robot performs the current behavior $B_i = \{N_s^i, N_g^i\}$ and an accidental fault $F_x$ is detected, where $N_s^i$ and $N_g^i$ indicate the starting and goal node, respectively. Subsequently, a new exploring node named $r$ for responding this fault is autonomously appended to original task graph $G$ as illustrated in Sect. 6.3, that is

$$r : \mathcal{T}_{B_1, B_i} \cup \mathcal{F}_x, B_i$$

(6.2)

![Fig. 6.1 Exploring the task via reverse execution when a fault event occurs](image)
As formulated in Eq. (6.4), the symbol $\mathcal{B}_* = \{\mathcal{N}_s^*, \mathcal{N}_g^*\}$ denotes an optimal way that consists of one or more selected movement primitives for retrying under current fault type and behavior situation. So, the critical problem is how to parameterize the transition probability $p(\mathcal{I}_{\mathcal{B}_i}, \mathcal{B}_*)$ given $\mathcal{F}_x$ and $\mathcal{B}_i$ when more than one ways to explore. To address this, a statistical distribution is introduced, which the instances are independent identically distributed and belong to a discrete distribution as well as the sum of the transient probabilities of all the samples after a fault occurrence is equal to 1. For these reasons, we define the $p(\mathcal{I}_{\mathcal{B}_i}, \mathcal{B}_*)$ is modelled with a multinomial distribution that the random variables is the respective frequency counted based on human intention when a fault occurs. For example, $\mathbf{N}_{\mathcal{F}_x} = (N_1, N_2, \ldots, N_K)$ is a frequency distribution of a random variables vector, where $K$ represents the total number of movement primitives that from beginning to current movement $\mathcal{B}_i$ with fault $\mathcal{F}_x$ (including the $\mathcal{B}_i$) and $N_i, i \in \{1, 2, \ldots, K\}$ denotes how many times of movement $\mathcal{B}_i$ is successfully executed. Therefore, the probability mass function of transition $p(\mathcal{I}_{\mathcal{B}_i}, \mathcal{B}_*)$ is formulated as a multinomial distribution that model a total $N = \sum_{i=1}^{K} N_i$ reverse executions, that is,

$$p(N | N, \theta) = \frac{N!}{N_1! \ldots N_K!} \prod_{i=1}^{K} \theta_i^{N_i}$$  

where, $\theta_i$ indicates the probability of movement primitive $\mathcal{B}_i$ is selected, which subject to $\theta_i \in [0, 1]$ and $\sum_{i=1}^{K} \theta_i = 1$. Therefore, we use the multinomial distribution not only to intuitively depict the expectation of human intention on the recovery behavior when an abnormal occurrence is detected in human robot interaction scenarios, but also to provide an indirect way to express the intuitive understanding of human for expecting the motion of the robot end-effector, the related manipulation objects as well as the complex relationship between “human-robot-environment”.

### 6.6 Human Interaction Policy for Persistent Anomalies

Human interaction allows the robot explore the current movement by human-assist demonstration for resolving the persistent faults that can’t be restored by reverse execution, such as tool collision and wall collision. The key question is how to fast capture and learn from the interactive demonstration base on the defined task representation. This exploring policy is activated when the system fail more than two times when reverse executions is carried out for resolving the fault $\mathcal{F}_x$. The human interaction policy is another exploring way that through human assisted demonstration when robot encounter a persistent fault. It’s essential that synchronously capture the human kinesthetic demonstration that not only the kinematic variables but also the relative coordinate frame (updating the task structure defined in Sect. 6.3 by understating the transformation relationship between demonstration and original movement) at each time step. After that, the movement is formulated by the DMP techniques intro-
duced in Chap. 3. In particular, the human interaction policy is not only limited to the originally designed movements, but also can be applied to the exploring movement (reverse execution or human interaction). Theoretically, our system can handle any kinds of failure situation with those aforementioned two exploring policies such that potentially achieve the long-term autonomy during robot manipulation task. Additionally, experiential verification indicates that another problem arises, there are faults also exist in reproducing the human interaction. To address this problem, we introduce a data augmentation method (without detail statement in this paper) for training a new exploring movement and introspective model (described in Chap. 4) when the same persistent fault is continuously encountered more than three times in the same (multimodal signals are recorded at each time). Consequently, we can achieve the critical behavior for incrementally addressing the faults by “exploration of exploration” and “anomaly monitoring and diagnose in exploration”.

Without loss of generality, an intuitive illustration of human interaction is presented in Fig. 6.2. We assume the robot performs the current behavior $B_j = \{N_j^s, N_j^g\}$ and an accidental fault $F_y$ is detected, where $N_j^s$ and $N_j^g$ indicate the starting and goal node, respectively. Subsequently, a new exploring node named $A$ for responding this fault is autonomously appended to original task graph $G$ as illustrated in Sect. 6.3, that is

$$A : \mathcal{T}_{B_j, B_h} | F_y, B_j$$

where, $B_h$ indicate the human interactive demonstration $B_h = \{N_h^s, N_h^g\}$, that have the same starting node $N_h^s$ with movement $B_j$ and a new goal node $N_h^g$ is derived from demonstration. Equipping this exploring movement $B_h$ by formulating it using the dynamical movement primitive. Additionally, we define the goal pose (end-effector or joint variables) of demonstration as $P$ that only task kinematic variables into consideration for task exploration. The $P$ can be adapted by the transformation relationship between $N_h^s$ and $N_h^g$, then $N_h^g$ derived by

$$N_h^g = P \mathcal{T}_{N_h^s, N_h^g}$$
6.6 Human Interaction Policy for Persistent Anomalies

Fig. 6.3 Illustrates the human interactive demonstration in our kitting experiment when the robot collides with the packaging box during transporting an object.

As Fig. 6.3 shown a human robot collaborative task is designed in latter experimental verification, where a Baxter robot encounter a persistent fault (wall collision) during transporting object from human-over. In this situation, the robot likely to encounter a fault along the deficient movement (as shown in Fig. 6.3a, and the fault can’t be eliminated by reverse execution because of the fixed obstacle (box) on the right hand side of robot. An modulated trajectory should be introduced for updating the original task representation, as shown in Fig. 6.3b a human interactive demonstration for exploring the failure task. Subsequently, a transformation in Eq. (6.5) is derived by learning from interactive demonstration, which can be explored in a new scenario when encounter a same fault, as shown in Fig. 6.3c.

6.7 Experiments and Results

6.7.1 Platform Setup

To evaluate the proposed method for incremental learning the introspective movement primitives. We designed a HRC task for picking and placing object into a
An autonomous kitting experiment by Baxter robot: The robot is designed to transport 6 marked objects with variable weights and shapes to a container, where the external anomalies may arise from accidental collisions (human-robot, robot-world, robot/object-world), mis-grasps, object slips. So as to identify those unexpected anomalies, the robot arm is integrated with multimodal sensors (shown in bottom-right), including internal joint encoders, F/T sensor, tactile sensor.

In consideration of the redundant features would aggravate computational efficiency and increase false-positive rate (fault occurs even when robot's movement is
Fig. 6.5 The Snapshots of human-robot kitting experiment with a Baxter robot. We subjectively assume this task consists of 5 individual movements, including a Home $\rightarrow$ Pre-pick; b Pre-pick $\rightarrow$ Pick; c Pick $\rightarrow$ Pre-pick; d Pre-pick $\rightarrow$ Pre-place; e Pre-place $\rightarrow$ Place; f Reset

Fig. 6.6 Multimodal signals in kitting task, where the grey represents the invalid robot movements including the adjustment, orientation alignment, reset, etc, other colors are empirically indicated an potential movement primitive along with obvious signals fluctuations

(normal), we perform empirical features extraction on the original observation vector to improve identification performance. Specifically, we compute the norm of both the force $n_f$ and the moment $n_m$ as features in wrench modality, take the norm of both the Cartesian linear $n_l$ and angular $n_a$ velocities in velocity modality, and consider the standard deviation for each tactile sensor $s_l$ and $s_r$. Therefore, our feature vector $y_t$ of length 6 and formulated as $y_t = [n_f, n_m, n_l, n_a, s_l, s_r]$, evolving extracted features of ten nominal executions are illustrated in Fig. 6.6.
6.7.2 Parameter Settings for Anomaly Monitoring and Diagnose

We need both qualitative and quantitative analysis of the proposed method. To evaluate the whole performance of IMPs in unstructured scenarios with unexpected faults that including both of the accidental and persistent causes. We organized 5 participants as collaborator (one expert user who confidently know this implementation and other four novice users) in our designed kitting experiment. Novice users first learned from the expert to induce fault during robot executions, which would aggravate the external uncertainty and increase the modeling difficulties. During data collection, each participant performed 1 nominal and 6 executions that at least one fault event by placing the set of 6 household objects in a one-by-one fashion. Consequently, totally perform 30 nominal executions and 180 failure executions. We induce the fault manually for each movement primitive, where including but no limited to the Human Collision (HC), Object Slip (OS), Tool Collision (TC), Wall Collision (WC), etc.

We first evaluate the complex task segmentation from twenty whole kinesthetic demonstrations using Bayesian nonparametric methods. As shown in Fig. 6.7, illustrating the complex task segmentation by learning the underlying dynamics using HDP-VAR-HMM, where each row indicates an independent demonstration and different color represents a specific movement primitive, which no need to guarantee each demonstration have the same segmentation order. After then, we concentrate the ordered hidden state sequence and group them by hidden state transition pair \( t_{Pair} \), where the total number of pair is computed using the permutation combination algorithm, i.e. \( t_{Pairs} = C_5^2 A_2^2 = 20 \), where \( K \) is the dimension of hidden space. Thus, the frequency among the possible transition of the learned five movement primitives is illustrated in Fig. 6.8, where the kitting experiment is definitely begin with movement 2 (clustered by the hidden state 2) and then the successor should be movement 1, subsequent movement is 4 for most cases or movement 3 is the second-choice, and so on. Until now, we can effective learning the complex task representation for a set of nominal unstructured kinesthetic demonstration in a manipulation graph way. With this representation, we can evaluate the following the anomaly monitoring, diagnose as well as task exploration, respectively.

According to Chaps. 4 and 5, our proposed movement primitive equipped with two introspective capabilities: anomaly monitoring and diagnose, that necessary for endowing robot long-term autonomy and safer collaborative interaction in human-robot collaborative scenarios. Particularly, our anomaly monitoring and diagnose are implemented based on the Bayesian nonparametric models proposed in Chap. 2 that using the HDP-VAR-HMM with a first-order autoregressive Gaussian likelihood, each state \( k \) has two parameters to be defined: the regression coefficient matrix \( A_k \) and the precision matrix \( \Lambda_k \) as well as the four parameters \( \nu, \Delta, V, M \) of the conjugate prior MNIW are assigned in advance. To guarantee the prior has a valid mean, the degrees-of-freedom variable is set as \( \nu = d + 2 \) and \( \Delta \) is set by constraining the mean or expectation of the covariance matrix \( \mathbb{E}[\Lambda_k^{-1}] \) under the Wishart prior in
Fig. 6.7 Complex task segmentation by learning the underlying dynamics using HDP-VAR-HMM. Each row indicates an independent demonstration and different color represents a specific movement primitive, which no need to guarantee each demonstration have the same segmentation order.

Fig. 6.8 Illustrates the transition frequency among the derived five movement primitives. The hidden state sequences is first taken to extract the movement primitives from a set of demonstrations. A uncompleted manipulation sequence is ordered with $2 \rightarrow 1 \rightarrow 0 \rightarrow 4 \rightarrow 3$ based on the derived sequences. The numbers on the transitions correspond to the transition frequency for weakly representing the task at the beginning, and subsequently strengthening by incremental learning the introspective capabilities and exploration from practical applications.
Learning Policy for Robot Anomaly Recovery Based on Robot Introspection

\[ \mathbb{E}[A_k^{-1}] = s_F \sum_{n=1}^{N} \sum_{t=1}^{T_n} (y_t - \bar{y})(y_t - \bar{y})^T. \]  

(6.6)

Assume that we record \( N \) sequential data for each skill and the length of sequence \( n \in N \) is \( T_n \). Thus, we can easily define the parameter \( \Delta \) accordingly as

\[ \Delta = (\nu - d - 1)\mathbb{E}[A_k^{-1}]. \]  

(6.7)

We placed a nominal prior on the mean parameter with mean equal to the empirical mean and expected covariance equal to a scalar \( s_F \) times the empirical covariance, here \( s_F = 1.0 \). This setting is motivated by the fact that the covariance is computed from polling all of the data and it tends to overestimate mode-specific co-variances. A value slightly less than or equal to 1 of the constant in the scale matrix mitigates the overestimation. Also, setting the prior from the data can move the distribution mass to reasonable parameter space values. The mean matrix \( M \) and \( V \) are set such that the mass of the Matrix-Normal distribution is centered around stable dynamic matrices while allowing variability in the matrix values (see Chap. 2 for details).

\[ M = 0_d, \]
\[ V = 1.0 \ast I_d. \]  

(6.8)

where \( 0_d \) and \( I_d \) are the matrix of all zeros and the identity matrix, respectively, each of size \( d \times d \). For the concentration parameters, a \( \text{Gamma}(a, b) \) prior is set on HDP concentration parameters \( \alpha \). A \( \text{Beta}(c, d) \) prior is set on the self-transition parameter \( \mu \) and the degree of self-transition bias \( \kappa \) is set to 50. We choose a weekly informative setting and choose:

\[ a = 0.5, b = 5.0, c = 1.0, d = 5.0 \]  

(6.9)

where, the initial transition proportion parameter is defined as \( \mu \sim \text{Beta}(1, 10) \) and the Split-Merge Monte Carlo method sampling parameter maximum iterations is set to 1000. The truncation states number is set to \( K = 5 \) for anomaly monitoring, and \( K = 10 \) for fault diagnose.

The anomaly monitoring is implemented by comparing the cumulative likelihood \( L \) of observed observations, where the fault threshold is calculated from a set of \( L_i, i \in \{1, 2, ..., 20\} \) nominal demonstrations, and formulated with the expected cumulative likelihood \( \mu(L) \) minus the standard deviation \( \sigma(L) \) that multiply by a constant value \( c \). Additionally, the fault diagnose is activated when fault detected, which mainly implemented by comparing the sum of log-likelihood of a failure sample in a supervised fashion. Particularly, we use the K-Folders Cross Validation.
6.7 Experiments and Results

As shown, the five normal IMPs is derived by learning from unstructured demonstrations that successfully performing human-robot collaborative kitting task and the dark-blue trajectories are explored with the maximum probability in our defined reverse policy once the fault detected defined in sklearn⁴ (here, \( K = 3 \)) for model selection, in which the objective is the accuracy of anomaly monitoring with a fixed constant value \( c = 3 \).

To achieve the incremental learning introspective movement primitive from unstructured demonstrations, another critical capability would be the task exploration under the unseen and unexpected situation, especially when robot encounter a failure event. There are two independent exploring policies: reverse execution and human interaction, are proposed for responding the external accidental and persistent faults, respectively. We test the whole performance in two scenarios for evaluating how to combine the anomaly monitoring, fault diagnose, and exploration in practical applications along with the quantitative performance. As illustrated in Fig. 6.9, a set of collective 3D moving trajectories is presented for evaluating the exploring reverse execution when robot encounter accidental faults. Where the kitting task is first represented by five introspective movement primitives, and then the accidental faults (marked in yellow dots) randomly happened during robot execution, the robot immediately perform the reverse execution (dark-blue in color) according the frequency distribution shown in Table 6.1.

Additionally, As illustrated in Fig. 6.10, a set of collective 3D moving trajectories is presented for evaluating the exploring human interaction when robot encounter an wall collision that is a persistent fault. To evaluate the overall performance after human one-shot interaction, we get a 80.95% success rate when the wall collision induced by repeating the experiment 30 times.

---

⁴https://scikit-learn.org.
### Table 6.1
The frequency distribution of reverse execution policy in kitting experiment, where we record the data from five participants in total of 30 times of each accidental fault. Additionally, the 2 is an union of robot performing $Pre \rightarrow \text{pick} \rightarrow \text{Pick}$ and $\text{Pick} \rightarrow Pre \rightarrow \text{pick}$

| Fault diagnose results | Fault occurs | Goal IMP | Frequency |
|------------------------|--------------|----------|-----------|
| Human collision        | 1            | 1        | 25        |
|                        | 2            | 2        | 30        |
|                        | 3            | 3        | 25        |
|                        | 4            | 4        | 25        |
| Tool collision         | 2            | 1        | 2         |
|                        | 3            | 3        | 20        |
| Object slip            | 2            | 2        | 5         |
|                        | 3            | 1        | 25        |

**Fig. 6.10** Visualizing the robot trajectories after human interaction when persistent faults are detected. As shown, the five normal IMPs is derived by learning from unstructured demonstrations that successfully performing human-robot collaborative kitting task and the dark-blue trajectories are incrementally explored by updated the demonstrated IMP parameters (including the start and goal pose, shape) once the fault detected

### 6.7.3 Discussion

In our previous work, we focus on the robot introspective capabilities in robotics, which only take the kinematic variables into consideration such that the task is restricted to be applied in human-robot collaborative scenarios that generalization (including motion, introspection, and decision-making, etc.) as a desirable characteristic. To address this problem, robot movement primitive augmented with introspective capacities IMPs is investigated in this paper, which by associating the generalization, anomaly monitoring, fault diagnoses, and task exploration during robot manipulation task. We mainly introduce the IMPs can be acquired by assessing the quality of multimodal sensory data of unstructured demonstrations using a nonpara-
metric Bayesian model, named HDP-VAR-HMM, and IMPs can incremental learning the exploring policy using multimodal distribution and human one-shot modification when robot encounter fault. Particularly, reverse execution and human interaction are two independent policies for task exploration, which proposed to respond the external accidental and persistent fault, respectively. Experimental evaluation on a human-robot collaborative packaging task with a Rethink Baxter robot, results indicate that our proposed method can effectively increase robustness towards perturbations and adaptive exploration during human-robot collaborative manipulation task. We need to emphasize that our method presents a solution for endowing robot with introspection in sense-plan-act control methodology for robot manipulation task.

Recently, we are working on extending IMPs to be more human-like manipulation that including the visual and audial information for human robot collaborative electronic assembly task. We are also investigating the use of variational recurrent autoencoder neural network to facilitate our proposed framework for more complex scenarios.

6.8 Summary

In this chapter, we present two policies to deal with accidental and persistent anomalies respectively: movement reverse execution and human interaction. Reverse execution allows the robot retry the current movement or several movements that independent with the current state for resolving the accidental faults by redo current or some previous movement primitives by updating the parameters of primitives (including shape, starting point and target point, etc.), and to use polynomial distribution to realize the modeling of a primitive after the occurrence of an anomaly occurrence. Human interaction allows the robot explore the current movement by human-assist demonstration for resolving the persistent faults that can’t be restored by reverse execution, such as tool collision and wall collision, which mainly to learn the anomaly behavior from the human demonstration, and can realize the growth of the task representation of the recovery behavior. Importantly however we learned that the recovery ability of the system grows in difficulty with an increased number of adaptations as variations in sensory-motor signals increase as more recoveries are attempted.

The proposed recovery policies are verified on a Baxter robot performs kitting experiment tasks, results indicate that the proposed policies meet the extensibility and adaptability for improving the long-term autonomy, an integrated to the SPAIR framework. Consequently, this book provides an efficient theoretical framework and software system for the implementation of the longer-term autonomy and a safer environment for human-robot interaction scenarios. Ultimately the system presented in this book significantly extended the autonomy and resilience of the robot and has broad applicability to all manipulation domains that suffer from uncertainties in unstructured environments: making industrial and service robots prime candidates for this technology.
In this book, anomalies are and will continue to be a reality in robotics despite increasingly powerful motion-generation algorithms, to address them explicitly, we presented a tightly-integrated, graph-based online motion-generation, introspection, and incremental recovery system for manipulation tasks in loosely structured co-bot scenarios, which consist of the movement identification, task representation, anomaly monitoring, anomaly diagnoses, and anomaly recovery.

References

1. Gutierrez RA, Chu V, Thomaz AL, Niekum S. Incremental task model updates from demonstration.
2. Pastor P, Righetti L, Kalarkishnan M et al. Online movement adaptation based on previous sensor experiences. In: 2011 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE;2011. p. 365–71.
3. Pastor P, Kalarkishnan M, Righetti L et al. Towards associative skill memories. In: 2012 12th IEEE-RAS international conference on humanoid robots (humanoids). IEEE;2012. p. 309–15.
4. Xu Z, Li S, Zhou X, Wu Y, Cheng T, Huang D. Dynamic neural networks based kinematic control for redundant manipulators with model uncertainties. Neurocomputing. 2019;329(1):255–66.
5. Xu Z, Li S, Zhou X, Cheng T. Dynamic neural networks based adaptive admittance control for redundant manipulators with model uncertainties. Neurocomputing. 2019;357(1):271–81.
6. Salazar-Gomez AF, DelPreto J, Gil S et al. Correcting robot mistakes in real time using eeg signals. In: 2017 IEEE international conference on robotics and automation (ICRA). IEEE;2017. p. 6570–77.
7. Xu Z, Li S, Zhou X, Yan W, Cheng T, Dan H. Dynamic neural networks for motion-force control of redundant manipulators: an optimization perspective. IEEE Trans Ind Electron, Early access. 2020. https://doi.org/10.1109/TIE.2020.2970635.
8. Niekum S, Chitta S, Barto AG et al. Incremental semantically grounded learning from demonstration. Robotics: science and systems. Berlin, Germany:Springer;2013. p. 9.
9. Mao R, Baras JS, Yang Y et al. Co-active learning to adapt humanoid movement for manipulation. In: 2016 IEEE-RAS 16th international conference on humanoid robots (humanoids). IEEE;2016. p. 372–8.
10. Karlsson M, Robertsson A, Johansson R. Autonomous interpretation of demonstrations for modification of dynamical movement primitives. In: 2017 IEEE international conference on robotics and automation (ICRA). IEEE;2017. p. 316–21.
11. Karlsson M, Robertsson A, Johansson R. Convergence of dynamical movement primitives with temporal coupling. In: 2018 European control conference (ECC). IEEE;2018 Jun 12. p. 32–9.
12. Rodriguez A, Mason MT, Srinivasa SS, Bernstein M, Zirbel A. Abort and retry in grasping. In: 2011 IEEE/RSJ international conference on intelligent robots and systems. IEEE;2011 Sep 25. p. 1804–10.
13. Laursen JS, Schultz UP, Ellekilde LP. Automatic error recovery in robot assembly operations using reverse execution. In: 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE;2015 Sep 28. p. 1785–92.
14. Laursen JS, Ellekilde LP, Schultz UP. Modelling reversible execution of robotic assembly. Robotics. 2018;36(5):625–54.
15. Muxfeldt A, Steil J. Fusion of human demonstrations for automatic recovery during industrial assembly. In: 2018 IEEE 14th international conference on automation science and engineering (CASE). IEEE;2018 Aug 20. p. 1493–00.
16. Muxfeldt A, Steil JJ. Recovering from assembly errors by exploiting human demonstrations. Proc CIRP. 2018;1(72):63–8.
17. Arne M. Automatic error recovery during industrial assembly operations based on human demonstrations. PhD diss.: Technische Universität Carolo-Wilhelmina zu Braunschweig; 2018.
18. National Research Council. Modeling human and organizational behavior: application to military simulations. National Academies Press; 1998 Aug 31.
19. Wu H, Xu Z, Yan W, Su Q, Li S, Cheng T, Zhou X. Incremental learning introspective movement primitives from multimodal unstructured demonstrations. IEEE Access. 2019;15(7):159022–36.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.