A Robotic Mind Model for Affective Decision Making and Behaviour Generation

Jinwei Zhang

Master of Science
School of Informatics
University of Edinburgh
2023

Abstract

Affective robots are of interest for many applications, to accompany people, assist in education, and reassure patients. However, the principles for control of such a robot are not yet established, and its specifications are not clear. Here we propose a framework for a lively robot that can reason to make decisions and maintain emotions and intentions. This Affective RObotic Mind (AROM) model is designed and implemented, aiming to enhance the interpretability of affective robot behaviour for humans in various social situations encountered by the robot. Inspired by biological examples, AROM integrates robot decision-making and behaviour-generation based on consistent affective support for taking action based on the decisions and in order to express internal states via body movements. The core factors that adjust decision-making and movement patterns are abstracted from the physiological system that is shared by most vertebrates. The system is thus intended as an exemplary architecture for affective robot design. To evaluate the achieved affective behaviour in the robot, experiments on human subjects are conducted related to the function of the decision-making module, the emotion-related enhancement of interpretability of the behaviour-generation module, and on the holistic performance of the combination of these modules. The project does not address affect issues from the aspect of hardware design or long-term effects of robot interaction such as the idea of robot personality.

Acknowledgements

Many thanks to my supervisor Michael Herrmann. He gave me much valuable guidance and support. Also thank Isobel Voysey, Minqiu Zhou, and all the people who helped me and warmed me. Finally, I would like to thank my family.

Table of Contents

Chapter 1

Introduction

With the rapid development of robotics, robots are useful in a multitude of tasks in the current era, no matter in industry or people's daily lives. Powered by Artificial Intelligence (AI), the application fields of robots are expanded greatly, from healthcare and medical services [37, 42, 25], transportation and road traffic [29, 17], manufacturing industry [11, 31] to education [5, 15]. The ever-growing human-robot interaction makes people rethink questions: What kind of robots do we humans want? And whether it is necessary to assign robots features of emotion, to promote potential delightful experiences while interacting with robots? What else can affective robots bring to humans? Facing these questions, many researchers propose hypotheses and approaches. Stock et al. [44, 50] believes that affective robotics can reduce human cognitive burdens in Human-robot Collaboration workspace. Cabibihan et al. [5] insist that social robots have significant benefits in the therapy of autism children because they are more predictable and adaptable. Under such kind of belief, relative research fields like Affective Computing and Human-Centered Calculation thrived, shedding new light on Robotics and promoting development of Affective Robotics.

Despite multiple attemps are made in the past decades, the theories and methods of Affective Robotics are still limited and under explored. For emotion model functioning as the core of robot systems, there is no universal model which supports systematic applications, i.e. from behaviour generation to making decisions. In addition, there lacks a general control logic for robot activities according to emotions. Many studies apply the basic discrete emotion model to robotics systems [18, 2, 51], though achieved meaningful results, are restricted on the potential of generating and controlling movement, because of the nature defects of discrete models on control issues. Research of generating behaviours are mostly focused on outputting facial expressions [16, 51], and the studies that do enable robots to move body parts emotionally are usually just mimicking human or a specific type of animal [2], lacking of adaptivity. For emotion-intervening robot decision-making however, as far as the author knows, is still a very fresh field. Aiming to tackle these problems, this project proposed *AROM*, an Affective RObotic Mind model for emotion-intervening decision making and behaviour generation. It provides a general, adaptive and consistent solution for emotion model, robot behaviour-generating and decision making, with enough interpretability for human to understand.

The *AROM* system applies a novel biomimetic continuous emotion model, and retains the ability to generalize at the same time, through combining with the widely used PAD emotion model. Inspired by the emotion theory proposed by Rolls [35], the emotion model of *AROM* is able to fluctuate emotions according to the rewards and context information, enabling being utilized in decision-making of learning tasks. In the behaviour-generating module, *AROM* defines several affective motion parameters, as well as an algorithm for generating behaviours with only robot body parameters required, and can be applied on alien robots. In addition, the decision-making module is able to conduct emotionless calculations as well as emotional ones with constantly changing emotions, where MDP and Thompson Sampling (TS) are applied as the core methods. All parts of the system share the same affective framework, and are able to combine and generate complex motions under affective decisions and posture expressions. In order to evaluate the performances of *AROM* system, human experiments are conducted to examine perceptions of emotions that the robot reveals. Simulations of robot decision-making when facing an unexpected adversity are also implemented under different settings of emotion model . The aim is to test whether emotions can help robot adapt to the changing environment.

Centering on the *AROM* system, two main research questions will be explored in this thesis:

1. Except for promoting communication and mutual understanding, does emotion has other benefits if set into robot systems? Can the principle of emotions, or the physiological reactions behind it help robots to adapt to the environment?

2. Although people already can design emotional expressions for specific affect states on various types of robots, those expressions do not follow general principles, and are usually not designed for continuous emotions. Does there exist an

emotion-behaviour mapping framework that is general and adaptive for robots, and can be easily applied for diverse tasks and trivial emotion states?

The overall contribution of this project is to propose a novel framework for affective robot decision-making and behviour generation, the results of which are adaptable and interpretable. Related works of this project are demonstrated and discussed in Chapter 2. In Chapter 3, the methods of this project are illustrated, with Chapter 3.1 showing the list of Symbols of the methodology section. Chapter 3.2 illustrates the novel emotion model of *AROM*, with Chapter 3.3 demonstrating the decision-making module. In Chapter 3.4, the behaviour-generation module is illustrated. Then, in Chapter 4, which is the chapter of experiment, the experiments of decision-making and behaviour-generating are shown in Chapter 4.2 and Chapter 4.3 respectively. The discussion of this project is illustrated in Chapter 5. Finally, the conclusions of this thesis are drawn in Chapter 6.

Chapter 2

Related Work

2.1 Robotic Emotion Models

Most of the Affective Robotics studies utilize emotion models from psychology. Commonly used models can be divided into two types: discrete model and dimensional model, which is also mentioned as the continuous model. The most known discrete emotion model is Paul Ekman's theory of six basic patterns of expression (happiness, anger, disgust, fear, sadness, and surprise) [13]. With emotional parameters being clearly defined and categorized, Ekman's model is applied in many robotics study and help generation and recognization of emotional behaviours. However, as a discrete model, the expression space of it is limited compared with dimensional model, where emotions can be described continuously using dimensions.

The existing dimensional model can have distinctive numbers of dimension, but the most widely applied one in robotics is the PAD emotional state model developed by Mehrabian. It uses three numerical dimensions of Pleasure, Arousal and Dominance to express all the emotional states [28]. Another three-dimension model is Lövheim cube [27], which is an explanatory model for emotions and monoaminergic neurotransmitters. Lövheim picked monoamines Dopamine, Serotonin and Noradrenaline as the three dimensions, and illustrated the emotions arisen with fluctuations of the three monoamines. Whereas, these two emotional models both have disadvantages on Human-Robot Interaction practice. It's hard to find traces and research on how emotions influence decision-making process and provide robotics a good reference if only based on the PAD model. Under the construction of Lövheim cube and monoamine, though there is sufficient evidence on how monoaminergic neurotransmitters influences decision-making process, the empirical study of relationship between Lövheim's emotion model and behaviour-generation is very rare. We resolve this problem by abstracting featured functions of monoamines and combining them with emotional dimensions that are massively used in behaviour experiments, and enable our model to support both decision-making and behaviour generation for robots.

2.2 Affective Decision Making

Historically, most research on affective robot decision-making was carried out by coalescing physiological features to decision-making systems. Lewis et al. [24] proposed a robot architecture to investigate the roles of Pleasure plays in the action selection process in survive tasks. During the decision-making tasks, the pleasure hormone alters the incentive salience of perceived stimuli by acting on the 'subjective assessment' or 'assignment of value'. Cos et al. [8] designed a homeostatic adaptive mechanism based on RL to modulate the internal deficits of a social agent. The robot tries to maintain a good physiology state by learning the optimized actions in variant situations, and the physiological dimensions of Tiredness and Hunger evolve over time. Together these studies provide important insights into the functions of physiological factors in survive tasks, whereas the affect dimensions involved are limited, unable to describe more complex states of robots.

A more comprehensive set of basic emotions is utilized in Ref. [14], which are hope, fear, happiness and sadness. In this study, an affective model for decision agent in non-expensive robotic platforms is demonstrated. Under the Adversarial Risk Analysis framework, the agent makes decisions according to emotional factors in a human-robot interaction scenario. However, this affective model is mainly oriented towards robot decision-making, and the emotional dimensions are insufficient to be used in comprehensive behaviour generation, which requires more affective factors and cannot function well based on vague states of hope, specifically. Motivated by establishing a framework which both possesses physiological features of affective decision-making, and also utilize the same emotion model which supports robot behaviour-generating, we designed the Decision-Making module of *AROM*, inspired by functions of three monoamines and supports updating emotions for generating behaviours.

2.3 Affective Behaviour Generation

Many research works on affective robot behaviour generation are conducted by mimicking human behaviours. Suguitan et al. [45] proposed a method for automatically synthesizing affective robot movements based on generative adversarial neural networks (GANs). It utilizes source human behaviours to generates robot motions which matches the affective labels of human movements. Ali et al. [2] focused on affective non-verbal human-robot interaction, applying 28 overlapping emotions including 8 primary categories based on Russell's circumplex model. Sourced from emotional behaviour research on human, it links robot physical postures and movements with emotions using features such as speed, frequency, and joint angles. However, such studies remain narrow in focus dealing only with robot body behaviours, without paying attentions to trajective generation and affective navigation.

To fill such gap, research Ref. [38] studied robot motion and kinematics variables' impact on people in practice of navigation. The trajective movements are specified by velocity profiles i.e. velocity curves of the mobile robot over time, using periodic perturbation as variants to the profiles. Results of it show that using saccadic or smooth motion can impact people's perceptions of robots' emotions, like confident, frail, aggressive and confident. Nevertheless, the trajectory movements in this study is still simple, only supporting moving in a straight line. To enable richer affective navigation expression as well as according body movements, the Behaviour-Generating module of *AROM* is designed, which also aims at promoting adaptivity to more general types of robots, except for humanoid ones.

Chapter 3

Methodology

The affective behaviour-generation integrated robot decision-making model (*AROM*) is inspired by the monoamine neurotransmitter systems shared by virtually all vertebrates, functioning to promote the adaptability to various environments [6, 33]. Monoamines like dopamine, serotonin and noradrenaline adjust physiological reactions, and influence the motivations and emotions on the levels of both decision-making and behaviour adjustment. Targeting on similar functions, the *AROM* model provides a novel affective robot control architecture, enabling the module of decision-making and behaviour generation be integrated, and influenced together by specific emotional states on the current time.

Diffent from traditional methods in robotics, which usually compute decisions of actions and design emotional motion patterns separately, the *AROM* model supports display of consistent affects on both of these levels, and enables robots to react to environmental stimuli on a vivid and quasi-natural way. The hypothesis is, under *AROM* model, affective robots can be perceived as more emotionally prominent and their intentions are easier to understand.

As illustrated in Figure 3.1, the *AROM* model is composed of the following three main components.

- *Affect Calculation*: Core module of the *AROM* model, it influences both the Decision-Making and Behaviour Generation modules.
- *Decision Making*: Mainly based on a Markov Decision Process (MDP), it utilises parameters of physiology and emotion to decide structural actions such as to approach the human partner or to play alone.

- *Behaviour Generation*: Controlled by physiological and emotional parameters, it generates behaviour details based on the decisions according to certain motion parameters like velocity, force and stretching level.

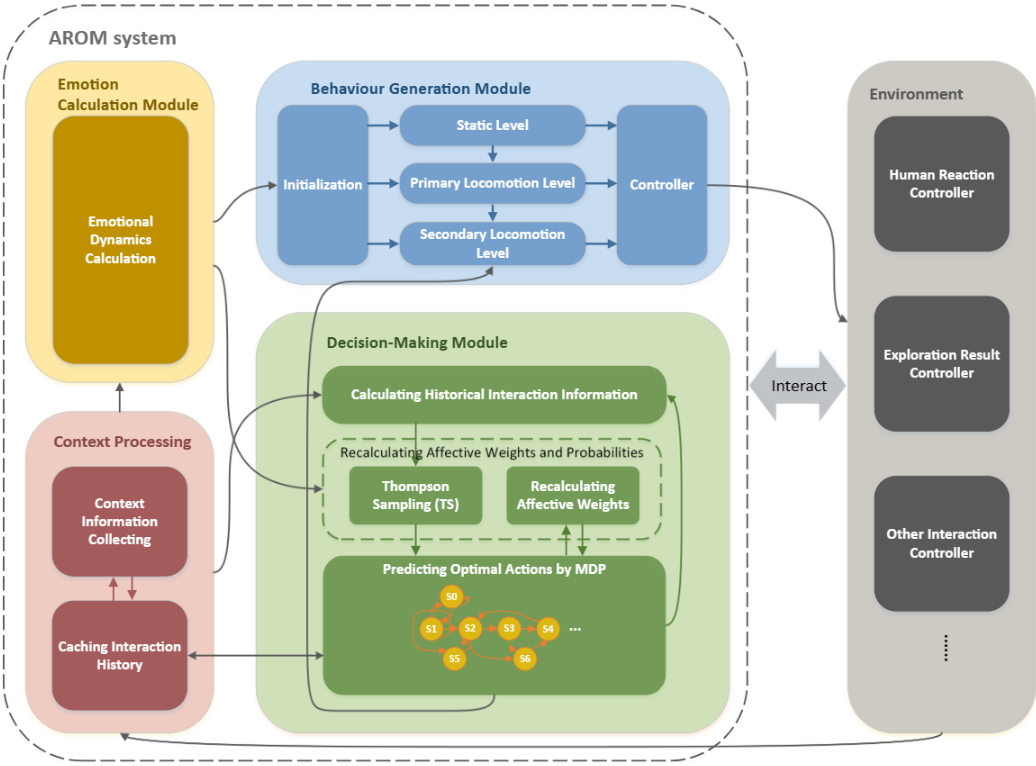


Figure 3.1: Global system structure of the Affective RObotic Mind (AROM) model that is proposed here, featuring modules for emotion calculation, behaviour generation, and decision making. Interaction with the environment is based on sensory input entering via context processing unit and behavioural output via the behaviour generation module. For details, see the respective sections below.

In addition to its main components, the model makes use of a *Context Processing* module and an Environment module which are also shown in Figure 3.1. The Context Processing module is built in the *AROM* system, which continuously collects context information and caches interaction history. The *Environment* module shows the origin of possible interaction results, such as human reaction controller. The system keeps interacting with the *Environment* in the Human-Robot Interaction scenario.

3.1 List of Symbols

See Table 3.1.

Table 3.1: List of symbols in the Methodology section

Symbols	**Meaning**
Parameters for emotional dynamics and interactions	
E	level of one dimension of affect
E_t	level of one dimension of affect at time t
E_0	level of original affect
P	level of PLEASURE
A	level of AROUSAL (Noradrenaline)
D	level of DOMINANCE (Serotonin)
D^A	level of DOPAMINE
t	time
R	value of stimuli
S^A	level of satiety
k_e	rate constant of damping of E
k_p	rate constant of damping of P
k_a	rate constant of damping of A
k_d	rate constant of damping of D
k_{da}	rate constant of damping of D^A
m_p	sensitivity of P to changes of the value of stimuli R
m_a	sensitivity of A to changes of the value of stimuli R
m_d	sensitivity of D to changes of the value of stimuli R
κ	sensitivity of D^A to changes of the value of satiety S^A
q	number of time steps of historical information that need to be remembered in calculation of D
Parameters for decision-making	
W	weight value for general states in the MDP
W_{idle}	weight value for idle state specifically
W_{bs}	original basic weight value without influences of affect
$W_{bs-idle}$	original basic weight value for idle state
λ	degree that DOPAMINE D^A can influence the values of weight W
μ	degree of Serotonin (DOMINANCE) D influencing weights W
δ	level of Serotonin (DOMINANCE) D influencing weights for IDLE state W_{idle}

Continued from Table 3.1

U	threshold for Noradrenaline (AROUSAL) A to influence Thompson Sampling (TS) and calculation of probability
α_{x_t}	parameter of beta function from Thompson Sampling (TS), see Algorithm 2
β_{x_t}	parameter of beta function from Thompson Sampling (TS), see Algorithm 2
r_t	reward in Thompson Sampling (TS), see Algorithm 2
P	transition probability in the MDP
P_{TS}	basic probability generated by Thompson Sampling (TS)
Parameters for behaviour generation	
H_{Sp}	degree of PLEASURE P influencing suppression
H_{Sd}	degree of DOMINANCE D influencing suppression
H_{Va}	degree of AROUSAL A influencing velocity
H_{Vd}	degree of DOMINANCE D influencing velocity
H_{Ap}	degree of PLEASURE P influencing amplitude
H_{Aa}	degree of AROUSAL A influencing amplitude
H_{Ad}	degree of DOMINANCE D influencing amplitude
H_{Fa}	degree of AROUSAL A influencing frequency
v_{Sd}	damping constant of DOMINANCE D regarding to suppression
v_{Va}	damping constant of AROUSAL A regarding to velocity
v_{Vd}	damping constant of DOMINANCE D regarding to velocity
v_{Aa}	damping constant of AROUSAL A regarding to amplitude
v_{Ad}	damping constant of DOMINANCE D regarding to amplitude
v_{Fa}	damping constant of AROUSAL A regarding to frequency

3.2 Emotion Model and Affect Calculation

3.2.1 Emotion Model

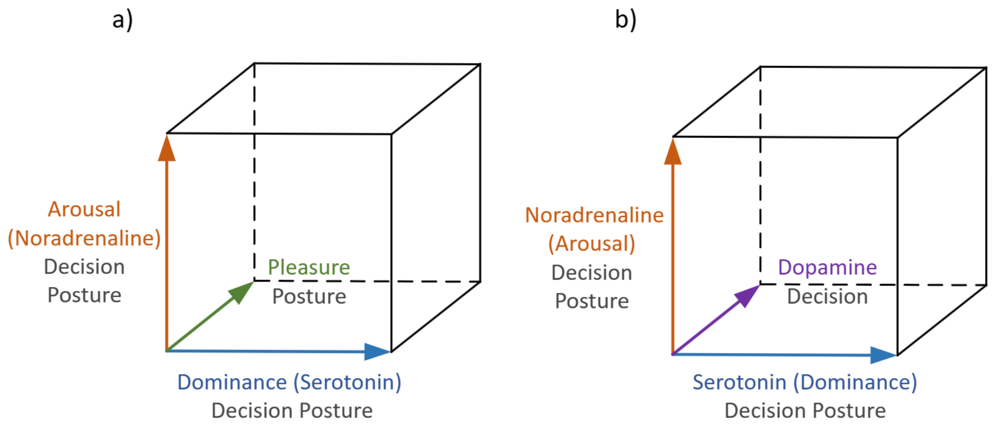


Figure 3.2: Monoamine-based emotion-behaviour model. a) Original PAD model [28] including monoamine-based interpretation. b) Lövheim cube of the relationship between the monoamine neurotransmitters and emotions [27]. In both models the information relevant for the AROM model (see Figure 3.1) is added.

Inspired by the three-dimensional monoamine and emotion model [27], the emotion model of *AROM* merges the monoamine model and PAD emotion model. The monoamines like Dopamine, Serotonin and Noradrenaline engage the adjustment and regulation of robot decision-making and behaviours, just as they do in vertebrate body.

The PAD emotion model is a generally used robot affect model, by utilising which is easier to refer and compare different standards of emotional motions. Although just as [27] declared, the three-dimensional monoamine model does not fully match the PAD dimensional emotion model, and there exists rotation of model comparing of two of these. In addition, the pleasantness (or valence) dimension are commonly taken in many emotion models, while the monoamine model [27] does not utilize it, thus there is a gap need to be filled.

In the *AROM* model, we construct an emotion model possessing features from both the monoamine system and from the PAD emotion model. In order to simplify the model and to reduce the number of dimensions, the SEROTONIN dimension in monoamine model is merged to the DOMINANCE dimension in PAD model, and the Noradrenaline dimension is merged to the AROUSAL.

As illustrated in Figure 3.2, the original monoamine dimensions are used to adjust the process of decision-making in the robot, while the original PAD emotion dimensions are to regulate parameters in the behaviour generation process. From the perspective of the monoamine model, the DOPAMINE dimension mainly influences the motivation level when making decisions, and the Serotonin (DOMINANCE) dimension adjusts the confidence and satisfaction of the decision-making process, while the Noradrenaline (AROUSAL) level has effects on making rational or risky choices. From the point of view of the PAD emotion model, the PLEASURE dimension can influence behavioural patterns as well as some motion parameters, such as speed. The AROUSAL dimension adjusts the degree of excitement of a movement by regulating parameters like speed, acceleration, time gap to change motion patterns, and force. Finally, the DOMINANCE (Serotonin) level influences the posture of behaviour generation, e.g., shrink body when losing DOMINANCE.

3.2.2 Affect Calculation and Response to Stimuli

Because affects are not always static, the calculation of how affect fluctuates and responses to stimuli are significant, in particular for a robot that can react to the external world and has various emotional expressions. During the update of emotions, two components are to be considered: One is the natural decay of emotions, and the other relates to the change of emotions when external stimuli arrive and interfere with the affective system.

For the damping of affect, we use an exponential function to mimic the kinetics of monoamine elimination:

$$E_t = E_0 e^{-k_e t} \tag{3.1}$$

Here E denotes the level of one of the affect dimensions, which can be DOPAMINE, Serotonin (DOMINANCE), Noradrenaline (AROUSAL) or PLEASURE. E_0 is the base affect level, and k_e is the rate constant of damping.

To tailor to the MDP-based affective decision-making process (see Section 3.3), the variation of emotions can be expressed as:

$$E_t = E_{t-1} e^{-k_e} \tag{3.2}$$

The time t in the expression $e^{-k_e t}$ is initially assigned a value of 1, which serves also as time step in the MDP dynamics.

For changes of emotions when exposed to various external stimuli (as illustrated in Table 3.2), the four dimensions are under the influence of the positive or negative

Table 3.2: The response to stimuli and satiety state in affect calculation. The symbol S represents stimuli. Sat refers to satiety as a physiological state. Since the three-dimensional monoamine system and PAD emotion model are merged here, some of the names of monoamines are shown in brackets. 5-HT represents serotonin, and NA represents Noradrenaline.

Various Situations and Affects								
Stimulus/Satiety	S−	S+	S+−↓	S+−↑	S+!	S−!	Sat+	Sat−
Direction of Change	−	+	−	+	−	+	−	+
Emotional States	VALENCE		AROUSAL (NA)		DOMINANCE (5-HT)		DOPAMINE	

nature of stimuli, the change rate of stimuli, and the fluctuation of satiety. Succeeding the reinforcer-emotion association hypothesis that Rolls proposed in Ref. [35], the emotion model in *AROM* system introduces reinforcers and stimuli to help regulate variation of PAD emotions. The $S+$ means a positive stimulus, while $S-$ means a negative one. The symbol ! means the termination of such stimuli, e.g., $S+!$ means the termination of a positive stimulus. $S+-\downarrow$ means the intense of the stimulus is decreased, no matter the stimulus is positive or negative. Similarly, $S+-\downarrow$ means the intense of the stimulus is increased.

The levels of affect can thus be determined accordingly. Starting from PLEASURE, the degree of pleasantness is expected to decrease, if the robot is exposed to negative stimuli and increases, if the robot is exposed to positive stimuli. The dynamics of PLEASURE P can thus the modeled by

$$P_t = P_{t-1}e^{-k_p} + m_p\left(R_t - R_{t-1}\right), \tag{3.3}$$

where P is the degree of pleasantness, $P_{t-1}e^{-k_p t}$ is the damping of PLEASURE, m_p is the parameter adjusting sensitivity to stimuli, and R is the value of stimuli, here we refer to as *reward*, similar to the concept in Reinforcement Learning [46]. The determination of the level of PLEASURE only considers the temporal changes according to stimuli, so we only compare the reward of current and previous time.

The degree of AROUSAL (Noradrenaline) is influenced by the intensity of any sudden change on the stimulus. If the changes are always slow, then the level of AROUSAL will gradually decay, while, if the changes are large and quick, the AROUSAL will rise. Here, the level of AROUSAL A will be determined by the following relation,

$$A_t = A_{t-1}e^{-k_a} + m_a|R_t - R_{t-1}| \tag{3.4}$$

where A is the degree of AROUSAL, m_a is a parameter adjusting the sensitivity of AROUSAL dimension of the stimuli. The damping act in the same way in the dynamics of PLEASURE given by equation (3.3). Since the changes to AROUSAL is also generally sudden, so we only consider the absolute changing value of rewards between current and previous time step.

The dynamics of DOMINANCE is a simplified model of the serotonin system and is affected by the gap between expectation and reality. Here we use the termination of positive or negative stimuli to govern DOMINANCE. If the robot is always getting positive stimuli, but suddenly receives a negative stimulus, then the DOMINANCE decreases a little bit, expressing in a sense the expectation that things are starting to go wrong. Similarly, if at one time the robot receives a stimulus that is more positive than usual, the robot may feel more confident about this task. The update for DOMINANCE D is thus:

$$D_t = D_{t-1}e^{-k_d} + m_d\left(\frac{\mathrm{d}\left(R_t - R_{t-1}\right)}{\mathrm{d}t} - \frac{\mathrm{d}\left(R_{t-1} - R_{t-q}\right)}{\mathrm{d}t}\right) \tag{3.5}$$

where D is the degree of DOMINANCE, m_d is the parameter adjusting the sensitivity to stimuli, q is the parameter controlling how much historical information we want to count in to calculation. The damping way is still similar to the previous affect calculation. Since the variation of DOMINANCE is related to the changes of new stimulus from the old ones, here we compare the changing rate of reward between the current and last time as well as the past several times of changing rate of rewards. By adjusting parameter q, the robot can be set to have better or worse memory.

The calculation of DOPAMINE is in a system distinct from the aforementioned affect dimensions, and since the *AROM* model mainly utilizes the function of DOPAMINE to regulate motivation, here DOPAMINE is under influence of the level of various forms of satiety. If the robot has low satiety in respect to the battery level or social contact, then the DOPAMINE will increase and provide higher motivation for robot to reach a higher level of satiety and to obtain what it wants. If the robot is more satisfied with the present conditions, then the DOPAMINE level will decrease and the robot will be less inclined to achieve those goals. The calculation function for DOPAMINE D^A is:

$$D_t^A = D_{t-1}^A e^{-k_{da}} + \kappa S_t^A \tag{3.6}$$

where D^A is the level of DOPAMINE, with S^A representing the level of satiety. Inspired by research on incentive salience of DOPAMINE in the field of computational biology [53], the rate parameter κ represents the sensitivity of satiety in the present context.

3.3 Decision-Making Model

The decision-making process of the *AROM* model is influenced by both the interaction context information and the affect factors. As illustrated in Figure 3.1, the decision making module is composed of two procedures, one of which is calculating weight values, and the other one employs an MDP-like dynamics to determine decision probabilities. The computations in this module are based on the context and on historical information, and specifically the weight values are calculated based on emotional and cognitive WANTING factors. After decision-making, the decided action will impact the behaviour generation module, along with the calculated affect value.

For a biomimetic decision-making module, the goal is to enable the robot to decide reasonably and in a way similar to a human (or another animal). As a simple realisation of such a module, we refer to Markov Decision Process (MDP) which were employed as a basic model for human decision-making process in neuroscience [34] and psychology [41]. The key factors in this process are the reward (in the sense of Reinforcement Learning) and the probabilities of state transitions. Here, we the MDP model is embedded in the *AROM* model which interferes with the Markov property and the matrix-linearity of the state transitions. Because of its basic structure and its relevance in the theoretical study of decision making, we will still refer to it as an MDP even if this holds in a strict sense only for the isolated system.

We proceed using the original concept "reward" as a weight value, which represents how the robot value the states to reach, as a combination of rational judgement, its physiological "wanting", and affective influence to the states. The probability distribution, which is the probability of taking action for given current states to next states as originally declared in Reinforcement Learning, represents the "understanding" of the world by the robot, e.g., to what probability can the robot reach this state from the last state by taking this action?

Compared to the original MDP, the states in action prediction model are also modified, to represent the positive or negative features of consequences of actions. For instance, for a robot with action of going towards a human, the corresponding states are the positive results, e.g., being accepted by the human and playing with the human, as well as negative results, such as being refused by the human. By setting states with these features, it is easier to calculate the effect of stimuli as illustrated in Section 3.2, for instance, after some actions the robot reaches the state with a negative feature, then it can be assumed as receiving one kind of negative stimuli. The weight value here for each state can be seen as the degree that this state is positive or negative. As a result, the weight value can not only be utilized to engage in the decision-making process through MDP model, but also can be utilized to calculate the updated affects, as declared in Section 3.2.

3.3.1 Inspirations from Emotions and their Influence on Decisions

The emotions applied in the Decision-Making module are DOPAMINE, Serotonin DOMINANCE and Noradrenaline (AROUSAL). The functions of them are abstracted from the realistic functions of those monoamines on decision-making. The monoamine dopamine plays an integral role in coding of rewards in decision-making [48], and during decision-making, it controls the incentive salience function, which triggers a pulse of "wanting" towards the cues of rewards [53, 4]. We borrow the incentive salience as the feature of DOPAMINE in the *AROM* system, and utilize it to adjust the weights in MDP.

The monoamine serotonin functions to constrain the response of external stimuli and adjusts ongoing processing of sensory input [26]. Considering that the mechanism and influences of it are very complex, we only borrow parts of features of it overlapping with emotional dimension dominance, and apply DOMINANCE to adjusting action weights.

The monoamine noradrenaline, as far as researched, can cause to a reduced usage of decision-relevant information when choosing actions. There also exists evidence suggesting noradrenaline functions in randomly deciding, instead of directed exploration in exploration-exploitation dilemma [9, 30]. From here, we apply the features of neglecting relevant information and making random decisions to our model, enabling AROUSAL impact probability distribution in MDP.

3.3.2 Structure of the Decision-Making Process

The process of MDP decision-making sub-module is shown in Figure 3.3. Starting from calculating historical interaction information, the module recalculates affective weights and probabilities. After that, the affective probability weights are transferred to MDP, and utilized to predict optimal actions with emotions. The MDP uses Value Iteration Algorithm 1 to calculate the best action to take.

After the decision is made, the robot will conduct this decision and interact with the external environment, and receive the results of this interaction. The results will be matched to the states with positive or negative features, and then the robot will be assumed at this matched state, and use this state as the "last state" for the next MDP best action calculation. During this process, each time the robot obtains interaction results from outside, it will calculate them as historical information and cache them. From this information, the robot calculates the historical transition probability distribution, which is the expectation of robot whether specific actions can be successful or are likely to fail. This distribution will be introduced into the MDP model as the updated transition probability distribution for the next MDP calculation.

Algorithm 1 Value Iteration Algorithm [3, 46]

Initialize $V(s)$ to arbitrary values
Repeat until $V(s)$ converge
for all states **do**
  for all actions **do**
    $Q(s,a) \leftarrow \sum_s P^a_{ss'}(r(s,a) + \gamma V(s'))$
  end for
  $V(s) \leftarrow \max_a Q(s,a)$
end for

The calculation of historical transition probability distribution uses Thompson Sampling (TS) to avoid frustration caused by cold boot and significantly bad responses in a constantly changing system. TS is used to model the uncertainty or belief about the reward probability of each arm in multi-armed bandit problems, and also used as a heuristic solution of the exploitation-exploration dilemma [47, 32]. The beta distribution provides a way to represent this uncertainty in a probabilistic manner with

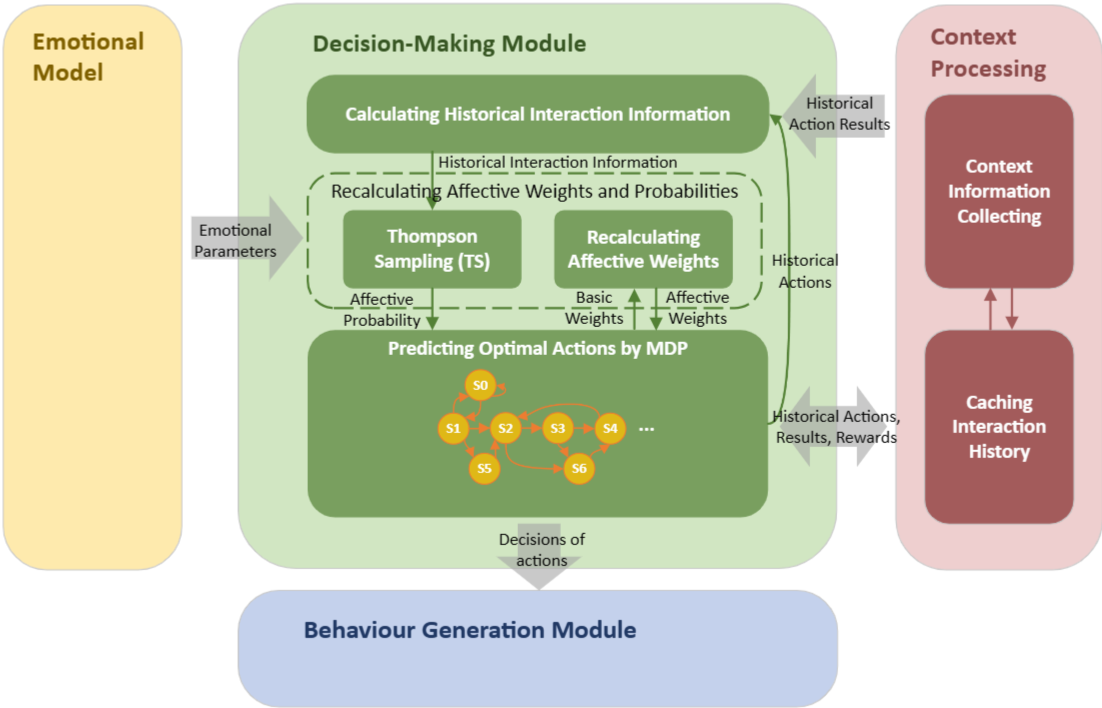


Figure 3.3: Decision-Making Process is based on an MDP as its core which is embedded in the decision-making module. This module receives inputs from the emotion module and is affected by context information. In this way, the weights of the MDP are dynamic and allow for flexible transition probabilities on longer time scales.

two parameters α and β. The more rewards obtained relative to selections, the higher the shape parameter α will be, reflecting a higher estimated reward probability. Here, TS is not directly applied and controls the robot to choose the action with the highest estimated reward probability. It calculates the *reward probability* as the transition probability in the MDP functioning as the robot's understanding of the world. As the robot has more probability to obtain a certain state S, the belief of get S at the next action is getting higher and more stable. But if the robot is continuously refused by human for example, and cannot get certain state S by taking action A, then TS will sway the belief higher to prompt the robot to keep trying, instead of directly give up. The algorithm of TS is shown in Algorithm 2.

3.3.3 Calculation of Intervening Affects

Based on the aforementioned MDP decision-making system, the affect factors are calculated and combined during the decision-making process. There are two main types of variables that influence how decisions are produced, which are the weight values

Algorithm 2 Thompson Sampling in *AROM* (K, α, β) [36]

for t=1,2,... **do**
  for k=1,..., K **do**
    Sample $\hat{\theta} \sim beta(\alpha_k, \beta_k)$
  end for
  $P_{x_{kt}} \leftarrow argmax_k \hat{\theta}_k$
  Apply $P_{x_{kt}}$ as transition probability in the MDP and observe r_t
  $(\alpha_{x_t}, \beta_{x_t}) \leftarrow (\alpha_{x_t} + r_t, \beta_{x_t} + 1 - r_t)$
end for

and transition probabilities. The weight values represent which states the robot considers as important, whether they "want" them and to what extend they need them. While the transition probability distribution, is the robot representation of the world, describing whether it is possible to reach a certain state S by conducting action A. Here, we show how the emotions function on these two variables, and thus influence the decision-making process.

The weight value of the MDP is mainly affected by the factors of DOPAMINE and Serotonin (DOMINANCE). One of the functions of DOPAMINE is controlling "wanting" parameter, and the larger "wanting" parameter is, the higher weight value to be set when considering making a decision. Another factor that intervened into decision-making is Serotonin (DOMINANCE). Just consider how Serotonin is used to treat depression: through increasing intensity of Serotonin, the slothful and "fragile" performance is improved, and the subjects become more energetic. We can thus utilize this feature and assign a higher weight value to states (except to IDLE), if with a larger Serotonin (DOMINANCE) factor, representing that robots are activated to make decisions and move around.

For the IDLE state, the weight value is of reverse relationship with Serotonin (DOMINANCE), to show a lazier preference of choice under low level of it. The calculation of weight values is illustrated as below:

$$W = \begin{cases} W_{bs} e^{\lambda D^A} e^{\mu D}, & \text{if } D < 0 \\ W_{bs} e^{\lambda D^A}, & \text{if } D \geq 0 \end{cases} \tag{3.7}$$

where W represents weight value for general states, W_{bs} represents the original basic weight value without influences of affects. D^A is the value of DOPAMINE, and λ represent the degree that DOPAMINE can influence the values of weight; D is the value of Serotonin (DOMINANCE), and similarly μ represents the degree of it to influence weights. The effects from Serotonin (DOMINANCE) D are only valid when D is negative, and if D is neutral or positive the weight keeps as default. For weights of IDLE particularly, the calculation is:

$$W_{\text{idle}} = \begin{cases} \frac{W_{\text{bs-idle}}}{\delta e^{D}}, & \text{if } \ D < 0 \\ \\ W_{\text{bs-idle}}, & \text{if } \ D \geq 0 \end{cases} \tag{3.8}$$

where W_{idle} represents weight value for the IDLE state specifically, and $W_{\text{bs-idle}}$ represents the original basic weight value for the IDLE state. D is the value of Serotonin (DOMINANCE), and δ is the level that Serotonin (DOMINANCE) can influence weights for the IDLE state. What should be emphasized is that here IDLE is not totally equal to sleeping or taking a rest, and we temporarily ignore the ”wanting” of robot to rest.

The transition probability distribution is mainly under the impact of Noradrenaline (AROUSAL). Noradrenaline, as far as researched, can influence rational and irrational decisions, and one reason that it causes this change is that it may slow the reaction to new changes of the environment, e.g., not sensitive to new information and making decisions according to old learned knowledge and natural ”wanting”. Thus we utilize this feature of Noradrenaline (AROUSAL) to modify the transition probability distribution in the MDP model, by adjusting the parameters of α and β in Thompson Sampling (TS) illustrated in Algorithm 2, as well as normalising the probability to relatively ignore the success probability and turn to targets that are more wanted. While Noradrenaline (AROUSAL) A exceeds a positive threshold value U, the calculation of transition probability are:

$$(\alpha_{x_t}, \beta_{x_t}) \leftarrow \left(\alpha_{x_t} + \frac{r_t}{A+1}, \beta_{x_t} + 1 - \frac{r_t}{A+1}\right) \tag{3.9}$$

$$P = P_{\text{TS}} - A\left(P_{\text{TS}} - \frac{1}{2}\right) \tag{3.10}$$

where $\alpha_{x_t}, \beta_{x_t}$ are the parameters from TS, r_t represents rewards in TS, A represents the value of Noradrenaline (AROUSAL), P represents the probability, P_{TS} is the basic probability, which is generated through TS, which can be influenced by affects or not based on the value of affect. Equation 3.9 illustrates that in Thompson Sampling (TS), with Noradrenaline (AROUSAL) A increasing, the rewards r_t added to beta function in TS will be decreased in updating, so as to make the robot less sensitive to past successes or failures. Equation 3.10 shows the normalisation of P from a range of $[0,1]$ towards $\frac{1}{2}$ (here $A > U$, and $U > 0, A \in [0,1]$), representing relevant ignorance of success probability. The higher A is, the closer the probability approaches 0.5; and if $A = 0$, then the probability keeps not change as the basic probability P_{TS}.

3.4 Proposed Behaviour Generating Process

Except for emotion emotion-intervening decision-making process, affective robots also require a module to generate interpretable and vivid behaviours. The importance of proper design of robot emotional expression in behaviours has been researched and admitted by many researchers. Tsiourti et al. argue that incongruence of robot emotional expressions will confuse human observers and result in impressions that the robot is unintelligent or not empathetic [49]. Proper designed robot emotional expressions can also help humans acquire friendly and kind impressions of robots. As proposed by Koschate et al., the displays of emotions can reduce the uncanniness of human-like robots [22].

Whereas, except for generating vivid and properly designed emotional behaviours, and giving people friendly impressions of robots, in this project we also target on establishing a general and adaptive robot behaviour-generating mechanism, which is designed facing to continuous emotion states, and can be easily adopted to various types of robots in different tasks. Through adjusting motion parameters according to emotional dynamics, this behaviour generating process of *AROM* provides a mapping and generating architecture for various emotions and robot behaviours, including large-scale trajectory movement, body postures and possible detailed distal motions. The generating model is inspired by physiological observations and empirical relationship between emotions and movements, and it aims to be compatible with human intuition and to reach a level of generality that enables comparisons between machines and humans.

3.4.1 Inspiration from Physiological Behaviours

The physiological behaviours are influenced by monoamines, especially by serotonin. It affects the muscular tension, and causes ”5-TH posture” when injected into animals, which is ”holding a stance elevated above the substrate with the abdomen flexed

Table 3.3: General influences of emotions to motions

	Influences (+)		**Influences** (−)	
PLEASURE	Behaviour inhibition	↓	Behaviour inhibition	↑
	swing angle	↑	swing angle	↓
AROUSAL	Gait velocity	↑	Gait velocity	↓
	Swing	↑	swing	↓
	Thigh angle	↑	Thigh angle	↓
	Step length	↑	Step length	↓
DOMINANCE	Default		Inhibiting reactions	

under its body" [1, 21]. Inspired by this we design the DOMINANCE to influence the STRETCH and SHRINK of robot body. Aside from this, we abstract emotional behaviour features from affective gait research [10, 20, 19]. Dynamic emotional parameters that participate in calculation are PLEASURE, AROUSAL (Noradrenaline) and DOMINANCE (Serotonin). The general functions of them towards generating robot behaviours are illustrated in Table 3.3. As the value of PLEASURE rises, it will impose an impact on robot behaviours to reduce the inhibition of movement, and lift the angle of swinging arms or other similar types of body. As PLEASURE value reduces, it will have a reversed influence. With the AROUSAL value increases, for situations where robots have legs, the gait velocity, swing movement, thigh angle, and step length will increase, and vice versa. For robots that do not have legs and move otherwise, the impact on aspects of legs will be altered correspondingly. At a higher value of DOMINANCE, we assume that the current robot movement is retained by default. However, at a lower value, DOMINANCE has the effect of inhibiting reactions of the robot both regarding speed and amplitude.

3.4.2 Calculation of Affect-Intervened Motion Parameters

Table 3.3 shows the simplified relationship model between emotional dynamics and robot behaviours. Based on this model, the affect-intervened motion parameters are calculated according to a definition involving three dimensions: Static posture, motion variability, and body movement. The motion parameters related to the dimension of static posture are STRETCH and SHRINK, which describe the direction of robot static postures, and STRETCH represents a direction outwards from the body, while SHRINK represents direction towards the inside.

For instance, for the head and neck part of postures, STRETCH means the robot raising its neck and chin up, while SHRINK refers to a downward movement of the head onto he chest (as if the robots intends to shrink its body). For the ears, for example, STRETCH means spreading the ears outwards, and SHRINK means turning ears inwards. The parameters of STRETCH and SHRINK compose the basic variables for controlling robot static postures, and they are combined to a united control parameter: SUPPRESSION. The parameter SUPPRESSION describes the static postures of movable joints analogous to the definition for STRETCH and SHRINK. In terms of robot movements, SUPPRESSION describes the medium degree of movement.

The eyelids of the robot, for example, may be static under some conditions, but can partially close under certain emotions to indicate a level of tiredness. However, if the robot is required to blink eyes, the eyelids will be dynamic and move upwards and downwards, so that the medium degree of movement appears as the static posture of eyelids, if they do not need to move, and this specific posture is defined by SUPPRESSION. Usually, according to the changing emotions the robot is going through, the SUPPRESSION parameter is usually not directly related to the mechanically neutral position of a posture. For instance, the lower value of PLEASURE may cause an increase of SUPPRESSION, which causes a deviation of the median positions from the neutral positions.

Under the BODY MOVEMENT dimension there are three motion parameters, which are VELOCITY, AMPLITUDE and Frequency respectively. Just as explained by their names, they described the velocity, amplitude and frequency of joints. For example, for swinging movements of the arm of the robot, the velocity of swing is controlled by VELOCITY, and the scale of swing is controlled by AMPLITUDE. If the swing movement continuously exists or happens regularly, then the frequency of the movement is controlled by FREQUENCY.

Expanded from Table 3.3, here in Table 3.4 the relationship between emotions and the motion parameters is illustrated. The emotion of PLEASURE influences the STRETCH, SHRINK, SUPPRESSION and AMPLITUDE. For positive values of PLEASURE (+), the STRETCH value increases, resulting the SUPPRESSION value decreasing, and increasing the AMPLITUDE value. For negative values of PLEASURE (−), the SHRINK value decreases, resulting in an increase of the SUPPRESSION value, while the AMPLITUDE value decreases. The AROUSAL value influences the VELOCITY, AMPLITUDE and FREQUENCY parameters. For positive values of AROUSAL (+), the VELOCITY, AMPLITUDE and FREQUENCY increase, while for negative values of AROUSAL (−), the three parameters decrease. The level of DOMINANCE affects the parameters of STRETCH, SUPPRESSION, VELOCITY and AMPLITUDE, but the influences are only effective when the DOMINANCE value is negative (−). The negative DOMINANCE values will cause the SHRINK to decrease, thus leading to an increase of SUPPRESSION. Finally, if VELOCITY decreases, then also AMPLITUDE decreases.

Table 3.4: Framework of generating emotional robot behaviours

Behaviour		**Static Posture**		**Deviation**	**Body Movement**		
		STRETCH	SHRINK	SUPPRESSION	VELOCITY	AMPLITUDE	FREQUENCY
PLEASURE	+	↑		↓		↑	
	−		↓	↑		↓	
AROUSAL	+				↑	↑	↑
	−				↓	↓	↓
DOMINANCE	+						
	−		↓	↑	↓	↓	

Based on Table 3.4, motion parameters are calculated according to emotional parameters of PLEASURE, AROUSAL and DOMINANCE. The effects of those emotional parameters are assumed that: PLEASURE has linear effects on behaviour parameters; AROUSAL has stronger influences at larger values, and the effects are nonlinear; only when DOMINANCE is in its negative range it has influences on behaviours, and the influences are in a form of nonlinear damping. Under such kind of hypothesis, the motion parameters are calculated as follows:

$$\text{Suppression} = \begin{cases} H_{Sp}\left(\frac{P+1}{2}\right) \times H_{Sd} e^{v_{Sd} D}, & \text{if } D < 0 \\ H_{Sp}\left(\frac{P+1}{2}\right), & \text{if } D \geq 0 \end{cases} \tag{3.11}$$

$$\text{Velocity} = \begin{cases} H_{Va} e^{v_{Va}(A-1)} \times H_{Vd} e^{v_{Vd} D}, & \text{if } D < 0 \\ \\ H_{Va} e^{v_{Va}(A-1)}, & \text{if } D \geq 0 \end{cases} \tag{3.12}$$

$$\text{Amplitude} = \begin{cases} H_{Ap} \left(\frac{P+1}{2}\right) \times H_{Aa} e^{v_{Aa}(A-1)} \times H_{Ad} e^{v_{Ad} D}, & \text{if } D < 0 \\ \\ H_{Ap} \left(\frac{P+1}{2}\right) \times H_{Aa} e^{v_{Aa}(A-1)}, & \text{if } D \geq 0 \end{cases} \tag{3.13}$$

$$\text{Frequency} = H_{Fa} e^{v_{Fa}(A-1)}, \text{ for all } D \tag{3.14}$$

where in Equation 3.11, H_{Sp} is the degree at which PLEASURE P influences suppression, with H_{Sd} representing the degree of influence from DOMINANCE D. v_{Sd} is the damping constant of DOMINANCE D regarding to suppression.

Equations 3.12, 3.13, and 3.14 are understood in the same way, where the first uppercase subscript characters of H and v represent the motion parameters, V for velocity, A for amplitude and F for frequency. The second lowercase subscript characters describe the emotional dimensions the effects from, with p for PLEASURE, a for AROUSAL and d for DOMINANCE. The reason for applying $\frac{P+1}{2}$ and e^{A-1} is to regularize the ranges of calculated motion parameters, which are restricted to [0, 1], considering that the range of the quantities P, A, D is [-1, 1].

3.4.3 Behaviour Generating Algorithm

The Behaviour Generating algorithm of *AROM* utilizes the motion parameters calculated from emotions to let the robot produce affective behaviours, which can be combined with and controlled by decision-making module, being able to adapt to distinctive types of behaviours. The inputs of the algorithm are parameters of robot systems like types of joints, scales of movement, neutral positions of each joint or movable units. After calculated with the real-time motion parameters which contain emotional information, they transfer to control parameters of robots, and finally generate affective robot behaviours. The generated behaviours mainly belong to two aspects: the one is the large-scale movement trajectory, without considering the body postures or gait patterns, and the other one is the robot behaviours on the level of body part, like thigh elevation and arm swing. The macroscopic trajectory is controlled by the motion parameters of velocity, amplitude and frequency, and the body behaviours are controlled by all the motion parameters, if necessary.

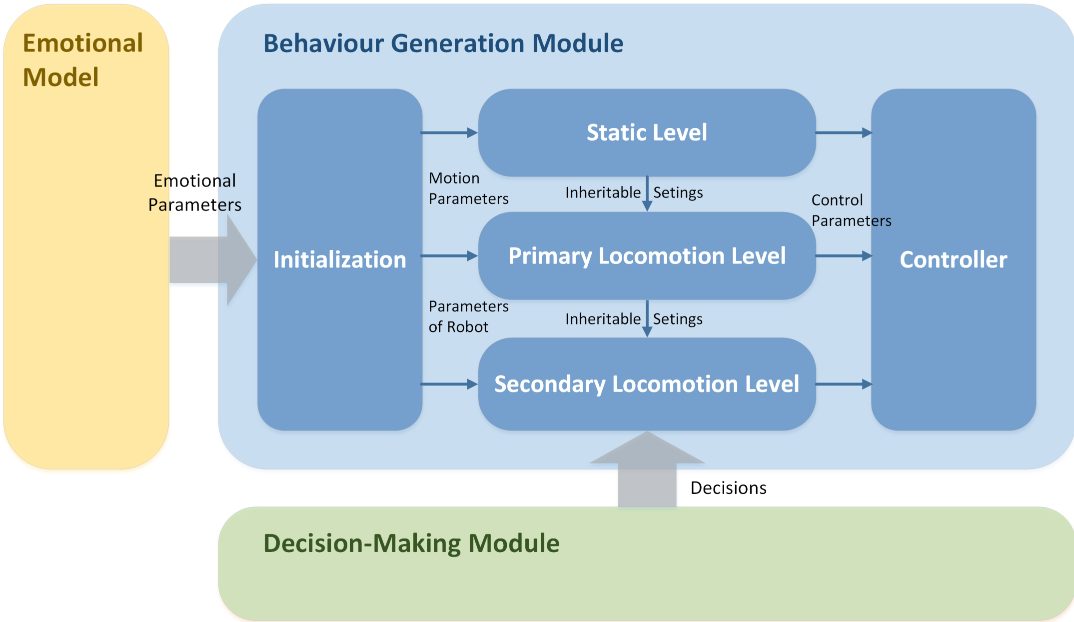


Figure 3.4: Structure of the the behaviour-generating process

As shown in Figure 3.4, the behaviour-generating module starts from an initialisation phase, and then calculates robot behaviour across several level, from the Static Level, the Primary Locomotion Level to the Secondary Locomotion Level, and finally sends the target behaviour parameters to the controller of the robot. At the initialisation stage, several inputs including emotional parameters and the basic parameters of the robot movement are received and processed, with movement parameters being affected also by emotion parameters. Also the basic movement parameters enter the process of the generation of behaviour.

In more detail, the three level of behaviour generation are as follows:

- *Static level*: to generate static postures, which change all the time with emotions
- *Primary locomotion level*: Generating movement parts that are not very significant for robots to complete tasks or so, but still have meanings such as expressing emotions. The robot components at this level will start to move or stop depending on triggers of specific strong emotions etc., and can be understood as almost fully controlled by behaviour-generating module.
- *Secondary locomotion level*: Generating important movements that are mainly required by decision-making module, of which some movement parameters may change according to emotions, but the joint or body parts on this level will keep moving or stop under control based on decision making.

While generating behaviours, these three levels of behaviour generation inherit settings from their previous level, and make use movement features that were calculated already. For instance, for a movement belonging to the Primary Locomotion level, a requirement may exist for the robot to remain static for a certain interval of time, in which case the calculated Static-Level behaviours can be directly applied and be sent with the movement generation output of the primary level to the controller. The analogous effect can be seen at the Secondary Locomotion Level, where the generation behaviour generation can refer to the results of the Decision-Making module. In intervals between two decisions, where such results are unavailable, the behaviours of the primary level can be applied.

Based on this structure for behaviour generating, the algorithm is given as pseudocode in Algorithm 3. As outlined in the algorithm, the inputs contain emotional parameters, neutral positions and movable scales of the joints of the robot, neutral velocity and velocity scale calculated from the max velocity of the robot, a basic continuous movement pattern function such as the trigonometric function for eyelid opening, tail wagging and so on, and the decisions made by the Decision-Making module. Through calculation, the algorithm outputs the parameters and functions for behaviour control, both for the movement trajectory of the robot and for movements of body appendages. On the Static Level, the joint positions are adjusted according to motion parameter SUPPRESSION. When a requirement for Primary Locomotion Level is met, which is here usually the presence of certain emotions, the robot start to conduct locomotion. The motion parameter FREQUENCY adjusts how frequently the robot conducts the locomotion, and the body movement amplitude is modified by the AMPLITUDE parameter, with VELOCITY modifying the linear velocity and AMPLITUDE any angular velocities. It should be noted that here the velocities do not directly relate to the velocities of the movements of body parts. Instead, they are defined at the level of full-robot trajectory, although for complex robots sometimes the velocity needs to be inferred from the velocities of robot body parts, according to the way the robot moves. For the Secondary Locomotion Level, the behaviour generation is based on the framework of the primary level, combined with decisions, which may assign the target movement directions for robots, depending on what decision was taken.

The algorithm is adaptive to variable task types. The movement parameters and patterns for each joint can be designed specifically, and at the same time still show affective features during the movements of the robot. For special robot body parts that not suit the framework of this algorithm, the appliers also can pick appropriate motion parameters of Suppression, Velocity, Amplitude and Frequency and design their own unique affective behaviours.

Algorithm 3 Behaviour-Generating Algorithm

Input: emotional parameters P, A, D, neutral positions and movable scales N_j, S_j for jth movable joints of the robot, neutral linear and angular velocity N_l, N_a and their scales S_l, S_a, basic continuous movement pattern function $T_j(t)$ at time t, Decision H

Output: static affective behaviour parameters B_j and movement function $B_j(t)$ for jth movable joints of the robot, linear and angular velocity v_l, v_a

Initialize motion parameters Sup, Vel, Amp, $Freq$ from P, A, D, requirement K for Primary Locomotion Level

for all j **do**
  $B_j \leftarrow N_j - 2(Sup - 0.5)S_j$ {Static Level}
end for
if K is met **then**
  for t per $\frac{1}{Freq}$ **do**
    for all j **do**
      $B_j(t) \leftarrow T_j(t) \times Amp$ {Primary locomotion level}
    end for
    $v_l \leftarrow N_l + 2(Vel - 0.5)S_l$
    $v_a \leftarrow N_a + 2(Amp - 0.5)S_a$
    if H exists **then**
      for all j **do**
        Update $B_j(t)$ based on $T_j(t) \times Amp$ and H {Secondary locomotion level}
      end for
      Update v_l based on $N_l + 2(Vel - 0.5)S_l$ and H
      Update v_a based on $N_a + 2(Amp - 0.5)S_a$ and H
    end if
  end for
end if

Chapter 4

Experiments

4.1 Ethics Statement

This study was conducted according to relevant University of Edinburgh, national and international guidelines.

4.2 Decision-Making Module and Emotion Model During Interaction

4.2.1 Experiment Design

As many physiologist proposed, physiology is about how the living organism adjusts to the adversities of the environment, and enable them to adapt to the changing world [39, 40]. The organism escapes from dying cold or heat, moves to find suitable surroundings, water and food. They are quickly aroused if meeting prayers, and calm down at their safe spaces. Inspired by such features, the Decision-Making module is designed. Thus in the experiment, we test the adaptivity of the module to the changing environment, and the benefits of emotions in robot decision-making in the scenarios of human-Robot Interaction.

The interaction scenario is shown in Figure 4.1, abstracting the three places that the robot can choose to locate itself: A position near the human, a playground that the robot likes to explore, and a home position for resting. The robot has thus to decide among three options: Playing with the human, playing alone at the playground, and moving to the home position to rest. When the robot approaches the human, the human can reject the robot or accept it and engage in playful interaction between the robot and the

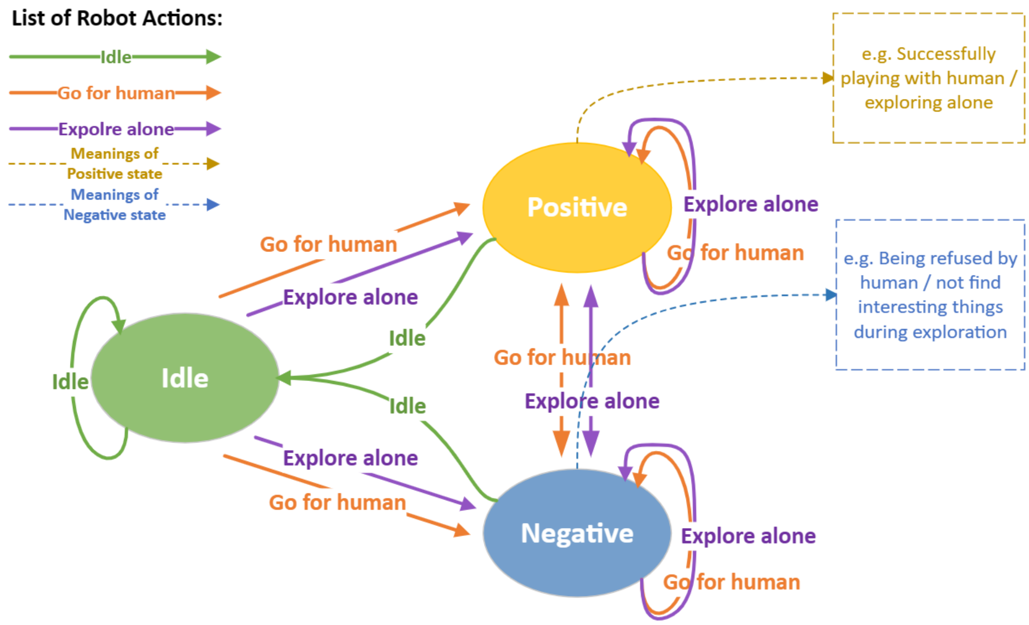


Figure 4.1: MDP structure in the experiment.

human. Illustrated in the form of an MDP, the robot can conduct three types of actions, which are IDLE, GO for human and EXPLORE alone. The states in the MDP describe the interaction states in an abstract way, which benefits the calculation and update of emotional dynamics. As shown in Figure 4.1, the IDLE state represents the state of the robot staying at the home position and rest. Whereas, the Positive and Negative states reflect the types of influences that the resulting states of actions on robot emotions, and Positive represents successful results, while Negative can be seen as a failure. For instance, the Positive state means successfully playing with humans or successfully exploring alone, and the Negative state means being refused by human or cannot find interesting things while exploring. In the MDP, the rewards, also mentioned as basic weights in this project, that lead to the Positive state is positive, and negative for the Negative state. Both GO for human action and EXPLORE alone action can lead to these two states, depending on the results of actions. But only the IDLE action can lead to the IDLE state.

In the experiment, we compare two models: One is our designed decision-making model with emotion dynamics, and the other is based on our decision-making framework, but without the influences of emotional factors. In order to enable a direct comparison between the two models, we posit that in default situation, the robot has the same scale values of WANTING, i.e. the sensitivity level of DOPAMINE to Satiety, towards states of playing with the human and playing alone at playground. The sensitivity of changes of DOMINANCE to stimuli, as well as the rewards for achieving these two states are also equal. For the scenario without intervening emotions, there are no "wanting" parameters, so we only set the rewards of the two aforementioned states to the same values. The two experiments are based on the framework of behaviour generation we proposed. Therefore, for the scenario without emotions, the results of interacting will influence the belief of the robot on the probability of success. While for the scenario with emotions, the interaction results will not only influence the belief of the robot regarding the probability of success from an aspect of interact history in reality, the accordingly variant emotions will also have influences on the probability as well as the weights (similar to rewards), illustrating the changes of the understanding to the world by the robot, and to what degree those states (e.g., being able to play with the human) are important to it.

4.2.2 Implementation Procedure

To evaluate the model performances facing a changing environment, the situation we simulate in the experiment is:

The human participant at first shows much interest to the robot, and continues interacting and playing with the robot for a while. But after that, maybe the human has some work to do, so the human refuses to play with the robot, and shows no interest in interacting with the robot.

The Decision-Making module will take decisions 60 times during each episode, and for each experiment group, we run 50 episodes to achieve average emotional values and probabilities of action conducted. In each episode, the human accepts the robot and makes the robot plays successfully with the human before the 15th state, as long as the robot conducts the GO for human action. After the 15th state, the human always refuses the robot if it turns to human. Under such situation settings, the experiment groups are divided according to combinations of emotional dynamics, as well as distinctive sensitivity of DOPAMINE and DOMINANCE, illustrated as follows:

- *Group 1*: Decision-Making model without influences of emotional factors
- *Group 2*: Decision-Making model with only D (DOMINANCE) and D^A (DOPAMINE), with equal sensitivities of D and D^A to states

- *Group 3*: Decision-Making model with only A (AROUSAL) and D^A (DOPAMINE), with equal sensitivities
- *Group 4*: Decision-Making model with only D (DOMINANCE) and A (AROUSAL), with equal sensitivities
- *Group 5*: Decision-Making model with all of the emotional factors functioning in model, which are A (AROUSAL), D (DOMINANCE), D^A (DOPAMINE), with equal sensitivities of D and D^A to states
- *Group 6*: Decision-Making model with all of the emotional factors, and the sensitivity of D^A (DOPAMINE) is higher towards human and lower towards exploration. This can be understood as easily having a stronger "wanting" to play with human, instead of exploring alone.
- *Group 7*: Decision-Making model with all of the emotional factors, and the sensitivity of D^A (DOPAMINE) is lower towards human and higher towards exploration.
- *Group 8*: Decision-Making model with all of the emotional factors, and the sensitivity of D (DOMINANCE) is higher towards human and lower towards exploration. This can be understood as the DOMINANCE to successfully playing with human is more easily to be influenced by what has happened, while the DOMINANCE to exploring is harder to influence.
- *Group 9*: Decision-Making model with all of the emotional factors, and the sensitivity of D (DOMINANCE) is lower towards human and higher towards exploration.

Apart from the emotional factor A (AROUSAL), D (DOMINANCE), D^A (DOPAMINE), the emotion P (PLEASURE) has no impact on robot decision-making, it is also calculated and updated according to changes of stimuli. As a result, it is also illustrated in the diagrams functioning as a reference to the mood of the robot.

4.2.3 Results

The results of Decision-Making model without emotional factors are shown in Figure 4.2. Though not modified by emotions, the model used in Group 1 also utilizes Thompson Sampling (TS) to regulate decisions and optimize obtained rewards. From the graph, it is clear to see the trend that the robot reduces the probability to taking action GO for human after 15th state, when the human stops to interact with it, and increases actions of EXPLORE alone and IDLE.

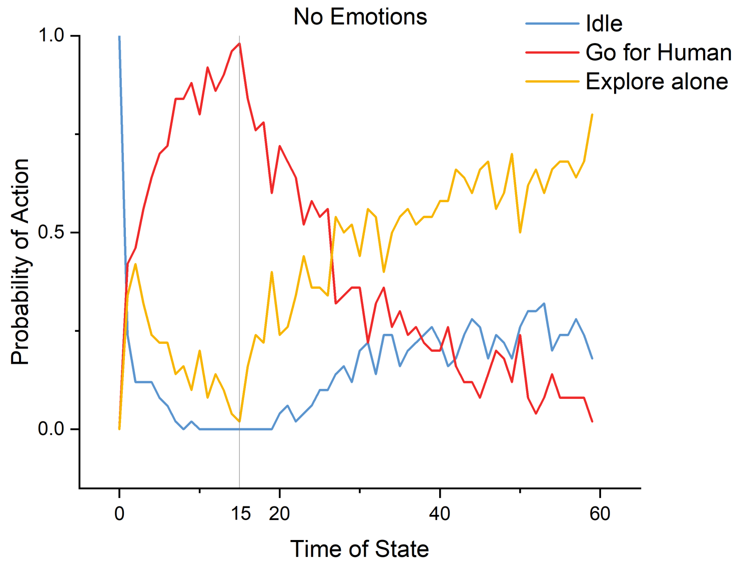
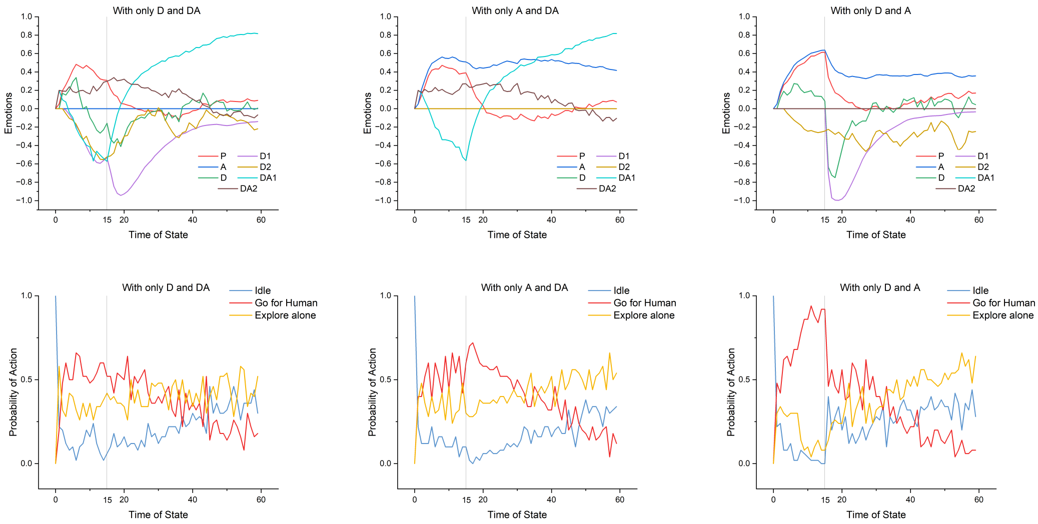


Figure 4.2: Probability of taking actions in Group 1

For the Decision-Making model with emotional factors, we assess the realistic effects of emotions when the robot faces a change of environment. Theoretically, the function of A (AROUSAL) is to reduce rational judgement according to interaction history, so the behaviours of robots will not converge very fast if A is large. Only A influences robots' understanding of success probability. The function of D (DOMINANCE) is to increase their understanding weights to staying idle (rest/home) and reduce weights to move, no matter towards human or explore alone, if D is low. D has three types, the total D, D_1 (D for playing with human), and D_2 (D for exploring alone). They follow the same generate method, with different types of stimulus values took into calculation. For example, D_1 only responds to stimuli related to humans, like being accepted by the human or being refused. The function of D^A (DOPAMINE) is to regulate robots' behaviours, and prevent robots always choosing one actions, aiming at showing some more natural and vivid robot behaviours. D^A influences robots' understanding to weights (rewards). When D^A is large and the robot has stronger wanting, the weights will increase and vice versa. D^A has two types, D^A_1 (for playing with human) and D^A_2 (for exploring alone). The changes of D^A depend on a type of variables named Satiety. Satiety is set to -1 if the robot failed to get positive rewards from the actions, no matter because the robot did not seek for it or being refused. Once the robot gets the positive rewards, (of course after seeking for it), the Satiety will $+2$. Thus, a loop of robot behaviours is built based on Satiety, to get rewards for one time and then take rest for 2 time steps unless influenced by any other factors.

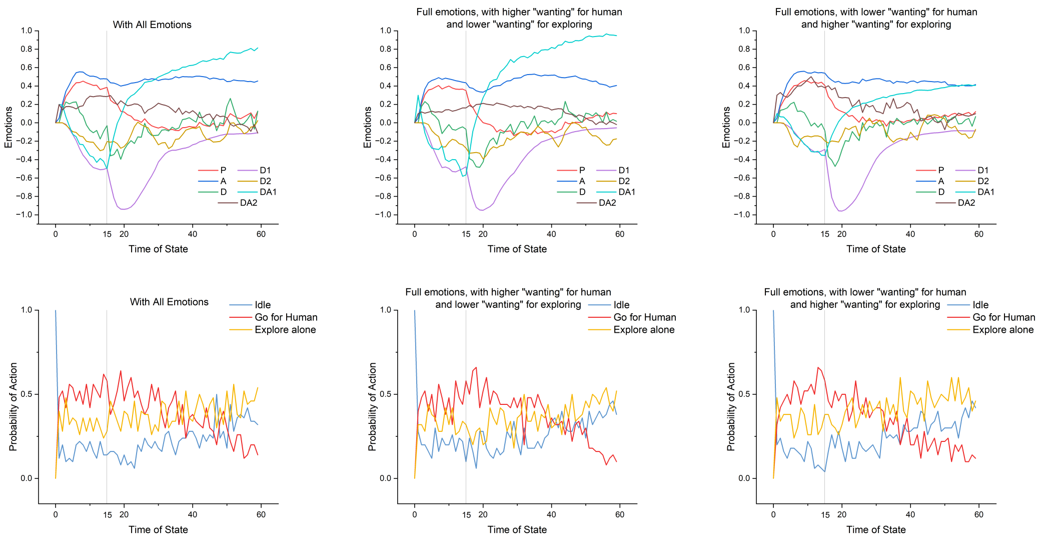
Referring to these inbuilt features of emotions in the model, we compare the emotionless model (Group 1) with the model with all of the emotions (Group 5) first. As shown in Figure 4.4, the trends of action probability in Group 5 is more flat than Group 1, reflecting a feature of stability of the emotional decision-making model. Yet, the emotionless model may seems more rational and can optimize obtained rewards quickly, the affective model provides a pattern where the robot always try to engage in interactions, without giving up once the environment changes negatively. To evaluate the impacts of single emotional factors, we compare the Group 2 to Group 4 to fully applied emotional model (Group 5). As shown in Figure 4.3, scales of action probability at model without A (AROUSAL) is broader than the model with AROUSAL. It can be understood as that the Group 5 is more irrational than Group 2, which makes impulsive decisions because of AROUSAL. For the model without D (DOMINANCE), the Group 3 has a higher peak value of action GO for human and lower peak value for action IDLE compared with the fully applied emotional model. In the meantime, the scales of Group 3 is also broader than Group 5. This illustrates that with DOMINANCE, the robot will be more conservative to try uncertain choices, and spend more time to rest, if the DOMINANCE level decreases and loses confidence. For the model without D^A (DOPAMINE), we can see that in Group 4, the front section of the figure of action probability is almost the same as the emotionless model, which converges quickly to the optimal choices before 15th state. But after that, the probability of action GO for human goes through a sudden and quick drop, with a significant increase of the probability of action IDLE occurring at the same time. Group 4 possesses a "sentimental" feature while making decisions, which reacts strongly when facing unexpected failures, significantly reducing choices related to failure and staying idle. There is also a stage of decrease of idle from state 16 to 24 approximately, which can be seen as the expression of frustration and ends with the promotion of DOMINANCE. Thus the function of DOPAMINE can be described here as regulating the decisions of the robot, and preventing endless seeking of one type of reward.

Aside from testing models with only two functional emotional factors, we also explore the various influences of sensitivities of D^A (DOPAMINE) and D (DOMINANCE) towards Satiety and Stimuli in Group 6 to Group 9. As illustrated in Figure 4.4, we



Group 2 Group 3 Group 4

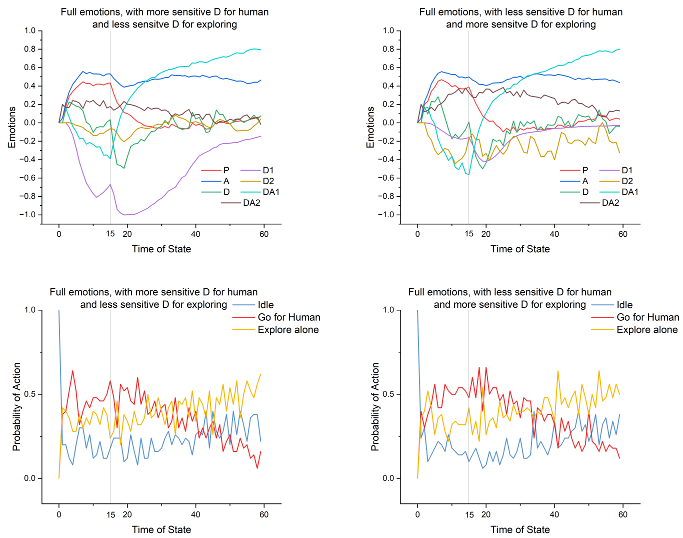
Figure 4.3: Average emotional values and probability of taking actions for models only apply two functional emotional factors



Group 5 Group 6 Group 7

Figure 4.4: Average emotional values and probability of taking actions for models with full emotions and different sensitivity of D^A (DOPAMINE)

apply higher sensitivity of D^A (DOPAMINE) towards playing with human, and less sensitivity towards exploring in Group 6, with reversed ones in Group 7. Comparing the action probability of Group 6 and 7, the peaks of action GO for human and IDLE in Group 6 occur later than Group 7, happening after the 15th state, while the peaks of Group 7 occur almost around the 15th state. This reflects a short-term insistence of going for human after being refused, when the DOPAMINE, or the ”wanting” towards human is highly sensitive. The holistic probability of EXPLORE alone in Group 7 is also slightly higher than that of Group 6, reflecting the increase of taking such action under higher sensitivity of DOPAMINE towards it. The values of DOPAMINE towards human (D_1^A) and DOPAMINE towards exploration (D_2^A) are also shown in the upper line of the pictures in Figure 4.4, illustrating an obvious increase of D_1^A in Group 6, caused by the combination of unsatisfying requirements and the highly sensitive feature of D_1^A.



Group 8 Group 9

Figure 4.5: Average emotional values and probability of taking actions for models with full emotions and different sensitivity of D (DOMINANCE)

In Figure 4.5, the models with different sensitivity of D (DOMINANCE) towards human and exploration are compared. In Group 8, the sensitivity of DOMINANCE to Stimuli towards human is higher, and the one towards exploration is lower, which are reversed in Group 9. The probability of action GO for human in Group 8 achieves a minor peak at around state 3, and after the 15th state, it levels off for around 10 states. While in Group 9, the probability of action Exploring alone achieves the minor peak at state 3, and the probability of action GO for human reaches a peak shortly after the 15th state, different from Group 8. The higher sensitivity of DOMINANCE makes the robot confident to choose choices when there are not much historical information can be referred to, which leads to the minor peak in Group 8 and 9 on action GO for human and EXPLORE alone. Whereas, when encountering failures, the models with higher DOMINANCE sensitivity are more easily to be frustrated, causing the levelling off stage of GO for human in Group 8.

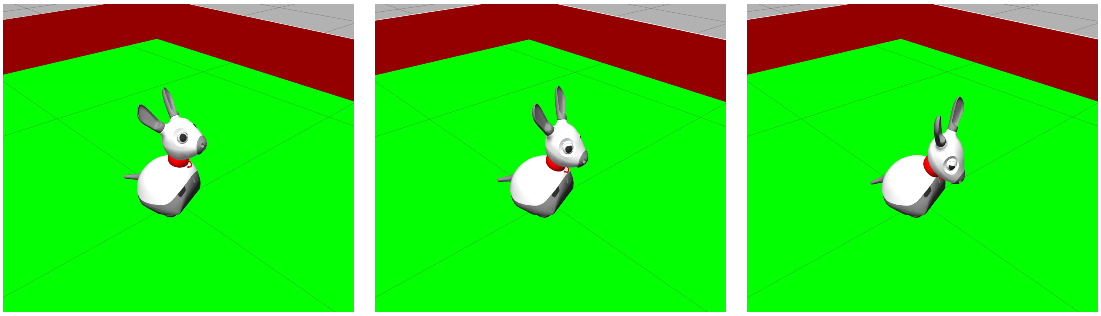
4.3 Behaviour-Generating Module

4.3.1 Experimental Design

In the experiment of Behaviour-Generating module, we evaluate the interpretability of robot generated affective behaviours for human, through asking participants watch videos of robot behaviours under different emotions. The scenario is demonstrated based on simulation, and the robot we use is MiRo, a bunny-like robot [7], which can move head, neck, ears and tail, although not in a complex way. Based on the principle rules of Behaviour-Generating module in Section 3.4, we simulate the scenario using ROS and Gazebo. In the experiment, he robots are programmed to move in a small piece of ground, and the standard behaviour for neutral emotions is circling slowly. Based on circling, the fundamental behaviour, the behaviours of the robot will vary and extend from circling according to their emotional dynamics. Since the functional emotional parameters in Behaviour-Generating part are only PLEASURE (P), AROUSAL (A) and DOMINANCE (D), the experiment groups are divided according to distinctive values of these three dimensions of emotion. There are three possible states for each dimension of emotional state in this experiment, which are positive state (emotional parameter value = 0.6), neutral state (emotional parameter value = 0), negative state (emotional parameter value = -0.6). For reference, the range of P, A, and D are all $[-1,1]$, with 1 standing for very pleased, aroused or confident, and -1 for very upset, sleepy, and lack of confidence. By combining these states, the experiment groups share the emotional values in 15 different combinations,

$$
(P,A,D) = \left\{
\begin{array}{rrrrrrrrr}
(0, & 0, & 0) & (0.6, & 0.6, & 0.6) & (-0.6, & -0.6, & -0.6) \\
(0.6, & 0.6, & 0) & (0.6, & 0, & 0.6) & (0, & 0.6, & 0.6) \\
(-0.6, & -0.6, & 0) & (-0.6, & 0, & -0.6) & (0, & -0.6, & -0.6) \\
(0.6, & 0, & 0) & (0, & 0.6, & 0) & (0, & 0, & 0.6) \\
(-0.6, & 0, & 0) & (0, & -0.6, & 0) & (0, & 0, & -0.6)
\end{array}
\right.
$$

so that we have 15 groups in total. A set of examples of static affective behaviours with such kind of emotional value sets are shown in Figure 4.6.



(a) P, A, $D = (0.6, 0.6, 0.6)$ (b) P, A, $D = (0, 0, 0)$ (c) P, A, $D = (-0.6, -0.6, -0.6)$

Figure 4.6: Affective behaviours of the MiRo robot.

During the experiment, the human participants were required to watch 15 videos based on the emotional sets aforementioned. The sequences of the videos are randomized, and the participants can watch the videos for multiple times if needed. After watching each video, the participants will be asked to rate their impressions on the robot behaviours shown in the video in a questionnaire, from the aspects of PLEASURE, AROUSAL and DOMINANCE. In the design of the questionnaire, we apply the questionnaire Table 4.1, where the description of the three emotional states refers to the study [23] which measures human emotions on aspect of P, A, D.

Table 4.1: Questionnaire

Please rate your impressions of the robot behaviours on these scales:

Unhappy	1	2	3	4	5	Happy
Sleepy	1	2	3	4	5	Wide-awake
Controlled	1	2	3	4	5	Controlling

4.3.2 Implementation Procedures

We recruited 21 people (8 Male, 11 Female, 1 Other, 1 Unwilling to say) speaking English to participate the experiment. Their knowledge and degree of familiarity to robots are that 38.10% of them know robots very well; 33.33% have some knowledge of robots; 23.81% know a little about robots; 4.76% barely know nothing about robots. The experiment is conducted in the form of an online survey, where participants first finish basic information, and then watch the 15 videos in a randomized sequence. During or after watching each video, the participants rate their impressions according to Table 4.1 immediately, and then enter the next video.

4.3.3 Results

Table 4.2: Pearson correlation coefficients and p-values of human emotional impressions in the experiments

Emotions	Pearson Corr.	p-value
P	0.93	<**0.0001*****
A	0.69	**0.004****
D	0.58	**0.022***

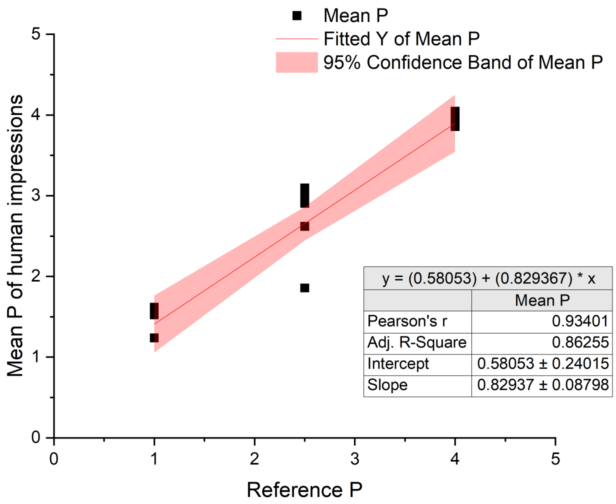
*p<0.05, **p<0.01, ***p<0.001

Bold value indicates statistically significant results

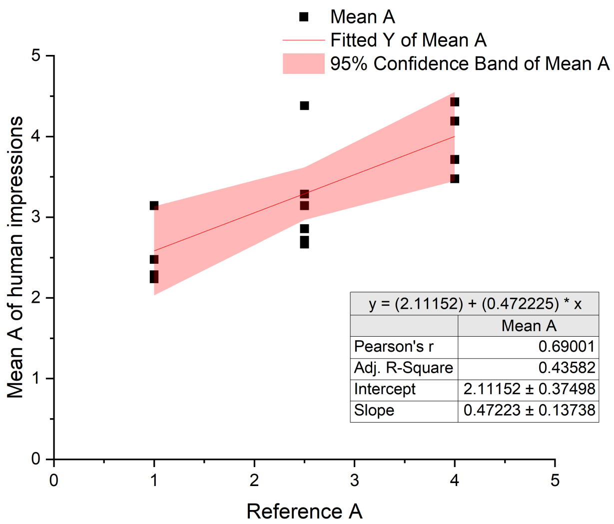
The results of each experiment group are processed and compared with the reference emotion values, i.e. the emotional values used to generate behaviours. Since the results have the scale of $(0,5]$, the reference emotion values are mapped from $[-1,1]$ to the same scale, where the value of 0.6 is mapped as 4, 0 is mapped as 2.5, and -0.6 is mapped as 1. The results of impression rating are also processed, utilising the mean values of each group to calculate the correlation. The Pearson correlation coefficients and p-values of human emotional impressions in the experiments and reference emotion values are illustrated in Table 4.2. We can see that the emotional dimension of PLEASURE matches the reference values best, with a Pearson correlation coefficient as 0.93, followed by AROUSAL and DOMINANCE, with Pearson correlation coefficients as 0.69 and 0.58 respectively. The p-values also share the same trend, with PLEASURE having the smallest p-value as < 0.0001, followed by AROUSAL and DOMINANCE. The p-values of which are 0.004 and 0.022 respectively. The results show that human participants are very good at distinguishing the behaviours of the robot in regard to the PLEASURE dimension, and the results are highly accurate, using values in the generating system as references. The second best emotional dimension that human can recognize is AROUSAL, and the worst one is DOMINANCE. More information of the impression rating results are illustrated in Figure 4.7.

In Figure 4.7, the mean values of human impressions of each experiment group is illustrated as black dots. The 95% confidence band are also shown in the pictures. We can see that the data distribution of PLEASURE is balanced and fit linearly well, meaning that people distinguish PLEASURE accurately. However, people usually overestimate the AROUSAL value of robot behaviours, and there exists an upward shift of the data distribution in the second graph. In the third graph, the distribution of DOMINANCE is narrow on the direction of y-axis, meaning that people are relatively not well at distinguishing behaviours on the aspect of DOMINANCE, compared with that of PLEASURE and AROUSAL.

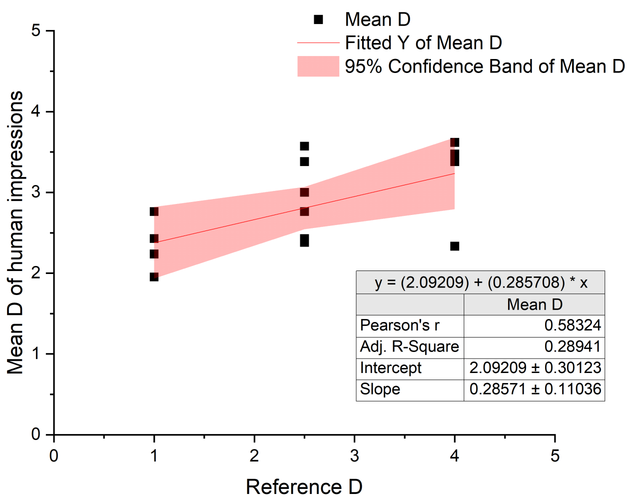
There may also exist mutual influences between emotional dimensions. For example, when people observe a robot behaviour with both high values of PLEASURE and AROUSAL, they may tend to overestimate the emotional value on the aspect of DOMINANCE. In the first graph of Figure 4.7, we observe the data dot which is obviously downward deviate from the band. It is the experiment group with reference values of $(P, A, D) = (0, -0.6, -0.6)$. In this group, people underestimate the value of PLEASURE, possibly being influenced by the joint low levels of AROUSAL and DOMINANCE. However, the PLEASURE value does not deviate in the group with values of $(P, A, D) = (0, 0.6, 0.6)$, so we assume that this influences of PLEASURE only happens when both values of AROUSAL and DOMINANCE are low. In the second graph, there is also a significantly deviated dot upward from the band, which is the group of $(P, A, D) = (0.6, 0, 0.6)$. Similarly, the AROUSAL value does not deviate much in the group of $(P, A, D) = (-0.6, 0, -0.6)$, so we deduct that the PLEASURE and DOMINANCE have a strong joint influence on AROUSAL when the values of P and D are large. Whereas, in the third graph, the pattern of influences to DOMINANCE is different. It seems that on the positive level, a large PLEASURE value itself can strongly influence the DOMINANCE value, given a neutral AROUSAL value. Moreover, large AROUSAL value with the neutral value of PLEASURE also have a sight influence of DOMINANCE. The two dots that most significantly deviate from the band upwards and downwards are from groups of $(P, A, D) = (0.6, 0, 0)$ and $(0, 0, 0.6)$. With the neutral AROUSAL, the single PLEASURE impacts people overestimate D value if high, and underestimate D value if



Impressions of *P*



Impressions of *A*



Impressions of *D*

Figure 4.7: Mean emotional impressions values of robot behaviours and the reference emotional values

low. The groups of (P, A, D) = (0, 0.6, 0) and (0, 0, 0.6) also show a similar impact from AROUSAL to DOMINANCE in the picture.

Chapter 5

Discussion

5.1 Affective Robotic Minds

Much work has been conducted on emotion recognition by robots as a pattern-matching problem or the expression of emotions by robots in human-robot interaction, see Reference [43], whereas the integration of such abilities in the behavioural context and its interpretability has found less attention. For integrated systems based on feedback loops, the cited paper refers only to the work of Dietrich Dörner, known as Psi-theory [12], and also to Ref. [52] who agree that "the ways in which humans and robots work together as a team" deserve more interest. The present study is more an attempt to show the complexity of the problem than a solution, but it also emphasises the possibility of separating the conceptual problem from the variety of useful developments in sensing, perception, planning, organisation of behaviour, control, learning, and evaluation. This is because the central organisation of a companion robot appears to be less easily realisable by utility-based machine learning approach [54], while a system-based affective architecture is not only natural, but tends also to be preferable by human users.

From the results of Decision-Making module, we've already been able to see how complex decisions can be by just adjusting several parameters in the emotion model. Personalities of social robots may have a more simple solution existing in this study. Through attributing variant values to the sensitivity parameters controlling influences of stimuli and satieties on emotional states, decision patterns like impulsive, conservative, sentimental and special preference for one choice are generated. This may shed a new light on robot personality design, which usually targets on obtaining humans' favour, and provides multiple possibilities to design robots for people with different preferences. The Behaviour-Generating module also presents a generation approach without referring to the behaviours of robot prototype organism, e.g., rabbits for the MiRo robot. Humans still being capable of percepting the emotions behind robots in this study, and interestingly the accuracy for distinguishing PLEASURE dimension of robot emotions is extremely high. In addition, the results reveal possible correlation of P, A, D dimensions in human emotion perception. This may be caused by the limited positions of natural emotions in the PAD emotion space, and humans tend to categorize emotions aside from that into those natural ones. Further work can be conducted and research whether this is a wide existing feature in human-robot interaction.

5.2 Limitations

The MiRo robot used in this study does not have much expressive options and the variety of movement patterns for expressive behaviour is limited. This situation was suitable for the present study, in line with the MiRo robot being designed for specific use cases, but requires a considerable extension, if direct and unsupervised applications of pet-like robots in educational or therapeutic situations are to be realised and, in particular, if a development towards commercialisation is intended.

Because only subjective decision making is considered here, the evaluation does not include the realism of the robot behaviour, its compliance with expectations, and efficiency of the behaviour, which is a problem in particular in the presence of an unpredictable human. Although some of these aspects have been addressed in the experiments included here, but the design of the experiments is biased toward simplicity and assessability which may not be indicative for the general richness of human interaction. Although only consider basic emotions and physiological factors were studied, it is still a valuable start that contributes to our understanding of the ways that human-robot interaction can develop. Further studies will focus on more complex emotion structures and more complex interactions among the modules of the AROM model.

We should also mention that the system is based on MDP instead of a POMDP which would appear to be more suitable for a realistic situation where many aspects of the mutual affective contexts are not directly observable. Here we need to choose among several options, whether the limitations of the continued use of an MDP are acceptable, whether a probabilistic sensor model is sufficient, whether more information is to be acquired by a wider action set, or whether it is sufficient to work with a small number of hypotheses in parallel. Similar options have been studied in the field of POMDP, but a decision is possible only in a practical setting.

Ethical questions related to the project have been addressed in the context of an Informatics ethics application which has been approved. However, long-term effects that are explicitly excluded from the present study, can be expected to be more critical. In the long-term interaction with a robot, well-meaning design features may have effects opposite to the original intentions, especially for an AI-based learning robot. For example, are robot supporting a human could affect lower levels of activity in the human, or it could become an annoyance for the human, if the robot has learned patterns in the human behaviour that the human prefers to keep private. These aspects are part of the on-going debate about the role of AI in the future, and are clearly beyond the scope of this project. It is, however, possible to predict that features of the current design, such as interpretability as a main characteristic of the robot mind will be useful also in the more general situation of use of such a robot in practice. Nevertheless, with the adoption of robots as companions, we are entering here uncharted territory that requires broader and deeper insights before it can be entered.

5.3 Future Work

Further experiments can be implemented based on the present study. To explore the boundaries of *AROM* system, other types of robots such as humanoid robots can be applied. It is also beneficial to conduct experiments on the combination of the Decision-making and Behaviour-Generating module, evaluating the joint interpretability of decisions and behaviours under the same affective context. Research concentrated on the influences of DOPAMINE and ”wanting” are also meaningful, which may fill the gap of human perception of robot intentions from an aspect of emotion. Furthermore, since the nature of the *AROM* emotion model is to build connection between stimuli and emotions, advanced exploration of “stimulus-reinforcement association” [35] may enable more complex behaviours on robots. For instance, by leaning the association between the stimulus PLAYGROUND with the emotional reinforcer PLAYING, affective robots can learn to get excited when they encountered with stimuli related to PLAYGROUND.

Chapter 6

Conclusion

Overall, the contribution of this project was to provide an architecture for robots to conduct appropriate actions in the context of a human-robot companionship. The system allows for a concurrent and adaptive diversity of behaviours, including the tendency of the robot to explore its environment, to execute actions for its own maintenance, and to engage with a human partner. This biomimetic approach to behaviour generation may provide a reference for a generic robot where copying of seemingly similar animal behaviours is not an option or for the design of a specifically animal-like robot that needs to include interpretability and adaptability within to context of a potentially vulnerable human partner. The novel affect model proposed here could also be useful for robot-assisted support for autism children, since it includes more dimensions such as WANTING compared to traditional models with core emotions only. Likewise, a framework is suitable by design as a companion robot for elderly people, as a social robot in various contexts or simply as a toy. It remains to be tested whether the proposed framework would be suitable also for a companion of solitary pet animals.

Bibliography

[1] Tierney A. and Mangiamele L. Effects of serotonin and serotonin analogs on posture and agonistic behavior in crayfish. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 187(10):757–767, December 2001.

[2] Sara Ali, Faisal Mehmood, Khawaja Fahad Iqbal, Yasar Ayaz, Muhammad Sajid, Muhammad Baber Sial, Muhammad Faiq Malik, and Kashif Javed. Human robot interaction: Identifying resembling emotions using dynamic body gestures of robot. In *2023 3rd International Conference on Artificial Intelligence (ICAI)*, pages 39–44, 2023.

[3] Richard Bellman. A Markovian decision process. *Journal of Mathematics and Mechanics*, 6(55):679–684, 1957.

[4] Kent C Berridge and Terry E Robinson. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, 28(3):309–369, 1998.

[5] Jogn John Cabibihan, Hifza Javed, Marcelo Ang, and Sharifah Mariam Aljunied. Why robots? a survey on the roles and benefits of social robots in the therapy of children with autism. *International Journal of Social Robotics*, 5:593–618, 2013.

[6] Sophie Callier, Marina Snapyan, Stéphane Le Crom, Delphine Prou, Jean-Didier Vincent, and Philippe Vernier. Evolution and cell biology of dopamine receptors in vertebrates. *Biology of the Cell*, 95(7):489–502, 2003.

[7] Emily C Collins, Tony J Prescott, Ben Mitchinson, and Sebastian Conran. MIRO: A versatile biomimetic edutainment robot. In *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology*, pages 1–4, New York, NY, USA, 2015. Association for Computing Machinery.

[8] Ignasi Cos, Lola Cañamero, Gillian M Hayes, and Andrew Gillies. Hedonic value: enhancing adaptation for motivated agents. *Adaptive Behavior*, 21(6):465–483, 2013.

[9] Anna Cremer, Felix Kalbe, Jana Christina Müller, Klaus Wiedemann, and Lars Schwabe. Disentangling the roles of dopamine and noradrenaline in the exploration-exploitation tradeoff during human decision-making. *Neuropsychopharmacology*, 48(7):1078–1086, June 2023.

[10] Fani Deligianni, Yao Guo, and Guang-Zhong Yang. From Emotions to Mood Disorders: A Survey on Gait Analysis Methodology. *IEEE Journal of Biomedical and Health Informatics*, 23(6):2302–2316, November 2019. Conference Name: IEEE Journal of Biomedical and Health Informatics.

[11] Alexandra Dobrokvashina, Shifa Sulaiman, Aidar Zagirov, Elvira Chebotareva, Hsia Kuo-Hsien, and Evgeni Magid. Human robot interaction in collaborative manufacturing scenarios: Prospective cases. In *2022 International Siberian Conference on Control and Communications (SIBCON)*, pages 1–6, 2022.

[12] Dietrich Dörner and C Dominik Güss. Psi: A computational architecture of cognition, motivation, and emotion. *Review of General Psychology*, 17(3):297–317, 2013.

[13] Paul Ekman. *Basic Emotions*, chapter 3, pages 45–60. John Wiley Sons, Ltd, 1999.

[14] Pablo Gómez Esteban and David Ríos Insua. A model for an affective non-expensive utility-based decision agent. *IEEE Transactions on Affective Computing*, 10(4):498–509, 2019.

[15] Michalis Feidakis, Iro Gkolompia, Athina Marnelaki, Konstantina Marathaki, Sofia Emmanouiilidou, and Eleni Agrianiti. Nao robot, an educational assistant in training, educational and therapeutic sessions. In *2023 IEEE Global Engineering Education Conference (EDUCON)*, pages 1–6, 2023.

[16] Enrique Fernández-Rodicio, Marcos Maroto-Gómez, Álvaro Castro-González, María Malfaz, and Miguel Á. Salichs. Emotion and Mood Blending in Embodied Artificial Agents: Expressing Affective States in the Mini Social Robot. *International Journal of Social Robotics*, 14(8):1841–1864, October 2022.

[17] Liang Gong, Changyang Gong, Zhao Ma, Lujie Zhao, Zhenyu Wang, Xudong Li, Xiaolong Jing, Haozhe Yang, and Chengliang Liu. Real-time human-in-the-loop remote control for a life-size traffic police robot with multiple augmented reality aided display terminals. In *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*, pages 420–425, 2017.

[18] Antonio Greco, Antonio Roberto, Alessia Saggese, Mario Vento, and Vincenzo Vigilante. Emotion analysis from faces for social robotics. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 358–364, 2019.

[19] Thierry Gélat and Carole Ferrel Chapus. Reaction time in gait initiation depends on the time available for affective processing. *Neuroscience Letters*, 609:69–73, November 2015.

[20] Shaun Halovic and Christian Kroos. Not all is noticed: Kinematic cues of emotion-specific gait. *Human Movement Science*, 57:478–488, February 2018.

[21] Jung Eun Kim, Sujin Chae, Sungsoo Kim, Yeon-Joo Jung, Myoung-Goo Kang, Won Do Heo, and Daesoo Kim. Cerebellar 5ht-2a receptor mediates stress-induced onset of dystonia. *Science Advances*, 7(10):eabb5735, 2021.

[22] Miriam Koschate, Richard Potter, Paul Bremner, and Mark Levine. Overcoming the uncanny valley: Displays of emotions reduce the uncanniness of humanlike robots. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, HRI '16, page 359–365, Christchurch, New Zealand, 2016. IEEE Press.

[23] Songpol Kulviwat, Gordon C. Bruner Ii, Anand Kumar, Suzanne A. Nasco, and Terry Clark. Toward a unified theory of consumer acceptance technology. *Psychology and Marketing*, 24(12):1059–1084, December 2007.

[24] Matthew Lewis and Lola Cañamero. Hedonic quality or reward? a study of basic pleasure in homeostasis and decision making of a motivated autonomous robot. *Adaptive Behavior*, 24(5):267–291, 2016. PMID: 28018120.

[25] Teng Li, Armin Badre, Hamid D. Taghirad, and Mahdi Tavakoli. Integrating impedance control and nonlinear disturbance observer for robot-assisted arthroscope control in elbow arthroscopic surgery. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11172–11179, 2022.

[26] Irwin Lucki. The spectrum of behaviors influenced by serotonin. *Biological Psychiatry*, 44(3):151–162, August 1998.

[27] Hugo Lövheim. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2):341–348, 2012.

[28] Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14:261–292, 1996.

[29] Justin Miller, Sanghyun Hong, and Jianbo Lu. Self-driving mobile robots using human-robot interactions. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1251–1256, 2018.

[30] David R. Montes, Colin M. Stopper, and Stan B. Floresco. Noradrenergic modulation of risk/reward decision making. *Psychopharmacology*, 232(15):2681–2696, August 2015.

[31] O.O. Ogorodnikova. Safe and reliable human-robot interaction in manufactory, within and beyond the workcell. In *19th International Workshop on Robotics in Alpe-Adria-Danube Region (RAAD 2010)*, pages 65–70, 2010.

[32] Kartik Anand Pant, Amod Hegde, and Kothapalli Venkata (KV) Srinivas. Thompson sampling with virtual helping agents, 2022.

[33] André Parent. Functional anatomy and evolution of monoaminergic systems. *American Zoologist*, 24(3):783–790, 1984.

[34] Antonio Rangel, Colin Camerer, and P Read Montague. A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7):545–556 (Box 3), 2008.

[35] Edmund T. Rolls. Précis of the brain and emotion. *Behavioral and Brain Sciences*, 23(2):177–191, 2000.

[36] Daniel Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, and Zheng Wen. A tutorial on thompson sampling, 2020.

[37] Ramesh Saini. Research on the recent trends and challenges on the adoption of robotics in the healthcare sector. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, pages 1140–1144, 2023.

[38] Philip Scales, Véronique Aubergé, and Olivier Aycard. Socio-expressive robot navigation: How motion profiles can convey frailty and confidence. In *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 1–8, 2022.

[39] Knut Schmidt-Nielsen. *Animal physiology: adaptation and environment*. Cambridge university press, 1997.

[40] Gary Sieck. Physiology in perspective: Adaptive responses: Changing to survive. *Physiology*, 29(3):157–158, 2014. PMID: 24789979.

[41] Alec Solway and Matthew M Botvinick. Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychological Review*, 119(1):120, 2012.

[42] Guangqing Song and Jingtao Lei. Control system design of heterogeneous master-slave robot for fracture reduction surgery of long bone. In *2023 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 2303–2308, 2023.

[43] Ruth Stock-Homburg. Survey of emotions in human–robot interactions: Perspectives from robotic psychology on 20 years of research. *International Journal of Social Robotics*, 14(2):389–411, 2022.

[44] S. Stork, C. StoBel, H. J. Muller, M. Wiesbeck, M. F. Zah, and A. Schubo. A neuroergonomic approach for the investigation of cognitive processes in interactive assembly environments. In *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pages 750–755, 2007.

[45] Michael Suguitan, Mason Bretan, and Guy Hoffman. Affective robot movement generation using cyclegans. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 534–535, 2019.

[46] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*. Adaptive computation and machine learning series. The MIT Press, Cambridge, Massachusetts, second edition edition, 2018.

[47] William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.

[48] Philippe N. Tobler, Christopher D. Fiorillo, and Wolfram Schultz. Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715):1642–1645, 2005.

[49] Christiana Tsiourti, Astrid Weiss, Katarzyna Wac, and Markus Vincze. Multimodal Integration of Emotional Signals from Voice, Body, and Context: Effects of (In)Congruence on Emotion Recognition and Attitudes Towards Robots. *International Journal of Social Robotics*, 11(4):555–573, August 2019.

[50] Valeria Villani, Fabio Pini, Francesco Leali, and Cristian Secchi. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55:248–266, 2018.

[51] Dongsheng Yang, Wataru Sato, Qianying Liu, Takashi Minato, Shushi Namba, and Shin'ya Nishida. Optimizing facial expressions of an android robot effectively: a bayesian optimization approach. In *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, pages 542–549, 2022.

[52] Sangseok You and Lionel Robert. Teaming up with robots: An imoi (inputs-mediators-outputs-inputs) framework of human-robot teamwork. *You, S. and Robert, LP (2017). Teaming Up with Robots: An IMOI (Inputs-Mediators-Outputs-Inputs) Framework of Human-Robot Teamwork, International Journal of Robotic Engineering (IJRE)*, 2(3):3308173(7), 2018.

[53] Jun Zhang, Kent C. Berridge, Amy J. Tindell, Kyle S. Smith, and J. Wayne Aldridge. A neural computational model of incentive salience. *PLOS Computational Biology*, 5(7):1–14, 07 2009.

[54] Minqiu Zhou, Isobel Voysey, and J Michael Herrmann. Is machine learning enough to train robotic pets? In *RO-MAN 2022 Workshop on Machine Learning for HRI: Bridging the Gap between Action and Perception*, pages 1–4. HAL, 2022.