

# 基于eBICA认知架构的具身智能数学模型

郑建拥\*, 张 晨, 王 茜, 范 翔

北京跟踪与通信技术研究所, 北京

收稿日期: 2026年8月19日; 录用日期: 2026年9月17日; 发布日期: 2026年9月23日

## 摘 要

近年来具身智能异军突起,但在复杂社会环境中的适应性,亟需超越纯粹数据逻辑的深度情感认知能力。现有具身智能架构常面临情感认知计算与躯体感知隔离、自适应调节机制不足等问题。提出一种基于eBICA (emotional Brain Inspired Cognitive Architecture)的具身智能体数学模型,旨在通过情感计算、脑认知语义映射与伦理图式的数学化与自我调节学习的深度耦合机制,实现社会情感智能的可计算化。首先,建立eBICA六因子和认知图式的统一数学表征体系,消除感知-情感-行动的语义鸿沟;接着加入情感调节的自我学习模型,实现策略参数和情感增益的自适应调节;最后构建身体-情感双向动力学方程,建立身体信号与情感状态的实时互反馈模型,并提出评价指标。为构建具有深度社会理解与情感交互能力的具身智能体提供可参考的数学基础与算法框架。

## 关键词

具身智能, 情感认知计算, eBICA架构, 身体-情感耦合, 数学模型

# Embodied Intelligence Mathematical Model Based on eBICA Cognitive Architecture

Jianyong Zheng\*, Chen Zhang, Qian Wang, Xiang Fan

Beijing Institute of Tracking and Telecommunications Technology, Beijing

Received: August 19, 2026; accepted: September 17, 2026; published: September 23, 2026

## Abstract

In recent years, embodied intelligence has emerged as a new force, but its adaptability in complex social environments urgently requires deep emotional cognitive capabilities that transcend pure data logic. Current embodied intelligence architectures often face issues such as the isolation of emotional cognitive computation from bodily perception and insufficient adaptive regulation mechanisms. A mathematical model of embodied intelligent agents based on eBICA (emotional Brain

\*通讯作者。

文章引用: 郑建拥, 张晨, 王茜, 范翔. 基于 eBICA 认知架构的具身智能数学模型[J]. 理论数学, 2026, 16(9): 104-113.

DOI: 10.12677/pm.2026.169203

Inspired Cognitive Architecture) is proposed, aiming to achieve the computability of social-emotional intelligence through a deeply coupled mechanism of emotional computation, brain cognitive semantic mapping, ethical schemas, and self-regulatory learning. Firstly, a unified mathematical representation system of eBICA's six factors and cognitive schemas is established to eliminate the semantic gap between perception-emotion-action. Then, a self-learning model for emotional regulation is incorporated to achieve adaptive regulation of strategy parameters and emotional gains. Finally, a bidirectional dynamic equation between body and emotion is constructed, establishing a real-time mutual feedback model between bodily signals and emotional states, and proposing evaluation metrics. This provides a referential mathematical foundation and algorithmic framework for constructing embodied intelligent agents with deep social understanding and emotional interaction capabilities.

## Keywords

Embodied Intelligence, Affective Cognition Computing, eBICA Architecture, Body-Emotion Coupling, Mathematical Model

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来,具身智能的研究与应用异军突起,其产生、发展与表达深刻依赖于智能体所拥有的物理躯体及其与环境的实时互动,表明真正的通用智能不仅需要逻辑推理,更需融合对社会情境的理解、情感的体验与表达,以及基于躯体状态的自适应决策[1]。但是当前具身智能与真实人脑或者广义上的生物神经系统的智能相比还存在一些缺点,如数据依赖性强、泛化能力弱、缺乏常识与因果逻辑、决策过程不可解释(也就不可评估)等核心瓶颈,导致缺乏自主性和联想等高级认知处理能力,根源在于现有范式主要依赖于数据驱动的端到端学习,其本质上是统计拟合与模式匹配,对人类智能底层认知结构的借鉴比较少,也就无法针对新颖、复杂、未知等情况或者问题做出类似人的思考与决策等[2][3]。针对这一问题,近年来学术界已经提出一些理论和方法,其中的类脑智能方法受到了关注。类脑智能方法是受大脑或者生物神经运行机制和认知行为机制启发,以计算建模为手段,通过软硬件协同实现的智能技术,旨在达到信息处理机制上类脑、认知行为表现和智能水平上达到或超越人(或其他仿生对象)的目的[4]。需要注意的是,类脑智能≠人工智能,当前的人工智能是一门结合了数学、计算机科学、心理学等多学科理论发展起来的新技术。而类脑智能是在脑科学、神经科学基本原理指导下或启发下进行的智能理论和技术研究与开发[5][6]。尤其是随着大数据、超级计算、智能信息科学等新理论新技术的快速发展,受神经生物学启发的深度神经网络等机器学习算法迅速崛起。当前人工智能的研究已经借鉴大脑的感觉信息处理等功能发展了深度学习、胶囊学习等前沿算法和理论,也得到了很多应用[7]-[16]。但如何借助脑启发实现具身智能的突破,使得智能体和生物大脑一样在面临新环境时具备高度的可塑性、创造性、适应性、灵活性和自学习等能力,仍然是一个巨大挑战[17]。

本文利用基于脑启发的情感认知架构建立具身智能体数学模型,在运用认知体系架构改进智能算法的道路上做有益的探索。

## 2. eBICA 理论背景

情感是个体的认知对外界刺激进行信息综合、推导后产生的,在现有的研究中,国内多数学者主要将认知理论嵌入到原有的情感计算模型中,研究和模拟情感的动态变化过程等[18]-[22]。然而这些研究没

有对引起情感变化的原因进行分析研究,且缺乏一定的系统性,不能简单地扩展应用。国外已有学者将成熟的认知体系结构与情感计算模型相结合,如将 SOAR 认知模型与 ALMA 情感计算模型相结合构建了认知情感架构,但此类研究范用性比较弱,通常仅应用于自适应智能体,使其具有一定表达情感的能力[23][24];ACT-R 模型基于记忆进行感情建模,通常用于理性思维的研究[25]-[29];此外还有基于 PMF-serv 和 BDI 等模型的研究,但都缺乏脑启发的情感认知计算相关研究支撑[30]-[36]。

基于脑启发的情感认知统一架构对于未来智能体更广泛地融入人类社会至关重要,该领域研究起始于通过人工大脑表征或映射一套与人类认知相统一的思维机制,促进人工智能与人类认知行为的一致性。但人工智能的认知模式较之于人类丰富的情感认知体系相差甚远[37]。Alexei Samsonovich 基于认知统一理论开发了一套受脑启发的情感认知架构(emotional Brain Inspired Cognitive Architecture, eBICA)旨在通过该系统表征并模拟人类社会情感机制,弥合人工智能和人类智能在情感认知上的差距[38],正适合用于增强具身智能体的社会情境交互和情感认知计算能力。

eBICA 侧重对情绪反应与情感认知进行表征与设计:一是标记接触新事件或目标的情绪感受,在之后处理类似事件时获取情感记忆从而增强决策效率;二是能够对某事件产生特定或稳定的情感态度与价值评价;三是通过语义映射与伦理图式产生对某事件或目标评价的差异,从而选择不同的情感表达路径。具体而言, eBICA 通过评价(Appraisal)、知觉标记(Somatic marker)、感受(Feeling)、情感状态(Emotion)、偏差(Bias)、反应(Reaction)六个因子的关系来实现人工智能社会情感进路设计[37][38]。基于这六因子,构建具身智能体数学模型。

### 3. 基本定义与符号集

$S \subseteq R^{n_s}$ :  $S$  环境状态空间,  $R^{n_s}$  为  $n_s$  维实数向量空间,环境状态由  $n_s$  个实数特征(如位置、温度、物体属性)描述;

$B \subseteq R^{n_b}$ :  $B$  身体状态空间,  $R^{n_b}$  为  $n_b$  维实数向量空间,身体状态由  $n_b$  个实数特征(如姿态、生理参数等)描述;

$A \subseteq R^{n_a}$ :  $A$  动作空间,  $R^{n_a}$  为  $n_a$  维实数向量空间,动作状态由  $n_a$  个实数特征描述;

$\mathcal{E} \subseteq R^{n_e}$ :  $\mathcal{E}$  情感状态空间,  $R^{n_e}$  为  $n_e$  维实数向量空间,情感状态由  $n_e$  个实数特征描述;

$G$ : 目标集合;

$M$ : 认知地图空间;

$s_t \in S$ :  $t$  时刻环境状态;

$b_t \in B$ :  $t$  时刻身体状态;

$e_t \in E$ :  $t$  时刻情感状态;

$a_t \in A$ :  $t$  时刻执行动作;

$m_t \in M$ :  $t$  时刻认知地图;

$g \in G$ : 当前活跃目标。

### 4. eBICA 六因子的数学实现

基于 eBICA 六因子的关系,设计智能体基本框架。

#### 4.1. 评价函数(Appraisal)

评价函数是多维价值评估的加权综合:

$$\text{Appraisal} = \Phi(s_t, b_t, e_t, a_t, m_t, g) = W \cdot \phi(s_t, b_t, e_t, a_t, m_t, g) \quad (1)$$

其中:  $W=[w_1, w_2, \dots, w_6] \in R^{n_e \times 6}$  是权重矩阵,  $\phi=[\phi_1, \phi_2, \dots, \phi_6]^T$  是六个评价维度:

① 安全性评价:

$$\phi_1 = \exp\left(-\frac{\min_{o \rightarrow O} \|p_o - p_b\|^2}{2\sigma^2}\right) \quad (2)$$

其中:  $O$  是障碍物集合,  $p_o$  是障碍物位置,  $p_b$  是身体位置,  $\sigma$  是 Sigmoid 函数。

② 能量效率评价:

$$\phi_2 = \frac{E_{\max} - \|a_{t-1}\|^2 \cdot c_{\text{action}}}{E_{\max}} \cdot \sigma\left(\frac{b_t^{\text{energy}} - \tau_{\text{low}}}{\tau_{\text{high}} - \tau_{\text{low}}}\right) \quad (3)$$

其中:  $E_{\max}$  表示最大能量,  $\|a_{t-1}\|$  表示上一时刻动作能耗,  $c_{\text{action}}$  表示动作成本,  $b_t^{\text{energy}}$  表示  $t$  时刻能量水平,  $\tau_{\text{high}}$ 、 $\tau_{\text{low}}$  表示能量上下警戒限。

③ 社会距离评价:

$$\phi_3 = \frac{1}{1 + \exp(-\kappa(d_{\text{act}} - d_{\text{exe}}))} \quad (4)$$

$d_{\text{act}}$  实际距离(标量),  $d_{\text{exe}}$  预期社会距离(基于文化规范),  $\kappa$  是 sigmoid 函数的灵敏度参数(正实数),  $\exp(\cdot)$  为 sigmoid 函数, 输出接近 1 表示实际距离过近, 接近 0 表示过远。

④ 目标进展评价:

$$\phi_4 = \frac{\Delta P_g}{\Delta t} \cdot e^{-\gamma'_g} \cdot I(P_g \leq P_{\text{threshold}}) \quad (5)$$

其中:  $\Delta P_g$  目标进度变化量,  $\gamma_g$  目标紧迫性衰减因子,  $I(\cdot)$  为指示函数(条件满足时为 1),  $P_g$  为当前目标进度,  $P_{\text{threshold}}$  为目标阈值。

⑤ 合规性评价:

$$\phi_5 = \sum I(\text{violation\_score}(a_i)) \quad (6)$$

其中,  $I(\cdot)$  为指示函数(违反规则时为 1)。

⑥ 舒适性评价:

$$\phi_6 = \sigma\left(\sum_i a_i \cdot (b_t^i - b_{\text{comfort}}^i)\right) \quad (7)$$

其中:  $a_i$  是权重,  $b_t^i$  是  $t$  时刻第  $i$  个身体状态分量,  $b_{\text{comfort}}^i$  是第  $i$  个身体状态最佳舒适值。

## 4.2. 知觉标记(Perceptual Tagging)

知觉标记是一个从原始感知到情感标签空间的映射:

$$T_t = \psi(s_t, b_t, e_t) = f_{\text{tag}}(s_t) \odot g(b_t, e_t) \quad (8)$$

其中:  $T_t$  是  $t$  时刻知觉标记,  $\psi$  是知觉标记函数,  $f_{\text{tag}}$  表示  $S \rightarrow R^n$  的映射, 是多模态感知编码器, 将环境状态映射到标签空间,  $g$  表示  $B \times E \rightarrow R^n$ , 为身体-情感调制函数。  $g(b_t, e_t) = 1 + \tanh(U_b b_t + U_e e_t)$ ,  $U_b$ 、 $U_e$  为权重矩阵, 用于调制感知。

## 4.3. 感受与情感状态(Feeling & Emotional State)

情感状态演化遵循带噪声的微分方程:

$$\frac{de_t}{dt} = -\Lambda e_t + \Gamma_A A_t + \Gamma_T T_t + \Gamma_D D_t + \xi_t \quad (9)$$

离散时间形式:

$$e_{t+1} = (I_0 - \Lambda \Delta t) e_t + \Gamma_A A_t \Delta t + \Gamma_T T_t \Delta t + \Gamma_D D_t \Delta t + \xi_t \quad (10)$$

其中:  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{n_e})$  为情感衰减矩阵, 对角矩阵;  $\Gamma$  为增益矩阵,  $A$ 、 $T$ 、 $D$  对应评价、知觉标记和偏差增益。  $I_0$  为单位矩阵,  $\xi_t$  为高斯噪声。

补充稳定性条件, 取 Lyapunov 函数:

$$V(e) = \frac{1}{2} e^T P e \quad (11)$$

$P$  为正定对称矩阵, 则:

$$V = e^T P [-\Lambda e + \Gamma_A A_t + \Gamma_T T_t + \Gamma_D D_t] + \frac{1}{2} \text{tr}(\sum \xi) \quad (12)$$

当满足  $\lambda_{\min}(\Lambda) > \frac{1}{2} \|\Gamma_A + \Gamma_T + \Gamma_D\|_2$  时, 存在  $P$  使  $V < -k \|e\|^2 + c$ 。

#### 4.4. 偏差计算(Deviation)

偏差是预测误差的情感加权:

$$D_t = V \cdot (s_t^{\text{predicted}} - s_t^{\text{actual}}) \quad (13)$$

其中:  $V$  为加权矩阵,  $s_t^{\text{predicted}}$ 、 $s_t^{\text{actual}}$  分别为预测状态和实际状态。

#### 4.5. 反应生成(Response)

反应生成是一个带约束的优化问题:

$$a_t^* = \arg \max_{a \in A} J(a; s_t, b_t, e_t, g) \quad (14)$$

$a_t^*$  为最优动作,  $J(a)$  为目标函数:

$$J(a) = \alpha_1 R_g(a, g) + \alpha_2 R_e(a, e_t) - \alpha_3 C_b(a, b_t) - \alpha_4 C_m(a, m_t) \quad (15)$$

其中:  $\alpha_i$  为权重系数,  $R_g(a, g) = E[P_g(t+1)|a] - P_g(t)$ , 代表目标奖励, 指采取动作  $a$  后目标进展的变化。

$R_e(a, e_t) = -\|\text{Express}(a) - e_t\|^2$ , 表示情感一致性,  $\text{Express}(a)$  表示动作  $a$  表达的情感状态。

$C_b(a, b_t) = \|a\|^2 \cdot (1 + \beta \cdot I(b_t^{\text{energy}} \leq \text{threshold}))$ , 表示身体动作成本。

$C_m(a, m_t) = \sum \text{violation\_score}(a, m_t)$ , 表示社会成本。

### 5. 认知地图的数学结构

认知地图是一个图结构, 是 eBICA 伦理图式的数学化:

$$m_t = (V, X, W) \quad (16)$$

其中:  $V$  是顶点集合, 包括目标、其他个体和自身;  $X$  是边集合, 表示关系;  $W$  则为权重函数, 表示关系强度。

认知地图的更新公式如下:

$$m_{t+1} = \text{Update}(m_t, s_t, a_t, s_{t+1}, r_t) \quad (17)$$

具体更新:

$V_{t+1} = V_t \cup \text{NewObjects}(s_{t+1})$ , 表示新检测到的对象。

$X_{t+1} = X_t \cup \{\text{Int}(v_i, v_j)\}$ ,  $v_i \in V$ ,  $\text{Int}(v_i, v_j)$  表示顶点  $v_i$  和  $v_j$  产生的交互。

$W_{t+1}(x) = (1 - \eta_w)W_t(x) + \eta_w \cdot \text{Rl}(x | a_t, r_t)$ ,  $\eta_w$  为学习率,  $x \in X$ ,  $r_t$  为  $t$  时刻奖励,  $\text{Rl}(x | a_t, r_t)$  表示基于动作和奖励计算与边  $x$  的相关性。

## 6. 情感调节学习的数学模型

### 6.1. 策略参数自适应

使用策略梯度方法更新评价权重  $W$ , 实现策略参数自适应调节, 最大化累积奖励:

$$\nabla_W J = \mathbb{E} \left[ \sum_{\tau=t}^T \gamma^{\tau-t} (r_\tau - V(s_\tau)) \cdot \frac{\nabla_W \pi_W(a_\tau)}{\pi_W(a_\tau)} \right] + \lambda_W \|W\|_F \quad (18)$$

其中,  $\nabla_W J$  目标函数  $J$  对参数  $W$  的梯度。  $\mathbb{E}[\bullet]$  为期望算子, 在策略分布下求平均。  $\tau$  是时间步索引(从当前  $t$  到终局  $T$ )。  $\gamma$  为折扣因子(降低未来奖励的重要性)。  $r_\tau$  为时间  $\tau$  时刻的奖励(标量)。  $\pi_W(a_\tau)$  为策略函数(在给定状态下选择动作  $a_\tau$  的概率),  $\lambda_W$  是  $W$  的正则化系数,  $\|W\|_F$  是  $W$  的 Frobenius 的范数。

### 6.2. 情感调节学习

情感增益矩阵的自适应学习:

$$\Gamma_{t+1} = \Gamma_t + \eta_\Gamma \cdot \frac{\partial}{\partial \Gamma} \left( \|e_t - e_{\text{optimal}}(g)\|^2 + \lambda \|\Gamma\|_F^2 \right) \quad (19)$$

$\eta_\Gamma$  为学习率,  $\Gamma = [\Gamma_A, \Gamma_T, \Gamma_D]$  为增益矩阵。  $e_{\text{optimal}}(g)$  为目标  $g$  的最佳情感状态。  $\lambda$  正则化系数。

## 7. 身体 - 情感耦合机制的数学建模

耦合机制微分方程:

$$\begin{cases} \frac{db_t}{dt} = -K_b b_t + L_{be} \cdot \sigma(U_e e_t) + B a_t \\ \frac{de_t}{dt} = -\Lambda e_t + \Gamma_A A_t + \Gamma_T T_t + \Gamma_D D_t + L_{eb} \cdot \tanh(V_b b_t) \end{cases} \quad (20)$$

$-K_b b_t$  表示稳态恢复,  $K_b = \text{diag}(k_1, \dots, k_b)$  身体状态恢复系数矩阵(表示身体状态回归稳态的速度);  $L_{be} \cdot \sigma(U_e e_t)$  表示情感影响;  $B a_t$  表示动作作用;  $-\Lambda e_t$  情感衰减;  $\Gamma_A A_t + \Gamma_T T_t + \Gamma_D D_t$  表示认知输入;  $L_{eb} \cdot \tanh(V_b b_t)$  表示身体感受;  $L_{be}$ 、 $L_{eb}$  耦合矩阵, 表示和身体之间相互的影响(如疲劳时,  $b_t$  低, 通过  $L_{eb}$  降低情感唤醒度)。  $B$  动作影响矩阵。  $b_t$   $t$  时刻身体状态。  $U_e$ 、 $V_b$  为权重矩阵, 用于特征提取。

## 8. 评价指标

1) 情感真实性: 衡量智能体外在行为表达与内部情感状态的一致性程度。值越接近 1, 表明智能体行为越真实反映其内在情感。  $\text{Expr}(a_t)$  表示从动作  $a_t$  中提取的情感表达特征(如语音语调、面部表情、肢体动作的数学表示)。

$$\text{Auth} = \frac{1}{T} \sum_{t=1}^T \exp \left( -\frac{\|\text{Expr}(a_t) - e_t\|^2}{\sigma^2} \right) \quad (21)$$

2) 社会适应性: 量化智能体在社会交互中获取积极反馈的能力, 反映其符合社会规范的程度。PositiveFB 表示交互中取得积极反馈, NegativeFB 表示消极反馈次数, Interactions 表示交互行为。

$$\text{SocAdapt} = \frac{\sum \text{PositiveFB} - \sum \text{NegativeFB}}{\sum \text{Interactions}} \quad (22)$$

3) 学习效率: 衡量单位训练步长下的性能提升速度, 反映学习算法的有效性。Performance<sub>*i*</sub><sup>final</sup> 表示第 *i* 个任务中的初始得分, Performance<sub>*i*</sub><sup>initial</sup> 表示最终得分, *N* 为任务总数, TrainSteps 表示完成训练所需的总步长。

$$\text{LearnEff} = \frac{\sum_{i=1}^N (\text{Performance}_i^{\text{final}} - \text{Performance}_i^{\text{initial}})}{N \cdot \text{TrainSteps}} \quad (23)$$

4) 认知地图精度: 量化智能体对环境空间关系的建模误差, 表示对环境中节点(如物体、位置)间连接关系的内部表征的准确度。W<sub>pred(v<sub>i</sub>, v<sub>j</sub>)</sub> 表示顶点 v<sub>i</sub> 和 v<sub>j</sub> 的期望权重, W<sub>true(v<sub>i</sub>, v<sub>j</sub>)</sub> 是真实权重。

$$\text{CMA} = E_{(v_i, v_j)} \in E \left[ \frac{\|W_{\text{pred}(v_i, v_j)} - W_{\text{true}(v_i, v_j)}\|}{1 + W_{\text{true}(v_i, v_j)}} \right] \quad (24)$$

5) 能量利用效率: 衡量智能体单位动作能耗所获的能量收益。I(TaskSuccess) 为任务成功指示函数, 成果为 1, 失败为 0, 仅奖励有效行为。

$$\text{EUE} = \frac{\sum \phi_2(a_t, b_t) \cdot I(\text{TaskSuccess})}{\sum \|a_t\|^2} \quad (25)$$

## 9. 执行程序

整个数学框架的执行流程是: 初始化→感知处理→评价计算→情感更新→决策生成→动作执行→学习调节→状态更新。

### 9.1. 初始化阶段

参数初始化:

身体状态  $b_0$ ;

情感状态  $e_0$ ;

权重矩阵  $W$ ;

情感增益矩阵  $\Gamma = [\Gamma_A, \Gamma_T, \Gamma_D]$ ;

学习率  $\eta_W$ 、 $\eta_\Gamma$ 。

安全阈值设定:

生理安全范围  $[b_{\min}, b_{\max}]$ ;

情感饱和阈值  $[e_{\min}, e_{\max}]$ ;

目标紧迫性衰减因子  $\gamma_g$ ;

合规性规则集 violation\_score。

### 9.2. 主循环

1、输入环境状态  $s_t$ 、身体状态  $b_t$ ; 输出感知特征向量。

- 2、六维度评价：
  - 1) 安全性评价：检测危险信号，计算安全距离；
  - 2) 能量效率评价：计算能量效率；
  - 3) 社交评价：识别其他个体情感状态，计算社交距离；
  - 4) 目标进展评价：计算目标进展；
  - 5) 规范评价：检测规范违反，计算规范一致性；
  - 6) 舒适性评价：计算身体不适度和环境适宜度。
- 3、更新情感状态，检查稳定性。
- 4、动作决策生成：
  - 1) 策略选择；
  - 2) 反应生成优化；
  - 3) 安全检查。
- 5、动作执行与环境交互：
  - 1) 执行动作  $a_t$ ；
  - 2) 接受奖励  $r_t$ ；
  - 3) 观察新状态  $s_{t+1}$ 、 $b_{t+1}$ 。
- 6、学习与调节：
  - 1) 策略更新；
  - 2) 情感增益更新。
- 7、身体 - 情感耦合更新：
  - 1) 身体状态更新；
  - 2) 情感状态再调整；
  - 3) 更新  $U_b$ 、 $U_e$  和  $L_{be}$ 、 $L_{eb}$  权重矩阵。
- 8、认知地图更新：
  - 1) 更新状态转移概率；
  - 2) 调整评价函数权重；
  - 3) 更新目标紧迫性衰减因子。
- 9、终止条件：
  - 1) 任务完成：达到目标状态；
  - 2) 安全违规：身体状态超出安全范围；
  - 3) 时间耗尽：达到最大时间步；
  - 4) 情感崩溃：情感状态持续不稳定或超出阈值。

### 9.3. 关键模块

- 1、情感自适应调节模块
  - 1) 计算情感偏差  $\Delta e$ ；
  - 2)  $\Delta e > e_{\max}$ ，则增强负反馈，降低学习率； $\Delta e < e_{\min}$ ，则反之。
- 2、记忆与学习模块
  - 1) 记忆存储：状态 - 动作集、情感变化轨迹和评价函数历史；
  - 2) 学习触发条件：周期性更新、重大事件(大奖励/惩罚)、稳定下降。

### 3、监控与评估

- 1) 实时监控五项评价指标;
- 2) 每  $K$  步评估: 平均奖励、任务完成率、约束违反次数、情感波动范围。

## 10. 结束语

本研究立足于具身智能对深度社会情感认知的迫切需求, 提出了一套基于 eBICA 情感认知架构的具身智能体数学模型。该模型系统地实现了感知、情感、认知与行动在统一数学框架下的深度耦合, 将脑启发的情感认知原理转化为一套可计算、可分析、可验证的数学框架, 构建了一个兼具生物与计算的数学模型。该模型为开发能够在复杂社会环境中进行自适应情感交互、具备深度情境理解能力的下一代具身智能体, 提供可参考的数学基础与可扩展的算法框架。

## 参考文献

- [1] 董金成, 周垒, 孙琦钰, 等. 仿生智能机器人技术: 从形态结构到具身智能[J]. 信息与控制, 2026, 55(2): 228-246.
- [2] 郑南宁, 杨劼, 姜维周, 孙宏滨, 丁宁. 具身智能发展趋势与展望[J]. 中国工程科学, 2026, 28(2): 1-13.
- [3] 陶永, 万嘉昊, 王田苗, 等. 构建具身智能新范式: 人形机器人技术现状及发展趋势综述[J]. 机械工程学报, 2025, 61(15): 121-147.
- [4] 王瑞东, 王睿, 张天栋, 王硕. 机器人脑智能研究综述[J]. 自动化学报, 2024, 50(8): 1485-1501.
- [5] 李迅典. “类脑智能领域是一片蓝海” [N]. 环球时报, 2025-06-11(008).
- [6] Qiao, H., Chen, J. and Huang, X. (2022) A Survey of Brain-Inspired Intelligent Robots: Integration of Vision, Decision, Motion Control, and Musculoskeletal Systems. *IEEE Transactions on Cybernetics*, **52**, 11267-11280. <https://doi.org/10.1109/tyb.2021.3071312>
- [7] 冯建峰. 从脑科学、人工智能到应用[J]. 国际人才交流, 2021(4): 2.
- [8] 王泽雨. 脑启发的视觉 Transformer 压缩算法研究[D]: [硕士学位论文]. 合肥: 中国科学技术大学, 2025.
- [9] 李茹一, 柯铭, 王路斌, 刘珮, 高晋, 王刚. 面向降质光电图像的脑启发无人机小目标鲁棒检测方法[J]. 信号处理, 2025, 41(5): 886-905.
- [10] 贺俊文. 胶质细胞启发的神经网络模型研究[D]: [硕士学位论文]. 武汉: 华中农业大学, 2024.
- [11] 陈亮宇. 基于忆阻的脑启发感知联想电路设计及其应用[D]: [硕士学位论文]. 武汉: 华中科技大学, 2024.
- [12] 刘睿. 自上而下注意力启发的脉冲神经网络终身学习研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2024.
- [13] 隋映, 杨昊保, 杨中海. 基于国产类脑大模型的视频监控事件智能处置与自动定责研究[J]. 中国安防, 2025(11): 29-32.
- [14] 王小毅, 邓万江. AI 大模型与智能体驱动的消费研究新范式: 消费感知、类脑模拟与自主演化[J]. 中国科学基金, 2025, 39(5): 736-747.
- [15] 杨子睿. 基于类脑记忆库的雷达目标检测跟踪方法研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2025.
- [16] 廖元浩, 苏春旺, 江卫国, 文华, 李允君, 张四平, 黄子罡. 类脑智能测评: 内涵、方法与应用[J]. 宇航计测技术, 2025, 45(2): 31-48.
- [17] 曾毅, 张倩, 赵菲菲, 赵东城, 李金东, 申国斌, 毕韦达. 从认知脑的计算模拟到类脑人工智能[J]. 人工智能, 2022, 31(6): 28-40.
- [18] Laird, J.E., Newell, A. and Rosenbloom, P.S. (1987) SOAR: An Architecture for General Intelligence. *Artificial Intelligence*, **33**, 1-64. [https://doi.org/10.1016/0004-3702\(87\)90050-6](https://doi.org/10.1016/0004-3702(87)90050-6)
- [19] Lieto, A., Bhatt, M., Oltramari, A. and Vernon, D. (2017) The Role of Cognitive Architectures in General Artificial Intelligence. *Cognitive Systems Research*, **48**, 1-3. <https://doi.org/10.1016/j.cogsys.2017.08.003>
- [20] Alexandre, F. and Carrere, M. (2016) Modeling Neuromodulation as a Framework to Integrate Uncertainty in General Cognitive Architectures. In: *Lecture Notes in Computer Science*, Springer, 324-333. [https://doi.org/10.1007/978-3-319-41649-6\\_33](https://doi.org/10.1007/978-3-319-41649-6_33)
- [21] Qureshi, R., Sapkota, R., Shah, A., et al. (2025) Thinking beyond Tokens: From Brain-Inspired Intelligence to Cognitive Foundations for Artificial General Intelligence and Its Societal Impact. arXiv:2507.00951.

- 
- [22] Huyck, C. (2017) Programming a Cognitive Architecture with Simulated Neurons, Chris Eliasmith. How to Build a Brain: A Neural Architecture for Biological Cognition. Oxford University Press.
  - [23] Pérez, J., Cerezo, E., Serón, F.J. and Rodríguez, L. (2016) A Cognitive-Affective Architecture for ECAs. *Biologically Inspired Cognitive Architectures*, **18**, 33-40. <https://doi.org/10.1016/j.bica.2016.10.002>
  - [24] 李凯. 具有情感机制的机器人认知模型与算法研究[D]: [硕士学位论文]. 北京: 北方工业大学, 2023.
  - [25] 陈冉, 郭松, 孙浩, 俞晓斌, 韩晓东. 基于 ACT-R 模型的大吨位履带起重机界面设计[J]. 机械设计与制造工程, 2025, 54(8): 107-111.
  - [26] 王茹冰. 基于 ACT-R 理论促进初中函数深度学习的教学研究[D]: [硕士学位论文]. 济南: 山东师范大学, 2025.
  - [27] 张淞琦. 基于 ACT-R 理论的指数函数与对数函数单元教学研究[D]: [硕士学位论文]. 重庆: 重庆三峡科技大学, 2025.
  - [28] 吴蕴超, 罗飞, 陶俊臣, 王晨, 张子亿. 基于 ACT-R 的认知间隔重复学习方法[J]. 华东理工大学学报(自然科学版), 2025, 51(3): 371-379.
  - [29] 张晶晶. 基于 ACT-R 理论的突发事件网民负面情感建模研究[D]: [博士学位论文]. 南京: 南京理工大学, 2022.
  - [30] 陆冰川, 侯键. 多模态 MRI 与人工智能融合在轻度认知障碍诊断及转化预测中的研究进展[J]. 磁共振成像, 2025, 16(12): 184-189.
  - [31] Ron, S. 计算认知架构、双过程理论和大语言模型(英文) [J]. 宁波大学学报(理工版), 2025, 38(2): 17-28.
  - [32] 刘伟, 符征. 深度学习认知架构的反表征主义转向[J]. 长沙理工大学学报(社会科学版), 2024, 39(4): 54-60.
  - [33] 梁亚丽, 张志昌, 张圆圆, 马璐璐. 基于自适应特征选择的情感感知认知诊断模型[J]. 计算机技术与发展, 2025, 35(12): 157-164.
  - [34] 吴春琼, 孙建红, 姜山, 李文, 刘颖琪, 黄玉玲. 数据新闻中 KOL 情感认知引导模型构建与实证分析[J]. 武夷学院学报, 2025, 44(7): 30-38.
  - [35] 张伟庆, 邱莉榕, 张熙. 多源认知知识融合的情感支持对话模型[J]. 东北师大学报(自然科学版), 2025, 57(1): 82-90.
  - [36] 张培杰, 孙颖, 张雪英, 贾海蓉, 段淑斐. 基于关联认知网络的语音情感识别模型[J]. 微电子学与计算机, 2023, 40(8): 1-9.
  - [37] 刘鸿宇. 认知统一理论视角下人工智能社会情感进路设计——基于萨姆索诺维奇的 eBICA 模型[J]. 自然辩证法研究, 2022, 38(8): 28-34.
  - [38] Samsonovich, A.V. and Eidlin, A.A. (2021) A Proposal for Modeling Cognitive Ontogeny Based on the Brain-Inspired Generic Framework for Social-Emotional Intelligent Actors. *Procedia Computer Science*, **186**, 449-455. <https://doi.org/10.1016/j.procs.2021.04.165>