

基于强化学习的机器人认知情感交互模型

黄宏程^{①②} 李净^① 胡敏^① 陶洋^{*①②} 寇兰^①

^①(重庆邮电大学通信与信息工程学院 重庆 400065)

^②(重庆市通信软件工程技术研究中心 重庆 400065)

摘要: 为增强机器人的认知情感计算能力,依据PAD情感空间建立结合即时反馈和长期趋势的机器人认知情感生成方法,该文提出一种基于强化学习的机器人认知情感交互模型。首先,依据人际交往心理学理论,模拟人类情感生成过程进行类人情感生成,并从中提取相似性、积极性、共情性3个影响因素;其次,利用强化学习的全局统筹特性,建立响应情感状态与上下文长期情感状态之间的关联关系,从而对机器人情感生成过程进行建模;然后,将3个因素纳入模型奖励机制用于交互情感状态评估,实现模型更新并得到最优情感策略;最后,利用所得最优情感策略对应的最优情感状态对机器人情感状态转移概率进行更新,并依据6种基本情感状态在空间中的情感值,将其映射到连续情感空间中得到机器人的最优响应情感值。主客观对比实验表明,该文模型能有效增加机器人情感表达的细腻性、连续性、积极性以及共情性,还能有效降低机器人对外界情感刺激的依赖性,进一步提升和谐友好的人机交互关系。

关键词: PAD情感空间;强化学习;情感状态转移;认知情感生成

中图分类号: TP242.6

文献标识码: A

文章编号: 1009-5896(2021)06-1781-08

DOI: [10.11999/JEIT191035](https://doi.org/10.11999/JEIT191035)

Cognitive Emotional Interaction Model of Robot Based on Reinforcement Learning

HUANG Hongcheng^{①②} LI Jing^① HU Min^① TAO Yang^{*①②} KOU Lan^①

^①(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

^②(Chongqing Engineering Research Center of Communication Software, Chongqing 400065, China)

Abstract: In order to enhance the cognitive emotional computing ability of robot, a cognitive emotional interaction model of robot based on reinforcement learning is proposed, which combines immediate feedback and long-term trend according to PAD(Pleasure-Arousal-Dominance) emotional space. Firstly, according to the psychology theory of interpersonal communication, the human emotion generation process is simulated to generate human-like emotions, and the three influencing factors of similarity, positivity and empathy are extracted. Secondly, the relationship between the response emotion+ state and the contexted long-term emotion state is established by using the global co-ordination feature of reinforcement learning, so as to model the robot emotion generation process. Then, three factors are incorporated into the model reward mechanism for the evaluate of the interactive emotion state, to update the model and get the optimal emotional strategy. Finally, the optimal emotional state corresponding to the obtained optimal emotional strategy is used to update the robot's emotional state transition probability, and based on the sentiment values of the six basic emotional states in space, they are mapped to continuous emotional space to get the optimal response emotional value of the robot. Subjective and objective comparison experiments show that the model in this paper can effectively increase the delicateness, continuity, positivity and empathy of the robot's emotional expression, and can effectively

收稿日期: 2019-12-24; 改回日期: 2021-02-22; 网络出版: 2021-03-17

*通信作者: 陶洋 taoyang@cqupt.edu.cn

基金项目: 国家重点研发计划(2019YFB2102001), 国家自然科学基金(61871062)

Foundation Items: The National Key Research and Development Project (2019YFB2102001), The National Natural Science Foundation of China (61871062)

reduce the robot's dependence on external emotional stimuli, further improving the harmonious and friendly human-computer interaction.

Key words: Pleasure-Arousal-Dominance (PAD) emotion space; Reinforcement learning; Emotional state transfer; Cognitive emotion generation

1 引言

近年来,随着“智能家居”、“智慧社区”以及“智慧城市”等概念的提出与落实,人机交互成为公众日常生活中不可或缺的一部分。人们期望机器人在满足日常交互需求的同时,具备生成高级拟人化情感的认知情感计算能力。同时随着心理学、认知科学与人工智能交叉研究的深入,研究者发现机器智能应体现在“智商”与“情商”两个方面。因此,认知情感计算成为当前智能机器人研究领域中的热点。

认知情感计算就是要赋予计算机类似于人一样观察、理解和生成各种情绪状态的能力,使其能够像人一样进行自然亲切、生动有趣的交互^[1]。近年来,在认知情感计算方面,涌现出众多具有参考价值的情感模型。文献[2]提出基于PAD (Pleasure-Arousal-Dominance)的个性化情感模型,建立个性空间、心情空间与情感空间的3层映射关系来描述人类情感变化规律。文献[3]提出在云、边及用户协作下利用迁移学习进行情感分析,能有效分析用户的情感状态。文献[4]提出面向时序感知的多类别商品方面情感分析推荐模型,可推断用户在任意时间对商品的偏好。文献[5]提出基于指导性认知重评策略GCRs的情感交互模型,能降低机器人对外界情感刺激的依赖性,并在一定程度上促进机器人的积极情感表达。文献[6]提出多情感对话系统MECS,倾向在对话中产生连贯的情感反应,选择最相似情感作为机器人响应情感。文献[7]提出情感驱动的自私MANETS节点协商机制,模拟人类出价心理和情感变化提高节点出价竞争力。文献[8]提出ECM情绪聊天机,可以在内容上和情感一致性上产生适当响应。文献[9]提出生成对抗网络SentiGAN模型,在无监督情况下生成不同情绪标签通用的高质量情感文本。文献[10]提出基于句法约束的双向异步情感会话生成方法E-SCBA,将情感与主题引入解码增加回复响应的多样性。文献[11]提出融合强化学习与情感编辑约束的对话生成模型,能同时保证回复生成的流畅度与情感度。以上工作在一定程度上考虑了情感生成影响因素,但多为在“单轮交互模型”中加入影响情感生成的不同特征进行认知情感计算,未全面考虑上下文情境对当前情感状态生成的影响,或仅考虑了某种情感生成影响因素,

容易使机器人情感回应合理性不高、参与人丧失交互意愿。

因此,针对人机交互过程中机器人的情感生成问题,本文依据PAD 3维情感空间提出一种基于强化学习的机器人认知情感交互模型,试图利用强化学习全局统筹特性,建立上下文多轮情感状态与机器人当前情感响应之间的长期关联关系;利用强化学习奖励引导特性,实现对参与人进行情感支持、积极性引导以及情感共鸣的情感交互动机。通过考虑多轮多层次情感影响因素对人机交互过程中的情感生成过程建模以实现情感决策问题,得到机器人在连续多情感状态空间中的最优响应情感值。

2 人机交互情感分析

2.1 基于强化学习的认知情感计算

在人类情感生成过程中,个体情感状态响应不仅与外界情感刺激相关,还与自身情感状态和情感交互动机有关。进行情感状态响应时,不仅要考虑上下文多轮交互情境对当前情感状态转移概率的影响,还应考虑当前情感状态响应后续交互关系的影响。因此,为有效进行机器人情感策略学习,本文提出利用强化学习特性建立上下文多轮情感状态与当前响应情感状态之间的关联关系,对机器人进行认知情感计算,计算框架如图1所示。

2.2 基于情感空间的情感状态分析

为便于实现参与人情感状态跟踪,对交互输入内容进行情感量化与状态评估。本文首先依据文献[12]提供的数据与方法,对交互输入内容进行情感量化,得到其在PAD连续情感空间中对应的情感值 $\mathbf{E}_i=(p, a, d)$ 。其次,依据参考文献[13],对交互情感值向量 \mathbf{E}_i 进行状态评估,得到其在PAD连续情感空间内6种基本情感状态作用下的情感状态向量 $\mathbf{I}(\mathbf{E}_i)$ 。情感状态评估函数定义为

$$\mathbf{I}(\mathbf{E}_i) = [i_1, i_2, i_3, i_4, i_5, i_6] \quad (1)$$

$$\left. \begin{aligned} i_j &= \frac{1/h_j}{\sum_{j=1}^6 1/h_j}, & h_j &\neq 0 \\ i_1 &= 0, i_2 = 0, 1, \dots, i_j = 1, 2, \dots, i_6 = 0, \end{aligned} \right\} \quad (2)$$

$$h_j = (\mathbf{E}_i - \mathbf{E}_j)C_j(\mathbf{E}_i - \mathbf{E}_j)^T, \quad j = 1, 2, \dots, 6 \quad (3)$$

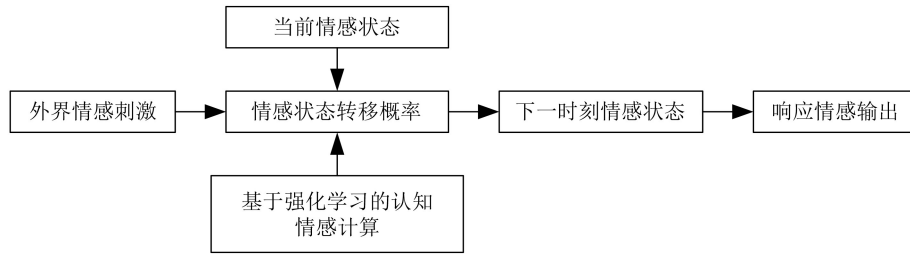


图1 机器人情感计算框架

其中， \mathbf{E}_i 表示交互输入情感值； $j = 1, 2, \dots, 6$ 分别表示高兴、惊讶、厌恶、生气、恐惧、悲伤6种基本情感状态； \mathbf{E}_j 表示基本情感 j 对应的情感值； \mathbf{C}_j 表示基本情感 j 聚类区域的协方差矩阵； h_j 表示 \mathbf{E}_i 到 \mathbf{E}_j 之间的距离； i_j 则表示 \mathbf{E}_i 在 \mathbf{E}_j 作用下的情感状态评估值。

3 基于强化学习的认知情感交互模型

3.1 强化学习定义

强化学习模型原理为：一个智能体(agent)，在当前状态(state)下，执行一个行为(action)与环境(environment)进行交互并进入一个新的状态，同时从环境中获得相应的即时奖励(reward)，再根据奖励评估此行为，利于目标实现的行为其奖励值增加，不利于目标实现的行为奖励值衰减，此过程不断循环到终止状态为止。

3.1.1 状态

状态 s 表示智能体所处的情感状态，通常由外部环境给出。为减小情感划分粒度，增加机器人情感表达连续性与细腻性，本文将含有151种情感状态的PAD连续情感空间作为智能体的情感状态空间，将空间中各情感状态在6种基本情感状态作用下的情感状态向量 $\mathbf{I}(\mathbf{E}_i)$ 作为可能的交互输入响应情感状态。

3.1.2 行为

行为 a 表示智能体在交互响应过程中，选择下一轮响应情感状态时执行的一个动作，其搜索空间为情感空间大小。智能体在情感空间中的活动过程即情感空间各情感状态间的马尔可夫转移过程。

3.1.3 折损因子

折损因子 γ 可在环境具有随机性的情况下，用于计算状态序列累积奖励的未来奖励衰减。本文考虑在距离当前会话越远的将来时刻，未来奖励对用于衡量下一轮会话情感状态的满意度的影响越小。其值介于0~1之间，考虑对未来奖励的重视程度越大， γ 值越大；反之， γ 值越小。

3.1.4 奖励

奖励 r 可在智能体执行相应动作 a 后，用于衡量

所获得情感状态的未来满意度。人机交互双方在交互过程中均存在一定的情感动机^[4]。因此，依据社会心理学中人际吸引原则将机器人的交互情感动机设定为能在一定程度上实现对参与人的情感肯定、情绪引导与情感共鸣，并据此对情感奖励函数进行构建。

相似性(similarity)：考虑人际交往过程中，人们往往希望对方能与自己产生相似的情感反应，即期望对方能够“乐他人之乐，忧他人之忧”。因此，为实现对参与人的情感肯定，本文依据文献^[14]，计算余弦相似度来度量情感状态向量间的相似性

$$r_1 = S(\mathbf{E}_{k+1}, \mathbf{E}_k) = \frac{\mathbf{I}(\mathbf{E}_{k+1}) \cdot \mathbf{I}(\mathbf{E}_k)}{\|\mathbf{I}(\mathbf{E}_{k+1})\| \|\mathbf{I}(\mathbf{E}_k)\|} \quad (4)$$

积极性(positive)：考虑人际交往过程中，人们会通过调整自身情感表达状态对他人的某种情绪引导。因此，为实现对参与人的情绪引导，本文通过设定机器人情感积极性引导来增加参与人交互意愿。实际上情绪引导并非积极度越高越好，尤其在参与人情绪比较消极时可能会适得其反。而积极性与相似性协同作用，恰好能有效解决引导过度问题。因此，本文对响应情感状态向量进行积极度计算

$$r_2 = P(\mathbf{E}_{k+1}) = P(\mathbf{I}(\mathbf{E}_{k+1})) = \sum_{j=1}^6 l_j i_j \quad (5)$$

共情性(empathy)：考虑人际交往过程中，人际吸引不仅与个体间相似性有关，还受彼此互补关系影响。受互补关系影响主要表现在人们有时会倾向于喜欢那些与自己能在某方面产生互补的人。在情感交互中，可以理解为期望对方具备“同理心”，与自己在情感表达上产生共鸣关系。因此，本文通过计算情感状态向量之间的相互关系来度量情感共情性

$$r_3 = M(\mathbf{E}_k, \mathbf{E}_{k+1}) = \frac{1}{1 + \text{rank}(\mathbf{E}_{k+1})} \log_2 P(\mathbf{I}(\mathbf{E}_k) | a) + \frac{1}{1 + \text{rank}(\mathbf{E}_k)} \log_2 P(a | \mathbf{I}(\mathbf{E}_k)) \quad (6)$$

其中， $P(a | \mathbf{I}(\mathbf{E}_k))$ 表示智能体在输入情感状态下

选择响应情感状态的转移概率； $P(\mathbf{I}(\mathbf{E}_k)|a)$ 表示由响应情感状态选择输入情感状态的后向转移概率； $\text{rank}(\mathbf{E}_k)$ 与 $\text{rank}(\mathbf{E}_{k+1})$ 分别表示情感状态 \mathbf{E}_k 与 \mathbf{E}_{k+1} 之间反向转移概率排名与正向转移概率排名，转移概率越大，排名越高。依据文献[2]，本文通过利用情感空间中各情感状态间转移概率与状态间欧氏距离成反比，对情感状态间的初始转移概率进行计算。

对于一个动作 a ，其获得的最终奖励为以上3个奖励衡量指标的加权和

$$R(a|\mathbf{I}(\mathbf{E}_k)) = \alpha_1 r_1 + \alpha_2 r_2 + \alpha_3 r_3 \quad (7)$$

本文参数权重设置为 $\alpha_1=0.4$, $\alpha_2=0.3$, $\alpha_3=0.3$ 。

3.1.5 策略(policy)

策略 P 用于表示在当前状态下，智能体选择下一情感状态时对应的概率分布，可用公式表示为 $\pi(a|s) = P_{\text{RL}}(\mathbf{I}(\mathbf{E}_{k+1})|\mathbf{I}(\mathbf{E}_k))$ ，初始值为情感状态间的初始转移概率。本文采用策略梯度算法对模型进行优化，因此其值与选择下一情感状态可获得的未来奖励值相关，获得未来奖励值大的动作出现概率就大，对应地获得未来奖励值小的动作出现概率则小。

3.1.6 模型优化

本文通过策略梯度算法将策略参数化实现模型更新训练，目的是通过优化模型参数 θ 使未来累积奖励期望值达到最大。因此，目标函数为最大化未来奖励的期望值，定义为

$$L_{\text{RL}}(\theta) = E_{\text{RH}(a_1:T)} \left[\sum_{k=1}^T R(a_k, \mathbf{I}(\mathbf{E}_k)) \right] \quad (8)$$

其中， $R_k(a_k, \mathbf{I}(\mathbf{E}_k))$ 表示在状态 $\mathbf{I}(\mathbf{E}_k)$ 下执行动作 a_k 获得的奖励值；再采用似然比技巧进行梯度更新

$$\nabla_{\theta} L_{\text{RL}}(\theta) = \sum_{k=1}^T \nabla_{\theta} \log_2 P(a_k, \mathbf{I}(\mathbf{E}_k)) R(a_k, \mathbf{I}(\mathbf{E}_k)) \quad (9)$$

最后利用所求得的梯度值对参数 θ 进行更新

$$\theta_{\text{new}} = \theta_{\text{old}} + \beta \nabla_{\theta} L_{\text{RL}}(\theta) \quad (10)$$

累积奖励期望值达到最大时，所得最优策略对应的情感状态为交互输入的最优响应情感状态。

3.1.7 情感交互过程模拟

本文利用两个智能体进行交互以模拟智能体与外界环境的情感交互过程：智能体1将初始交互输入情感 \mathbf{E}_1 通过情感评估为情感状态向量 $\mathbf{I}(\mathbf{E}_1)$ 后将其传送给智能体2，然后智能体2将得到的交互响应情感 \mathbf{E}_2 同样评估为情感状态向量 $\mathbf{I}(\mathbf{E}_2)$ 再回复给智能体1，不断重复此过程至模拟的最大交互轮数。

交互目标是在当前交互输入情感状态下能够选择获得未来奖励最多的最优情感状态。智能体之间的情感交互过程如图2所示。

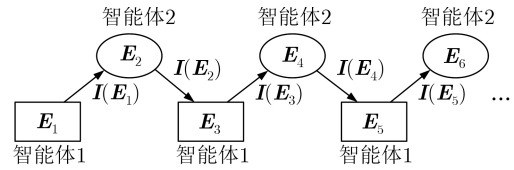


图2 情感交互过程

依据文献[15]，利用情感空间中情感状态间距离的远近来映射情感类别间的相似性。距离越近，相似性越大，情感状态间转移概率越大；距离越远，相似性越小，情感状态间转移概率越小。并且，某一情感状态到情感空间中其他各情感状态的转移概率之和为1^[16]。因此，为便于对情感刺激的响应情感状态进行计算，本文将空间中与外界情感刺激点欧氏距离最近的前 n 种情感状态作为智能体每轮交互的候选情感状态。

3.2 机器人情感状态更新

依据参考文献[17]，本文利用机器人受 k 轮交互参与人输入的外界情感 \mathbf{E}_{HR}^k 刺激后得到的 $k+1$ 轮最优响应情感状态向6种基本情感状态转移的概率 P_{RH}^{k+1} 和6种基本情感状态在空间中的坐标值 (p_j, a_j, d_j) 对 $k+1$ 轮机器人响应情感值 $\mathbf{E}_{\text{RH}}^{k+1}$ 在空间中的坐标位置进行标定，实现机器人在连续情感空间中的情感状态转移。首先，假设强化学习模型所得最优响应情感状态向量对应策略为 p ，可以得到依据参与人交互输入响应情感状态向6种基本情感状态转移的概率为

$$\begin{aligned} P(\mathbf{E}_{k+1}|\mathbf{E}_{\text{HR}}^k) &= \{P(1), P(2), \dots, P(6)\} \\ &= [i_{p1}, i_{p2}, \dots, i_{p6}] \end{aligned} \quad (11)$$

其次，结合 $k-1$ 轮机器人的情感状态转移概率 P_{R}^{k-1} 和 k 轮交互参与人输入最优响应情感状态的转移概率 $P(\mathbf{E}_{k+1}|\mathbf{E}_{\text{HR}}^k)$ 对 $k-1$ 轮机器人响应情感状态转移概率 P_{RH}^{k-1} 进行更新

$$\left. \begin{aligned} P_{\text{R}}^{k+1}(j) &= P_{\text{RH}}^{k-1}(j) + c_j P(j) \\ c_j &= \frac{P(j) - \min \{P(\mathbf{E}_{k+1}|\mathbf{E}_{\text{HR}}^k)\}}{\max \{P(\mathbf{E}_{k+1}|\mathbf{E}_{\text{HR}}^k)\} - \min \{P(\mathbf{E}_{k+1}|\mathbf{E}_{\text{HR}}^k)\}} \end{aligned} \right\} \quad (12)$$

其中， c_j 表示交互输入响应情感状态转移到6种基本情感状态的置信度。再将所得转移概率 P_{R}^{k+1} 归一化得到 $k+1$ 轮交互机器人响应情感状态的转移概率为

$$\left. \begin{aligned} P_{k+1}(j) &= \frac{P_{RH}^{k+1}(j)}{\sum_{j=1}^6 P_{RH}^{k+1}(j)} \\ \mathbf{P}_{RH}^{k+1} &= [P_{k+1}(1), P_{k+1}(2), \dots, P_{k+1}(6)] \end{aligned} \right\} \quad (13)$$

最后，依据所得机器人情感状态转移概率 P_{RH}^{k+1} 对 $k+1$ 轮机器人最优响应情感值 E_{RH}^{k+1} 在情感空间中的坐标位置 $(p_{k+1}, a_{k+1}, d_{k+1})$ 进行标定，计算公式为

$$\left. \begin{aligned} p_{k+1} &= \sum_{j=1}^6 p_j P_{k+1}(j) \\ a_{k+1} &= \sum_{j=1}^6 a_j P_{k+1}(j) \\ d_{k+1} &= \sum_{j=1}^6 d_j P_{k+1}(j) \end{aligned} \right\} \quad (14)$$

$$\mathbf{E}_{RH}^{k+1} = (p_{k+1}, a_{k+1}, d_{k+1}) \quad (15)$$

3.3 交互模型构建

本文输入为 k 轮交互参与人的输入情感值 \mathbf{E}_{HR}^k 以及 $k-1$ 轮交互机器人的情感状态转移概率 \mathbf{P}_{RH}^{k-1} ，输出为 $k+1$ 轮交互机器人的响应情感值 \mathbf{E}_{RH}^{k+1} 。依据PAD情感空间对输入内容进行情感的量化评估，并利用强化学习对机器人的情感生成过程建模，得到机器人最优响应情感。模型构建过程如表1所示。

4 实验与分析

4.1 实验设计

为便于对文本所提认知情感交互模型进行性能

分析与对比实验，依据文献[18]利用开源聊天机器人ChatterBot构建基于本文认知情感模型的文本聊天机器人。首先，利用聊天机器人逻辑适配器进行答案匹配，返回置信度较高的前 m 个答案作为候选答案集；然后，利用本文模型进行情感策略评估，选择最优情感策略。最后，依据本文模型响应情感对候选答案进行最优排序，并选择排序等级最高的答案作为机器人响应输出。此外，由于需要探索的情感状态数会随着交互轮数的增加呈指数增长，本文模型在进行情感状态评估时，设置两个智能体的最大交互轮数 $T=8$ (轮)，每轮候选情感状态选取数 $n=8$ (种)。

实验数据采用NLPCC2017共享任务Emotional Conversation Generation中的样本数据集，此数据集共包含1119207个问答对，随机划分8000个问答对作为验证集，5000个问答对作为测试集，剩余问答对用作聊天机器人的中文训练语料。

实验主要围绕情感准确度与人机交互会话实际效果展开，因此选取以下认知模型进行对比实验：

文献[18]提出机器人认知模型Chatterbot，根据候选答案集中各答案置信度高低进行输出响应。由于其不具备认知情感计算能力，只用于模型有效性验证对比实验；文献[8]提出情绪聊天机ECM，可以在内容相关语法和情绪一致性上产生适当的响应；文献[9]提出生成对抗网络SentiGAN模型，能够生成通用的、多样化的、高质量的情感文本；文献[10]提出双向异步情感会话生成方法E-SCBA，能够生成具有逻辑性和情感度的文本；

文献[5]提出基于指导性认知重评策略GCRs的情感交互模型，能够降低机器人对外界情感刺激的

表1 基于强化学习的机器人认知情感交互模型

| |
|---|
| 输入： k 轮交互参与人的输入情感值 \mathbf{E}_{HR}^k ； $k-1$ 轮交互机器人的情感状态转移概率 \mathbf{P}_{RH}^{k-1} ； |
| 输出： $k+1$ 轮交互机器人的响应情感值 \mathbf{E}_{RH}^{k+1} ； |
| 开始： |
| 参与人输入交互情感 \mathbf{E}_{HR}^k ； |
| 根据式(1)—式(3)对 \mathbf{E}_{HR}^k 进行情感状态评估得到 $\mathbf{I}(\mathbf{E}_{HR}^k)$ ； |
| 根据式(4)—式(7)计算在当前情感状态 $\mathbf{I}(\mathbf{E}_{HR}^k)$ 下，机器人执行某一动作产生下一情感状态时可获得的奖励值 $R(a \mathbf{I}(\mathbf{E}_{HR}^k))$ ； |
| 根据式(8)—式(10)对模型进行训练，通过实现累积奖励值期望最大化来获得最优参数，从而得到最优情感策略 p ； |
| 通过所选最优情感策略 p ，依据式(11)对交互输入响应情感状态转移概率 $P(\mathbf{E}_{k+1} \mathbf{E}_{HR}^k)$ 进行计算； |
| 通过获得的 $\mathbf{P}(\mathbf{E}_{k+1} \mathbf{E}_{HR}^k)$ ，依据式(12)—式(13)对机器人 $k-1$ 轮交互情感状态转移概率 \mathbf{P}_{RH}^{k-1} 进行更新，得到 $k+1$ 轮交互机器人的情感状态转移概率 \mathbf{P}_{RH}^{k+1} ； |
| 通过获得的 \mathbf{P}_{RH}^{k+1} ，依据式(14)—式(15)对最优响应情感值 \mathbf{E}_{RH}^{k+1} 的空间坐标位置进行标定，实现机器人在连续情感空间中的状态转移； |
| 机器人输出响应情感 \mathbf{E}_{RH}^{k+1} ； |
| 令 $k = k + 2$ ； |
| 直到参与人停止交互输入 |
| 人机交互结束 |

依赖性，并在一定程度上促使机器人的积极情感表达。其中ECM, SentiGAN和E-SCBA均为考虑一定情感因素的生成式聊天机器人模型，本文在进行情感准确度、信息检索有效性验证时需要将其作用下聊天机器人的响应文本量化为情感状态向量。

4.2 情感准确度分析

为避免机器人情感表达含糊不清使得参与人对响应情感状态识别困难，响应情感状态在预期情感类别的表达上应具备一定准确度。为直观对各模型作用下机器人情感生成状态的准确性进行评估，依据文献[11]，对响应情感的目标情感类别准确度进行计算

$$\text{Acc}(\mathbf{E}_{\text{RH}}^{k+1}) = P_{k+1}(i) - \frac{1}{5} \sum_{j \neq i} P_{k+1}(j) \quad (16)$$

$$i, j = 1, 2, \dots, 6$$

从测试集中随机划分出含有多个情感类别的100句对各模型进行响应情感状态准确率计算，结果见表2。由表2可见，本文模型在情感准确度方面均好于其他模型，这主要是由于本文在进行机器人情感状态转移概率更新时，将输入响应情感状态到各基本情感状态转移概率的置信度作为更新因子，有效地增加了输入响应预期情感类别对机器人情感状态转移概率的影响。

表2 不同模型情感准确度统计表

| 认知模型 | 准确度 |
|----------|-------|
| ECM | 0.775 |
| GCRs | 0.831 |
| E-SCBA | 0.792 |
| SentiGAN | 0.846 |
| 本文 | 0.865 |

4.3 模型有效性验证

4.3.1 信息检索有效性度量

为便于对模型答案检索有效性进行验证，依据文献[19]采用两个信息检索评价指标MRR(Mean Reciprocal Rank)和MAP(Mean Average Precision)对各个模型候选答案进行排序准确率计算，从测试集中随机选取60句进行试验，取排序准确率平均值作为实验最终结果，结果见表3。

表3为对不同认知模型答案($m=6$)排序平均准确率的统计结果，由表可见本文模型与其他模型相比取得了令人相对满意的结果。这是由于本文模型在对候选答案进行排序时，通过结合上下文情感状态的量化评估与类人情感状态影响因素的分析量化，利用强化学习建立上下文长期情感状态之间的

表3 不同模型排序准确率统计表

| 认知模型 | MAP | MRR |
|------------|-------|-------|
| Chatterbot | 0.466 | 0.438 |
| ECM | 0.608 | 0.587 |
| GCRs | 0.641 | 0.623 |
| E-SCBA | 0.634 | 0.625 |
| SentiGAN | 0.637 | 0.628 |
| 本文 | 0.657 | 0.646 |

关联关系，以实现对下文状态响应的综合最优评定，具备较好的认知情感能力。

4.3.2 交互会话有效性验证

为对交互会话有效性进行有效评估，本文邀请20位志愿者参与不同模型下的多次人机交互。同时，为增加模型间客观对比性，各模型每人均进行30次多轮人机交互会话实验。并从测试集中随机选取30句依次作为各模型中参与人进行交互会话的初始输入，统计各模型每次进行人机交互的会话轮数与交互时间。实验所得不同模型下平均会话轮数与平均交互时间统计结果见表4。

表4 会话轮数与交互时间统计表

| 认知模型 | N(轮) | T(s) |
|------------|------|--------|
| Chatterbot | 6 | 68.56 |
| ECM | 9 | 97.30 |
| GCRs | 13 | 132.84 |
| E-SCBA | 11 | 114.57 |
| SentiGAN | 10 | 107.29 |
| 本文 | 15 | 145.41 |

由表4可见，在平均会话轮数与平均交互时间上本文模型均优于其他模型，说明本文模型作用下的聊天机器人更不容易使聊天陷入尴尬境地，能有效延长人机交互会话时间。这是由于本文模型在多情感状态连续空间中考虑类人情感生成并结合机器人自身情感状态更新得到的响应情感的多样性更丰富、积极性与准确度更高，有效地引导了参与人参与人机交互。

4.4 模型满意度评估

为对模型满意度进行有效评估，本文从单轮对话主观满意度、多轮会话主观满意度两个方面进行问卷调查实验。单轮对话主观满意度评价指标为合理性、多样性、共情度。实验过程为：从测试集中随机选取100句用于测试，实验共计使用500个问答对，多渠道邀请200名志愿者进行线上线下问卷调查；多轮会话主观满意度评价指标为流畅度、积极

度、有趣度、参与度，具实验过程为：依据评价指标对交互会话有效性验证中的20位人机交互志愿者进行多轮会话满意度调查。同时，所有指标均采用三点量表(0,1,2)进行评估：0表示程度较低，1表示程度一般，2表示程度较高。最终统计结果取平均值，得分越高模型满意度越高。模型单轮对话主观满意度调查结果见图3，多轮会话主观满意度调查结果见图4。

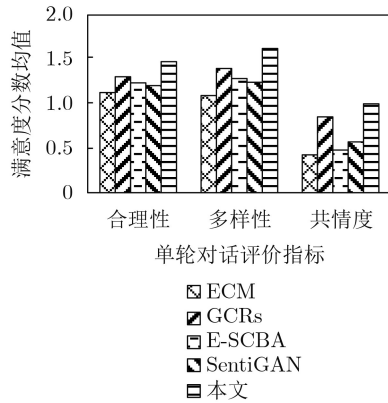


图3 单轮对话主观评估数据统计图

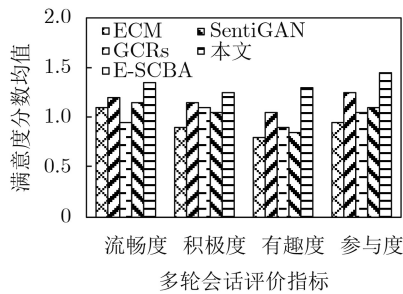


图4 多轮会话主观评估数据统计图

由图3可见，本文模型在对话合理性、多样性以及共情度上均明显优于其他模型，尤其在情感表达多样性上获得了很好的效果，这是由于本文在进行情感决策时充分利用了情感空间中的多种情感状态，结果表明本文模型能从多方面有效提升机器人单轮对话响应满意度。由图4可见，本文模型在机器人情感表达整体流畅度与积极度上、人机交互有趣度以及参与人参与度上较其他模型均取得有效提升，说明本文在进行情感交互模型构建时，建立的上下文长期依赖关系与考虑到的情感生成影响因素合理有效，能进一步增加参与人的人机交互意愿、构建自然和谐的人机交互关系。

5 结束语

本文提出一种基于强化学习的机器人认知情感交互模型，首先，利用强化学习对情感生成过程建模，将PAD情感空间作为机器人的情感状态空

间，情感划分粒度小，表达细腻；其次，考虑将相似性、积极性与共情性3个情感影响因素量化为进行情感状态评估的奖励函数，实现对参与人进行情感支持、情绪引导、情感共鸣的交互动机；最后，结合最优情感状态对机器人情感状态转移概率进行更新，从而进一步得到机器人在情感空间中的坐标位置，实现机器人在连续情感空间中的状态转移。实验从准确性、MAP和MRR等方面验证了模型有效性。由于人类情感生成过程具有复杂性、情感状态转移概率影响因素具有多样性，而本文模型只考虑了情感生成与状态转移过程中的部分影响因素。因此，未来工作还需全面考虑人类情感生成与状态转移过程中的影响因素以进一步优化类人情感状态生成。

参考文献

- [1] ZUCCO C, CALABRESE B, and CANNATARO M. Sentiment analysis and affective computing for depression monitoring[C]. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, USA, 2017: 1988–1995. doi: [10.1109/BIBM.2017.8217966](https://doi.org/10.1109/BIBM.2017.8217966).
- [2] 杨勇, 张志瑜. 基于PAD的个性化情感模型[J]. 重庆邮电大学学报: 自然科学版, 2012, 24(1): 96–103. doi: [10.3979/j.issn.1673-825X.2012.01.019](https://doi.org/10.3979/j.issn.1673-825X.2012.01.019).
YANG Yong and ZHANG Zhiyu. Personalized affective model based on PAD[J]. *Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition*, 2012, 24(1): 96–103. doi: [10.3979/j.issn.1673-825X.2012.01.019](https://doi.org/10.3979/j.issn.1673-825X.2012.01.019).
- [3] WU Dapeng, HAN Xiaojuan, YANG Zhigang, et al. Exploiting transfer learning for emotion recognition under cloud-edge-client collaborations[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(2): 479–490. doi: [10.1109/JSAC.2020.3020677](https://doi.org/10.1109/JSAC.2020.3020677).
- [4] 丁永刚, 李石君, 付星, 等. 面向时序感知的多类别商品方面情感分析推荐模型[J]. 电子与信息学报, 2018, 40(6): 1453–1460. doi: [10.11999/JEIT170938](https://doi.org/10.11999/JEIT170938).
DING Yonggang, LI Shijun, FU Xing, et al. Temporal-aware multi-category products recommendation model based on aspect-level sentiment analysis[J]. *Journal of Electronics & Information Technology*, 2018, 40(6): 1453–1460. doi: [10.11999/JEIT170938](https://doi.org/10.11999/JEIT170938).
- [5] LIU Xin, XIE Lun, and WANG Zhiliang. Empathizing with emotional robot based on cognition reappraisal[J]. *China Communications*, 2017, 14(9): 100–113. doi: [10.1109/CC.2017.8068769](https://doi.org/10.1109/CC.2017.8068769).
- [6] ZHANG Rui, WANG Zhenyu, and MAI Dongcheng. Building emotional conversation systems using multi-task Seq2Seq learning[C]. The 6th CCF International Conference

- on Natural Language Processing and Chinese Computing, Dalian, China, 2018: 612–621. doi: [10.1007/978-3-319-73618-1_51](https://doi.org/10.1007/978-3-319-73618-1_51).
- [7] 杨杨, 邱雪松, 孟洛明, 等. 情感驱动的自私MANETs节点协商机制[J]. 电子与信息学报, 2011, 33(6): 1294–1300. doi: [10.3724/SP.J.1146.2010.01072](https://doi.org/10.3724/SP.J.1146.2010.01072).
YANG Yang, QIU Xuesong, MENG Luoming, *et al.* An emotion-driven negotiation mechanism of selfish nodes in the MANETs[J]. *Journal of Electronics & Information Technology*, 2011, 33(6): 1294–1300. doi: [10.3724/SP.J.1146.2010.01072](https://doi.org/10.3724/SP.J.1146.2010.01072).
- [8] ZHOU Hao, HUANG Minlie, ZHANG Tianyang, *et al.* Emotional chatting machine: Emotional conversation generation with internal and external memory[C]. The 32nd AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th Innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, USA, 2018: 730–738.
- [9] WANG Ke and WAN Xiaojun. SentiGAN: Generating sentimental texts via mixture adversarial networks[C]. The 27th International Joint Conference on Artificial Intelligence, Stockholm, Sverige, 2018: 4446–4452. doi: [10.24963/ijcai.2018/618](https://doi.org/10.24963/ijcai.2018/618).
- [10] LI Jingyuan and SUN Xiao. A syntactically constrained bidirectional-asynchronous approach for emotional conversation generation[C]. 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 2018: 678–683. doi: [10.18653/v1/D18-1071](https://doi.org/10.18653/v1/D18-1071).
- [11] 孙晓, 李佳, 卫星. 基于强化学习的情感编辑约束对话内容生成[J/OL]. 自动化学报. doi: [10.16383/j.aas.c190058](https://doi.org/10.16383/j.aas.c190058).
SUN Xiao, LI Jia, and WEI Xing. Emotional editing constraint conversation generation based on reinforcement learning[J/OL]. *Acta Automatica Sinica*. doi: [10.16383/j.aas.c190058](https://doi.org/10.16383/j.aas.c190058).
- [12] RUSSELL J A and MEHRABIAN A. Evidence for a three-factor theory of emotions[J]. *Journal of Research in Personality*, 1977, 11(3): 273–294. doi: [10.1016/0092-6566\(77\)90037-X](https://doi.org/10.1016/0092-6566(77)90037-X).
- [13] PARK J W, KIM W H, LEE W H, *et al.* How to completely use the PAD space for socially interactive robots[C]. 2011 IEEE International Conference on Robotics and Biomimetics, Karon Beach, Thailand, 2011: 3005–3010. doi: [10.1109/ROBIO.2011.6181762](https://doi.org/10.1109/ROBIO.2011.6181762).
- [14] 邵俊健. 高维数据的聚类算法及其距离度量的研究[D]. [硕士学位论文], 江南大学, 2019.
SHAO Junjian. Research on clustering algorithm of high dimensional data and its Disatance metric[D]. [Master dissertation], Jiangnan University, 2019.
- [15] 孙佩宏, 陶霖密. PAD情感空间中情感距离度量方法[C]. 第四届和谐人机环境联合学术会议论文集, 武汉, 2008: 642–649.
SUN Peihong and TAO Linmi. Emotional distance measurement method in PAD emotion space[C]. The 4th Harmonious Human-Computer Environment Joint Academic Conference, Wuhan, China, 2008: 642–649.
- [16] 吴伟国, 李虹漫. PAD情感空间内人工情感建模及人机交互实验[J]. 哈尔滨工业大学学报, 2019, 51(1): 29–37. doi: [10.11918/j.issn.0367-6234.201807138](https://doi.org/10.11918/j.issn.0367-6234.201807138).
WU Weiguo and LI Hongman. Artificial emotion modeling in PAD emotional space and human-robot interactive experiment[J]. *Journal of Harbin Institute of Technology*, 2019, 51(1): 29–37. doi: [10.11918/j.issn.0367-6234.201807138](https://doi.org/10.11918/j.issn.0367-6234.201807138).
- [17] HUPONT I, BALDASSARRI S, and CERESO E. Facial emotional classification: From a discrete perspective to a continuous emotional space[J]. *Pattern Analysis and Applications*, 2013, 16(1): 41–54. doi: [10.1007/s10044-012-0286-6](https://doi.org/10.1007/s10044-012-0286-6).
- [18] GUNTHER C. ChatterBot tutorial[EB/OL]. <https://chatterbot.readthedocs.io/en/stable/tutorial.html>, 2019.
- [19] WU Yu, WU Wei, XING Chen, *et al.* Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots[C]. The 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada, 2017: 496–505.
- 黄宏程: 男, 1979年生, 副教授, 研究方向为认知情感计算、复杂网络与信息传播理论.
- 李 净: 女, 1995年生, 硕士生, 研究方向为认知情感计算.
- 胡 敏: 女, 1971年生, 副教授, 研究方向为信息通信网络体系结构、人机交互理论与技术应用.
- 陶 洋: 男, 1964年生, 教授, 研究方向为人工智能、大数据与计算智能.
- 寇 兰: 女, 1963年生, 副教授, 研究方向为D2D通信、人机交互理论与技术应用.

责任编辑: 马秀强