

基于动态 Q 学习算法的情感自动机模型研究 *

于冬梅 方建安

(东华大学信息科学与技术学院 上海 201620)

摘要 让计算机具有认知和表达自己情感的能力,培养其智能地对人类的情感做出反应是现阶段信息科学研究的热点内容。本文构建了基于动态 Q 学习算法的情感自动机模型,在该模型中,定义了情感元的概念,情感元应用动态 Q 学习算法来预测并感知环境的变化,从而改变自身情感来适应周围环境。

关键词 动态 Q 学习算法,情感自动机,情感元

Research on a Model of Emotion-automaton Based on Dynamic Q-learning Arithmetic

YU Dong-mei FANG Jian-an

(School of Information Science and Technology, Donghua University, Shanghai 201620, China)

Abstract Letting computers have the ability of cognizing, expressing their emotions and training them to act human emotions, is becoming hotspot of recent research. This paper, designs a model of emotion-automaton based on dynamic Q-learning arithmetic, in which defines an emotional unit. The emotional unit will change its emotion to forecast and apperceive the variety circumstance by dynamic Q-learning arithmetic, then change its action to adapt the circumstance.

Keywords Dynamic Q-learning arithmetic, Emotion-automaton, Emotional unit

情感在人们智能活动中扮演着重要的、不可缺少的角色。医学以及心理学研究表明:人们若丧失了一定的理解和表达情感的能力,他们的决策能力会遇到严重的障碍^[1]。情感能力与理性思维和逻辑推理能力不是绝对矛盾的,是可以相辅相成的。“情感计算”这一概念是由 MIT 的 Picard 教授在她的《Affective Computing》^[2]一书中最先提出来的。情感计算的研究重点就在于创建一个具有感知、识别和理解人类情感的能力,并能针对用户的情感做出理智、灵敏、友好反应的个人计算系统。基于情感具有“传染性”这一特性^[3,4],本文应用动态 Q 学习算法,构建了一个情感自动机模型,在模型的构建中,首先介绍了 Q 学习算法,然后提出了改进的 Q 学习算法,最后建立了情感自动机模型。

1 动态 Q 学习算法

1.1 动态 Q 学习算法原理

Q 学习就是一种典型的强化学习算法。本部分内容设计了一个个体与环境交互或者与环境中的其他个体交互从而改变自身情感状态的动态 Q 学习算法,根据当前更新了的环境得出个体的情感估计值 Q 。该方法采用的是估计即将发生在个体上的情感信息而非 Q 函数,因而就不用观察下一个兴趣项的实际回报值和它的 Q 学习参数,从而提高了学习效率。

这里采用联合策略(a_{this}, a_{other})对个体的情感信息—— Q 值进行估计,其中 a_{this} 表示该个体此刻所具有的情感信息(即个体的初始情感状态,假设其已知),而 a_{other} 表示个体与环境或者其他个体交互后的情感状态。

当个体在新环境状态 S (即情感状态与初始状态可以一致也可以不一致)下,学习 Agent(一种智能体)对个体行为进行分析,估计用户的下一情感状态信息应变为,基于随机策略

选择用户当前的情感状态 a_{this} :

$$prb[a_{this} | S] = \frac{e^{\bar{Q}(s, a_{this})/T}}{\sum_{a_k \in A} e^{\bar{Q}(s, a_k)/T}} \quad (1)$$

其中, $\bar{Q}(s, a_{this})/T$ 是基于状态 S 下的信念 P_s 的期望 Q 函数值,即:

$$\begin{aligned} \bar{Q}(s, a_{this}) &\equiv \langle Q(s, a_{this}, a_{other}) \rangle P_s \\ &= \int_B (Q(s, a_{this}, a_{other}) \times P_s) dB \end{aligned} \quad (2)$$

在时刻 t ,环境改变到新的状态 s' (即改变了环境或者遇到了其他的个体),得到行动的奖励值 γ_t ,智能体基于下面的公式更新 Q 的值:

$$Q_t(s, a_{this}, a_{other}^*) = (1 - \alpha_t) Q_{t-1}(s, a_{this}, a_{other}^*) + \alpha_t (\gamma_t + \max_{a'_{this} \in A} Q_{t-1}(s', a'_{this}, a_{other}^*)) \quad (3)$$

其中, $a_{other}^* = \arg \max_{a_{other} \in A} P_{s'}(a_{other})$ 为个体的实际情感状态, $\alpha_t \in [0, 1]$, 是学习率。 α_t 随着时间衰减,以利于学习算法的收敛。

在不断与环境的交互过程中,个体的历史经验可能会因为环境改变而过时,同时随着学习进程的进行,个体所取得的知识趋于精确(表现为 Q 值收敛于 Q^*)。如果这时仍然进行大量的探索活动,势必造成系统的性能下降。所以,在学习 Agent 与个体充分作用之后,适当的减少探索是必要的。这里将最近探索盈余(Recency Exploration Bonus)引入到 Q 学习中:

$$Q'_t(s, a_{this}, a_{other}^*) = Q_t(s, a_{this}, a_{other}^*) + 6\lambda (\rho_t(s, a_{this}, a_{other}^*)) \sqrt{\rho_t(s, a_{this}, a_{other}^*)} \quad (4)$$

其中, ρ 是等待时间的盈余, $b(\rho) = 6\lambda(\rho)\sqrt{\rho}$, 是探索的盈余。

这样,学习 Agent 就近探索那些即将发生的情感,并准备适应环境状态改变而做任意改变。

* 基金项目: 国家自然科学基金资助项目(60674088)。于冬梅 博士生,主研方向为模式识别与智能控制、情感计算、语音识别;方建安 教授,博导,主研方向为智能控制、情感计算、混沌理论。

1.2 动态 Q 学习算法步骤

一般在设计 Q 学习算法时,首先需要考虑:如何确定状态空间 S 、动作空间 A 以及回报值 r 。

定义 1 状态 S 定义为学习 Agent 接收到的个体的情感状态,它是一个 n 元组:

$$S = (x_1, x_2, \dots, x_j, \dots, x_n)$$

其中, x_j 是个体情感状态的取值。所有状态 s 构成的集合为状态空间 S 。

定义 2 情感定义为学习 Agent 推断出的个体情感 j 的取值,即 Agent 得到了个体的新的情感状态,所有的情感元素构成了情感状态空间 A 。

定义 3 回报值 r 定义为:

$$r = \sum_{j \in J} W_j v(x_j) \quad (5)$$

其中, $v(x_j)$ 表示情感 j 的值为 x_j 。

估计个体下一情感状态的动态 Q 学习算法:

①初始化: $\forall s \in S, \forall a \in A, Q_0(s, a) = 1, a_0 = 1, \gamma = 0.5$;

②在当前环境状态 s 下,对所有可能的情感状态 a ,基于信念 p ,根据公式(1),学习 Agent 将选择当前个体的情感状态 a_{this} ;

③执行步骤②中选择的 a_{this} ;

④在时刻 t ,个体与环境交互或者与环境中的其他个体进行交互,记为状态 s' ,该个体的情感状态发生变化,并变为 a_{other} ,根据公式(4),Agent 计算回报值 γ ,并修改当前学习率 α_t :

$$\alpha_t = \frac{\alpha_0}{n_t(s, a_{this}, a_{other})}$$

其中, $n_t(s, a_{this}, a_{other})$ 为到时刻 t 为止,Agent 得到的经验的次数 (s, a_{this}, a_{other}) ;

⑤根据公式(1),Agent 更新信念,得到 P_s ;

⑥根据公式(2),(3),Agent 依次更新 $Q_t(s, a_{this}, a_{other})$,并存储值 Q ;

⑦ $s \leftarrow s'$,根据个体的情感状态,产生可能出现的各个情感元素 a' ,并以缺省值存储 $Q(s, a')$;

⑧若学习 Agent 认定该时刻个体的情感状态项是正确的估计(收敛于 s^* 状态),则记录下来,否则转至步骤(2)继续执行。

2 基于动态 Q 学习算法的情感自动机模型

2.1 情感的交互特性

之所以建立“情感自动机模型”,其主要原因有三:第一,情感来源于认知,对于来源于外部的刺激能够做出本能的、快速的智能反应,因此,通常的情感模型必须把可计算的数学模型与认知科学结合起来;第二,情感比知识更能控制人群的氛围,在某些特定的场合情感比理智更加有效、更有感染力;第三,机器人学研究领域,缺乏真正具有人类情感的机器人,充其量只能说其具有某种智能,这些智能不但使机器人不能表达自己,更不能通过它们的情感与人类进行真正的“沟通”和“交流”^[5]。

心理学家普遍认为,情感是一种综合性的行为反应,每种情感都包括生理、行为和认知三种成分。它们在每种特定的情感中各自起着不同的作用,而相互作用互为因果。情感是有“本能的”生物成分和适应价值的,但更重要的是情感受社会文化情境的制约,受个体经验和人格结构的影响,而这一切都体现在每时每刻的变化之中,因此无论从心理学角度还是信息学角度,把某类情感单一地割裂开来进行研究,为了把某时间段发生的情感联合起来进行研究,本部分内容引入了情

感元的概念。

2.2 情感元及情感自动机模型

情感元是由一个二维向量组成 $\langle E, f \rangle$,其中 E 为情感状态描述,所谓的情感状态是指混合情感中的主要情感,即支配个体心智状态和行为行动的情感。 E 包括下列的情感状态:高兴(Happiness,简称 Ha)、愤怒(Anger,简称 An)、悲伤(Sadness,简称 Sa)、害怕(Fear,简称 Fe)、厌恶(Disgust,简称 Di)和惊奇(Surprise,简称 Su),因此 E 可用一个有限元集合来表示为 $E = \{Ha, An, Sa, Fe, Di, Su\}$ 。函数 $f: E \times E \rightarrow E$, f 表示情感元的下一个状态是由当前状态与输入状态共同决定的。用 s' 表示 t 段时间内的情感状态,其中 t 表示一个时间段,而不是一个时间点, y' 表示输入的情感状态,则 $s^{t+1} = f(s', y')$,其中 $s', y', s^{t+1} \in E$ 。这里之所以引入了输入状态的概念,是因为情感具有“传染性”,一个特定的环境下,人的心情会无意识地随着周遭环境的变化而变化,确切地说,是受着周围的人的情绪的影响。

本文定义了两种不同情感实体之间的交互(也包括实体与环境的交互),假设承担某两种情感元的实体分别用两个智能体(Agent)来代替,即 $Agent_i$ (简称 A_i) 与 $Agent_j$ (简称 A_j),即不同的情感将在 A_i 与 A_j 之间互相“传染”而产生变化, A_j 所具有的情感状态将由上文所引入的动态 Q 学习算法进行估计,估价结果如图 1 所示。该变化可表示为: $A_i^a \leftrightarrow A_j^b$,其中 $a, b \in \{0, +, -\}$,“0”意味着“中立”,即不受影响;“+”意味着“正”,即受到对方正面的影响,“-”意味着“负”,即受到对方负面的影响。为了说明问题,将 E 按照影响效应进行分类,其中 $E_+ = \{Ha\}$; $E_0 = \{Fe, Su\}$; $E_- = \{An, Sa, Di\}$,设任意集合 A, B, C 表示 $A = \{E_+, E_-\}$, $B = \{E_+, E_0\}$, $C = \{E_0, E_-\}$,然后由集合 A, B, C 中的情感元素分别作为 y' 和 s' 的值,则 $\alpha = A \times A$; $\beta = B \times B$; $\chi = C \times C$; $\delta = A \times B$; $\epsilon = B \times C$; $\varphi = A \times C$ 。其中 $\alpha, \beta, \chi, \delta, \epsilon, \varphi \in E$ 。

s^{t+1}		y'			
		Ha	An	Sa	Di
s'	Ha	Ha	Su	Su	Su
	An	Su	An	An	An
	Sa	Su	An	Sa	An
	Di	Su	An	Di	Di

(a)

s^{t+1}		y'		
		Ha	Fe	Su
s'	Ha	Ha	Su	{Su, An}
	Fe	Ha	Fe	Su
	Su	Ha	Fe	Su

(β)

s^{t+1}		y'				
		An	Sa	Di	Fe	Su
s'	An	An	An	An	{An, Fe}	{An, Su}
	Sa	An	Sa	An	{Sa, Fe}	{Sa, Su}
	Di	An	Sa	Di	An	An
	Fe	Fe	Sa	Su	Fe	{An, Su}
	Su	An	Sa	Di	Fe	Su

{χ}

(下转第 176 页)

由表3可以看出,在处理不平衡数据分类时,Biased-SVM要明显优于STSVM,尤其是对于属性数和不平衡程度较大的Satimage和Vehicle数据集。Sos+BSVM比单纯使用Biased-SVM的结果高1个百分点,而本文提出的对支持向量采用SMOTE和Biased-SVM相结合的方法(Support Vector SMOTE and Biased-SVM,SVSmt)都比其他方法优越。

表3 各种方法的g值(%)

	STSVM	Biased-SVM	Sos+BSVM	SVSmt
Phoneme	82.698	84.022	84.429	85.193
Pima	71.002	73.276	74.343	75.92
Satimage	75.737	86.764	88.597	89.782
Vehicle	72.158	81.417	82.086	83.232
平均值	75.399	81.370	82.364	83.532

结束语 本文首先介绍了目前对不平衡数据提出的一些处理方法。在此基础上,本文提出了一种基于支持向量的SMOTE向上采样和Biased-SVM相结合的分类方法。通过对UCI四种公开数据集上的对比实验,验证了基于支持向量的SMOTE向上采样和Biased-SVM相结合的分类方法比标准支持向量机方法、Biased-SVM方法和对数据利用SMOTE向上采样和Biased-SVM相结合的方法优越。

(上接第173页)

s^{t+1}		y^t			
		Ha	An	Sa	Di
s^t	Ha	Su	Fe	Su	Su
	Fe	HA	Fe	Sa	Su
	Su	Ha	An	Sa	Di
{ δ }					
s^{t+1}		y^t			
		Ha	Fe	Su	
s^t	Fe	Su	Fe	{Su, An}	
	Su	Ha	Fe	Su	
	An	Su	{An, Fe}	{Su, An}	
	SA	Su	{Fe, Sa}	{Sa, Su}	
	Di	Su	An	An	
{ ϵ }					
s^{t+1}		y^t			
		Ha	An	Sa	Di
s^t	Fe	Su	Fe	Sa	Su
	Su	Ha	An	Sa	Di
	An	Su	An	An	An
	Sa	Su	An	Sa	An
	Di	Su	An	Sa	Di
{ φ }					

图1 情感自动机内部的状态转化

着重对(α)和(β)两图进行解释说明,其他各图依此类推:

(α)中,仅仅当 $f(Ha, Ha)=Ha$,其余的情感元中均属于 E_0 和 E_- ,而且可以看出 $E_- \times E_- \rightarrow E_-$,由此可知,当带有负

参考文献

- [1] Chawla N, Bowyer K, Hall L, et al. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 2002, 16: 321-357
- [2] Kubat M, Matwin S. Addressing the Curse of Imbalanced Training Sets: One-sided Selection. ICML, 1997: 179-186
- [3] Joshi M, Kumar V, Agarwal R. Evaluating Boosting Algorithms to Classify Rare Classes: Comparison and Improvements//First IEEE International Conference on Data Mining. 2001: 257-264
- [4] Wu G, Chang E. Class-boundary Alignment for Imbalanced Dataset Learning//The Twentieth International Conference on Machine Learning (ICML) Workshop on Learning from Imbalanced Datasets. Washington DC, 2003, 8: 49-56
- [5] Huang Kaizhu, Yang Haiqin, King I, et al. Learning Classifiers from Imbalanced Data Based on Biased Minimax Probability Machine//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2004: 558-563
- [6] Veropoulos K, Campbell C, Cristianini N. Controlling the sensitivity of support vector machines //Proceedings of the International Joint Conference on AI. 1999: 55-60
- [7] Blake C, Merz C. UCI Repository of Machine Learning Databases. <http://www.ics.uci.edu/~mllearn/~MLRepository.html>. Department of Information and Computer Sciences, University of California, Irvine, 1998
- [8] Webb A R. 统计模式识别. 王萍, 杨培龙, 罗颖新, 译. 电子工业出版社, 2004, 10: 33-37
- [9] Vapnik V. The Nature of Statistical Learning Theory. NY: Springer-Verlag, 1995
- [10] Akbani R, Kwek S, Japkowicz N. Applying Support Vector Machines to Imbalanced Datasets. ECML 2004, LNAI, 3204: 39-50

面情感的智能体遭遇到另一带有负面情感的智能体时,它们彼此的情绪都不会有所改观,甚至于交互后,双方情绪会变得更遭,例如, $f(Sa, An)=An$,当处于 Sa (忧伤)状态的 A_i 遇到 An (生气)状态的 A_j 时,就变得十分 An (生气)。

(β)中, $f(Fe, Ha)=f(Su, Ha)=Ha$,由此可见当智能体处于 E_0 状态时,该智能体最容易受周围环境的影响,或者说受另一智能体的影响,从 $f(Fe, Fe)=f(Su, Fe)=Fe$ 和 $f(Fe, Su)=f(Su, Su)=Su$ 上都可以看得出来。特别需要说明的是当 $s^{t+1} \in \{Su, An\}$ 时, $s^{t+1}=Su$ 和 $s^{t+1}=An$ 的概率均为 $\frac{1}{2}$ 。

结束语 事实证明,情感计算的概念虽诞生不久,却面临着巨大的挑战,原因在于情感的复杂性导致至今为止都没有统一的建模方法,已有的方法大多将情感元素作为一个单一的变量来进行分析和建模,这不符合情感状态复杂多变的特性,本文构建的情感自动机模型不是将某类情感割裂开来,而是把分析个体的主要情感作为研究依据,通过与环境交互,通过动态Q学习算法,实时地对其自身的情感进行协调和整合,为情感计算提供了方便的研究平台和表现形式,体现了情感计算的发展潜力和应用价值。

参考文献

- [1] Michael A. Arbib and Jean-Marc Fellous, Emotions: from brain to robot. Cognitive Sciences, USA, 2004, 8(12)
- [2] Picard R W. Affective Computing. Cambridge, MA: MIT Press, London, 1997
- [3] Hatfield E, Cacioppo J T, Rapson R L. Primitive emotion contagion, Rev. Personal Social Psychol, 1992, 14: 151-177
- [4] Hatfield E, Cacioppo J T, Rapson R L. Emotional Contagion. Cambridge University Press, 1994
- [5] Andrew, Adamatzky. Affections: automata models of emotional interactions. Applied Mathematics and Computation, 2003, 146: 579-594