

人工情感Q学习的互联电网自动发电控制算法

殷林飞, 郑宝敏, 余涛[†]

(华南理工大学 电力学院, 广东 广州 510640)

摘要: 对互联电网中自动发电控制AGC中控制策略进行改进, 设计了人工智能中的人工心理学和人工智能中的机器学习结合的控制策略. 分别对Q学习算法和Q(λ)学习算法进行改进, 设计了具有人工情感的智能体. 提出了人工情感Q学习算法和人工情感Q(λ)学习算法. 且将人工情感分别作用于Q学习算法和Q(λ)学习算法中的输出动作、学习率和奖励函数. 最后在IEEE标准两区域和南方电网四区域的互联电网Simulink模型中进行数值仿真. 绘制并统计了控制性能指标、区域控制误差和频率偏差的值. 从仿真结果看, 所提人工情感Q学习算法和人工情感Q(λ)学习算法控制效果优于原有Q学习算法、Q(λ)学习算法、R(λ)算法、Sarsa算法、Sarsa(λ)算法和PID控制算法, 该数值仿真结果验证了所提算法的可行性和有效性.

关键词: 人工情感; Q学习; Q(λ)学习; 自动发电控制

中图分类号: TM734 文献标识码: A

Artificial emotional Q-learning for automatic generation control of interconnected power grids

YIN Lin-fei, ZHENG Bao-min, YU Tao[†]

(School of Electric Power, South China University of Technology, Guangzhou Guangdong 510640, China)

Abstract: Artificial psychology and machine learning are combined in the automatic generation control strategy of interconnected power grids. An agent obtaining artificial emotion is designed, and the Q-learning and Q(λ)-learning algorithms are improved by artificial emotion. The novel artificial emotional Q-learning and artificial emotional Q(λ)-learning algorithms are proposed. The artificial emotion is respectively applied to the selection of output action, learning rate and reward function in Q-learning and Q(λ)-learning, and then simulated on the standard IEEE two-area model and the China Southern Power Grid four-area model. The control performance standard, area control error and frequency deviation are figured. Simulation results verify the feasibility and effectiveness of the proposed algorithms and their superiority to the Q-learning, Q(λ)-learning, R(λ), Sarsa, Sarsa(λ) and PID algorithms.

Key words: artificial emotion; Q-learning; Q(λ)-learning; automatic generation control

1 引言(Introduction)

合格的频率是电网运行所必须的, 其能确保电网的稳定与安全^[1]. 为评价电网频率的优劣, 1997年北美电力可靠性协会提出了控制性能指标(control performance standard, CPS). 针对互联电网中自动发电控制(automatic generation control, AGC)问题, 目前普遍使用的是PI控制器, 而当扰动发生较大变化时, PI控制器需要重新调整参数. 需要自动调整参数则需要采用智能算法进行控制. 智能算法能在解决此问题的同时, 还能达到其CPS性能考核指标, 如模糊控制^[2]等, 同时有将许多种方法进行结合. Q学习算法则具有自学习功能^[3], 而且是在线学习的能力,

能适应于未知环境, 也能解决此问题. 而这些智能算法都属于人工智能中的一个分支, 即机器学习算法. 人工智能中的另一个分支为人工心理学.

目前, 针对人工智能的两个分支(机器学习和人工心理学)都有较多的研究. 机器学习方面分类较多, 其中的综合分类可将机器学习分为经验性归纳学习、分析学习、类比学习、遗传学习、联接学习和强化学习六大类^[4]. 近日非常热门的深度学习就属于联接学习. 而人工心理学则主要包括人工情感、人工意识、人工认知和人工情绪. 其中的人工情感研究较热^[5], 较多研究情感识别的研究者在研究生理信号识别、面部表情识别、姿态表情识别和语音识别^[6-9]. 研究者发现

收稿日期: 2016-05-20; 录用日期: 2016-12-20.

[†]通讯作者. E-mail: taoyu1@scut.edu.cn; Tel.: +86 13002088518.

本文责任编辑: 赵冬斌.

国家自然科学基金项目(51177051, 51477055), 国家“973”计划项目(2013CB228205)资助.

Supported by National Natural Science Foundation of China (51177051, 51477055) and National “973” Program (2013CB228205).

情感还可以进行分级,如分为初级情感和二级情感^[10],也可以进行分层,如多层情感模型^[11].也有专门研究情感的模型,如Ortony、Clore和Collins三人构建的OCC情绪模型、卡内基-梅隆Bates和Reilly构建的Oz-Tok结构、J.Velasquez提出的Cathexis模型、Breazeal设计的Kismet模型、隐马尔可夫模型(hidden Markov model, HMM)模型和Picard模型等^[12].人工情感与机器学习和非机器学习算法的结合也有一些,如和模糊控制进行结合^[13]、和进化算法结合^[14]和小脑模型控制器(cerebellar model articulation controller, CMAC)网络结合网络结合^[15].人工情感的Q-学习算法已有在机器人避障问题中进行应用^[16-17],应用于机器人的行为决策中,将人工情感作用于Q矩阵的更新,以经验Q矩阵或者不同子目标的Q矩阵乘以人工情感输出因子.但人工情感与Q学习的结合还应有许多种,如作用于Q学习算法中的动作集等.

原始Q学习算法中,也存在一定的问题,如动作的不连续、学习率的一定和奖励函数的不变化.引入人工情感后,将离散的动作值变为可取连续值,学习率和奖励函数随学习效果可变.目前尚未有直接利用人工情感Q学习算法来解决互联网的自动发电控制问题,但有将神经网络和模糊逻辑结合起来处理自动发电控制问题的文章^[18],且被称为“情感学习”智能控制.利用具有人工情感的智能体参与互联网的自动发电控制问题,该智能体控制器不仅具有逻辑控制能力,还具有情感控制能力.因此可从Q学习算法中的动作、学习率和奖励函数3个方面考虑人工情感因素,提出新的人工情感Q学习算法.

2 自动发电控制 (Automatic generation control)

对于互联网的自动发电控制 (automatic generation control, AGC)而言,其目的是使控制区域的功率偏差为0,频率偏差也为0.

2.1 自动发电控制(AGC)

发电机组参与的一次调频具有衰减性和有差性,因此必须进行系统频率的二次调节.而现代大型的互联网普遍采用联络线功率频率偏差控制(tie-line load frequency bias control, TBC).以两个区域互联电力系统(如图1所示)为例进行描述TBC的控制方式.

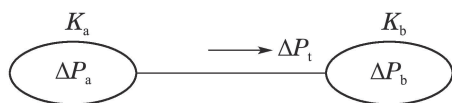


图 1 互联系统功率交换特性

Fig. 1 The power exchange feature of interconnected systems

系统A和系统B的调差系数分别为 K_a 和 K_b , 负荷

变化分别为 ΔP_a 和 ΔP_b , 两系统均设有二次调节的电厂,有功功率变化分别为 ΔG_a 和 ΔG_b , 负荷变化量分别为 ΔL_a 和 ΔL_b . 联络线功率变化为 ΔP_t . 当系统A和系统B发生功率变化而引起频率变化 Δf 时,

$$\begin{aligned} \Delta P_a &= \Delta G_a - \Delta L_a = K_a \Delta f + \Delta P_t, \\ \Delta P_b &= \Delta G_b - \Delta L_b = K_b \Delta f - \Delta P_t. \end{aligned} \quad (1)$$

由上式可得

$$\begin{cases} \Delta f = (\Delta G_a - \Delta L_a + \Delta G_b - \Delta L_b) / (K_a + K_b), \\ \Delta P_t = \frac{\{(\Delta G_a - \Delta L_a)K_b - (\Delta G_b - \Delta L_b)K_a\}}{(K_a + K_b)}. \end{cases} \quad (2)$$

区域控制误差(area control error, ACE) 则定义为

$$ACE = \Delta P_t - 10B\Delta f, \quad (3)$$

式中 B 为控制区的频率响应系数,为负值(MW/0.1 Hz).

2.2 控制性能指标(Control performance standard)

自动发电控制的评价标准有A1/A2和CPS两种.1997年由北美可靠性协会(North American Electric Reliability Corporation, NERC)提出的CPS评价标准中CPS1指标定义为一段时间内(典型值为10 min)某区域电网必须满足

$$\frac{\sum (ACE_{AVE-\min} * \Delta F_{AVE-\min})}{-10nB_i} \leq \varepsilon_1^2, \quad (4)$$

$ACE_{AVE-\min}$ 是1分钟ACE的平均值; $\Delta F_{AVE-\min}$ 是1分钟频率偏差的平均值; B_i 是控制区域的频率偏差系数,单位是10 MW/Hz; n 为该时段内的分钟数; ε_1 是互联网对全年1分钟频率平均偏差的均方根的控制目标值.CPS1指标的计算表达式为

$$CPS1 = (2 - CF1) \times 100\%, \quad (5)$$

$$CF1 = \frac{\sum (ACE_{AVE-\min} * \Delta F_{AVE-\min})}{-10B_i n \varepsilon_1^2}.$$

CPS2指标则要求区域控制误差ACE在考核时段(典型值为10分钟)内的均值小于 L_{10}

$$|\sum ACE_{AVE-\min}| / 10 \leq L_{10}, \quad (6)$$

式中: $L_{10} = 1.65\varepsilon_{10} * \sqrt{(-10B_i) * (-10B_s)}$ 为其范围, ε_{10} 为互联网对全年10分钟频率平均偏差的均方根的控制目标值; B_i 和 B_s 分别表示该区域电网和整个互联网的频率偏差系数;系数1.65是根据NERC的控制标准要求各控制区域的CPS2指标的合格率达到90%以上,且控制区域ACE的10分钟平均值 $\sigma = \varepsilon_{10} * \sqrt{(-10B_i) * (-10B_s)}$ 是符合的正态分布求得.依据统计学的CPS2考核指标为

$$CPS2 = \left[1 - \frac{T_u}{T_s - T_n} \right] \times 100\%, \quad (7)$$

式中 T_u , T_s 和 T_n 分别为考核不合格时段、总时段和非

考核时段. 考核不合格时段为ACE每10分钟的平均值大于 T_{10} 的考核时段数.

CPS指标的是否合格的判据:

1) $CPS1 \geq 200\%$, 则不论CPS2是否合格, 均判定CPS合格;

2) $CPS1 < 100\%$, 则不论CPS2是否合格, 均判定CPS不合格;

3) $100\% \leq CPS1 < 200\%$, 若CPS2合格, 则判定CPS合格, 若CPS2不合格, 则判定CPS不合格.

3 人工情感 Q 学习算法 (Artificial emotional Q-learning)

3.1 Q学习算法(Q-learning)

人工智能的一个重要分支为机器学习, 而机器学习又包含强化学习. 作为强化学习中最重要且最早的在线学习算法: Q学习算法^[19], 以智能主体通过环境给其的状态和奖励值来修正Q矩阵, 依据某种概率的方式从动作集中获取动作, 作为智能主体的输出. Q学习离散时间马尔可夫决策过程模型, Q学习算法是不基于对象模型的算法. 而Q学习中的Q矩阵则是以状态和动作组成的键值对, 即(状态-动作, Q值). 而Q矩阵的更新策略为:

$$Q(s, a) \leftarrow Q(s, a) + \alpha(R(s, s', a) + \gamma \max_{a \in A} Q(s', a) - Q(s, a)), \quad (8)$$

式中: s 和 s' 分别代表当前状态和下一时刻的状态, γ 为折扣因子, $R(s, s', a)$ 为从环境中获得的奖励值, 其计算公式为:

$$R(s, s', a) \leftarrow \frac{1}{\delta_s \sqrt{2\pi}} e^{-s^2/2\delta_s^2}, \quad (9)$$

式中 δ_s^2 为状态 s 的方差.

而概率矩阵 P 的更新策略为

$$P(s, a) \leftarrow \begin{cases} P(s, a) - \beta(1 - P(s, a)), & a' = a, \\ P(s, a)(1 - \beta), & a' \neq a, \end{cases} \quad (10)$$

式中 β 为概率分布因子. $P(s, a)$ 的初始值为 $P(s, a) = 1/|A|$, $|A|$ 为动作集中动作的数量, 且其范围为 $P(s, a) \in [0, 1]$.

在互联网的AGC控制中, Q学习算法中的智能体通过对输入(状态 s : CPS1和ACE与奖励值 R)和输出(动作 a : 机组功率指令)的学习, 不断强化对模型的认知, 在线地更新 P, Q 矩阵, 在随后的时刻里, 选择奖励值较大的动作进行执行. 强化学习中的Q学习的特点便是该算法有一个不断被强化的Q矩阵.

3.2 人工情感(Artificial emotion)

人工智能中的另一个较大分支为人工心理学, 而人工心理学则包括人工情感、人工意识、人工认知和人工情绪. 当智能体Agent具有某种特定事物的逻辑

处理能力时, 称为“弱人工智能”. 当智能体具有人工情感、人工意识和人工认知能力时, 称为“强人工智能”. 当智能体具有人工情绪时, 称为“超人工智能”.

人工情感是人工心理学目前较热的分支, 此时的智能体不仅仅包含以往的智能逻辑体, 还包含智能情感体. 而智能体的输出, 同时受到逻辑处理能力(类似人类的理性思维)和情感处理能力(类似人类的感性思维)的影响. 具有人工情感的智能体除了具有自治性、反应性、社会性和进化性外, 还具有情感性.

智能体具有情感后, 可设计出情感机器, 可用来模拟和分析人类的情感, 具有情感的机器成为情感机器. 智能体则通过环境和记忆中的情感因素共同决定当前输出的人工情感, 而该人工情感转换为工程问题时, 一般需将人工情感进行量化, 进行情感量化过程的量化器称为情感输出器, 示意图如图2所示. 简化处理的情感输出器输出的是情感输出值 η , 作为其他函数的比例因子, 该比例因子随智能体的情感更新而改变, 作为工程问题, 一般需将该比例因子限定在某个范围内.

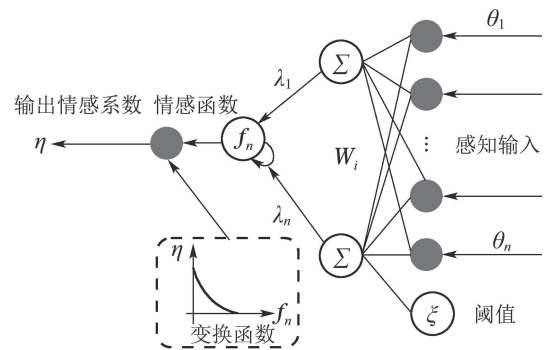


图2 人工情感输出示意图

Fig. 2 The output of artificial emotion

而人工情感 f_n 的计算可设计为

$$f_n = \sum_{i=1}^n \lambda_i = \sum_{i=1}^n \theta_i \omega_i, \quad (11)$$

式中 θ_i 和 ω_i 分别为人工情感感知到的信息和信息的权重, 而 λ_i 则为 θ_i 和 ω_i 的乘积, 人工情感 f_n 的总输出为信息素的积和.

人工情感的输出转换有多种方式, 较为简单的一种处理方式

$$\eta \leftarrow \begin{cases} k_\eta, & \frac{1}{f_n} \geq \eta_{\max}, \\ \frac{k_\eta}{f_n}, & \frac{1}{f_n} < \eta_{\max}, \end{cases} \quad (12)$$

式中 k_η 为人工情感最大化的范围系数, 即 $\eta \in [0, k_\eta]$.

3.3 人工情感 Q 学习算法 (Artificial emotional Q-learning, AEQL)

人工情感Q学习算法是利用人工情感更新Q学习算法中特定值的一种算法, 该算法示意如图3所示.

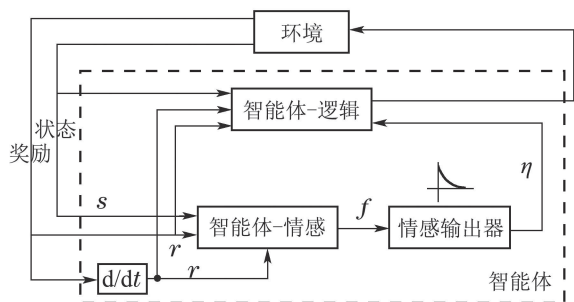


图 3 基于人工情感的Q学习算法示意图

Fig. 3 Artificial emotional Q-learning schematic diagram

智能体包含两个部分, 一是智能体逻辑思维部分 (agent logical part, ALP), 一是智能体情感思维部分 (agent emotional part, AEP). 而采用人工情感更新Q学习算法中特定值的策略可大致分为3种.

1) 人工情感直接作用于输出动作.

将得到的人工情感输出值直接对原有的动作值进行更改, 有利于形成连续的动作, 此时该智能体输出的动作为连续Q动作, 即属于连续Q学习算法中的一种策略. 此时的情感输出值 η 的范围一般是 $0 \leq \eta \leq 1$. 此时的智能体因为具有人工情感的原因, 能输出连续的动作, 弥补了Q学习算法中不能输出连续动作的不足.

设计一种人工情感直接作用于输出动作的策略为

$$a \leftarrow \eta a. \tag{13}$$

2) 人工情感作用于学习率.

将人工情感作用于学习率, 其作用在于加速其收敛速度, 在不断地迭代过程中, 使用人工情感的输出值更新Q学习的学习率. 也称为变学习率的Q学习算法. 此时的人工情感作用于Q学习算法起到了类似爬山算子的作用. 此时的情感输出值 η 的范围一般也是 $0 \leq \eta \leq 1$.

设计一种人工情感作用于学习率的策略为

$$\alpha \leftarrow \eta \alpha. \tag{14}$$

3) 人工情感作用于奖励函数.

将人工情感作用于奖励函数, 在不断迭代过程中修正奖励函数, 该算法优化难度大于前两者. 而奖励函数在Q学习算法中又非常重要, 奖励函数的设计将直接影响Q学习算法的稳定性. 也称为变奖励函数的Q学习算法. 此时人工情感不仅限于一个情感输出值 η , 而是对奖励函数的更新, 即更新集合键值对(状态变量集合, 奖励值).

人工情感Q学习算法的流程图如图4所示.

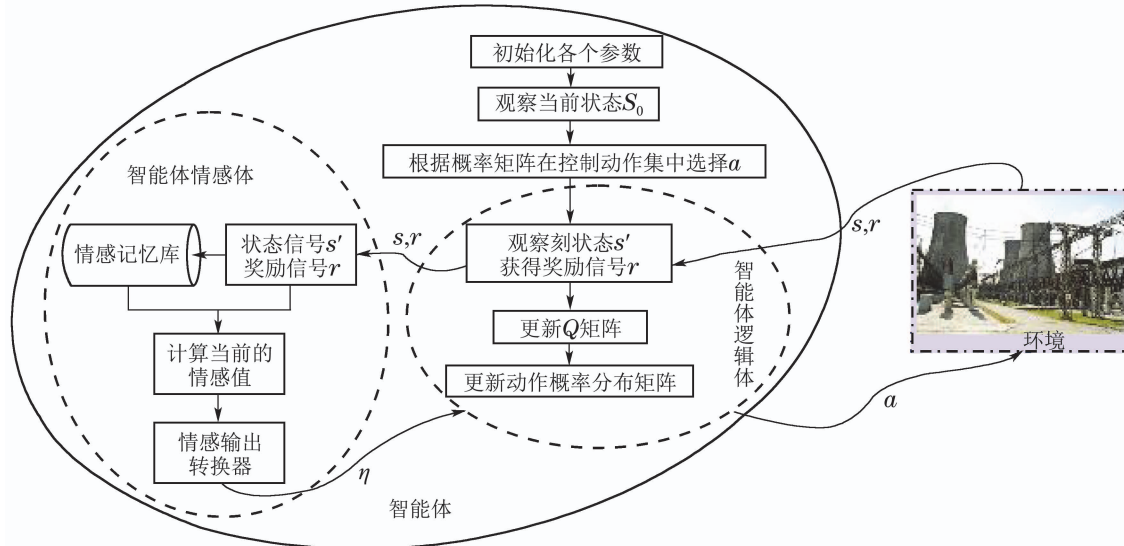


图 4 人工情感Q学习算法流程图

Fig. 4 The flow chart of artificial emotional Q-learning algorithm

设计一种人工情感作用于奖励函数的策略, 在状态 s 下智能体的奖励值为

$$R(s) \leftarrow \frac{1}{\delta_s \eta_\delta \sqrt{2\pi}} e^{-\frac{(s-\eta_s s)^2}{2(\delta_s \eta_\delta)^2}}, \tag{15}$$

式中: η_δ 为方差因子情感输出值, η_s 为状态因子情感输出值.

3.4 人工情感带资格迹的 $Q(\lambda)$ 算法 (Artificial emotional $Q(\lambda)$ with eligibility trace, EQ(λ)L)

$Q(\lambda)$ 算法是在Q学习的基础上改进的算法, 在原有Q学习算法的基础上引入 λ 参数, 再引入资格迹. 带资格迹的 $Q(\lambda)$ 算法更新Q值的策略为

$$Q(s, a) \leftarrow Q(s, a) + \alpha e_t (R(s, s', a) +$$

$$\gamma \max_{a \in A} Q(s', a) - Q(s, a), \quad (16)$$

式中 e_t 为资格迹, 其更新方式为

$$e_t(s, a) \leftarrow \begin{cases} \gamma \lambda e_t(s, a) + 1, & a = a', \\ \frac{1}{|A|}, & a \neq a'. \end{cases} \quad (17)$$

引入资格迹 e_t , 对智能体连续选择同一动作进行记录, 加快Q学习的收敛速度. 考虑到智能体中情感部分的存在, 人工情感Q(λ)学习的3种策略和人工情感Q学习中的3种策略一致.

结合互联电网的自动发电控制问题, 综合人工情感Q学习和人工情感Q(λ)学习算法, 形成算法步骤为:

Step 1 初始化矩阵(A, Q, P)和变量($\omega_i, \delta, \alpha, \beta,$

$\gamma, \lambda, e_t, \eta, \eta_{CPS1}, \eta_\delta, \eta_{ACE}$), 给出这些参数在本文的取值与取值范围(通过大量仿真获得), 如表1所示;

Step 2 获取当前的状态 s : CPS1, ACE, 作为控制器的输入变量;

Step 3 按式(11)(12)计算人工情感的输出值 η ;

Step 4 按式(9)或式(15)计算奖励值;

Step 5 按式(13)(8)或式(14)(8)或式(15)(8)或式(17)(13)(16)或式(17)(14)(16)或式(17)(15)(16)更新Q矩阵;

Step 6 按式(10)更新P矩阵;

Step 7 根据P矩阵从动作集A中选择动作 a 作为功率指令 ΔP_{Ci} , 也作为控制器的输出变量;

Step 8 循环执行Step2.

表 1 算法参数取值表

Table 1 Parameters of the proposed algorithm

参数	本文取值	取值范围
ω_i	0.8, 0.2	[0.1, 0.9], [0.1, 0.9]
δ	0.4	(0.4-1]
α	0.1	[0.05, 0.9]
β	0.5	(0, 1)
γ	0.3	(0, 1)
λ	0.9	(0.85, 0.95)
$\eta_{CPS1}, \eta_\delta, \eta_{ACE}$	0.1	[0.1, 1)

4 数值仿真(Numerical simulation)

所有算例均在CPU为i7-2760QM @ 2.40GHz内存为8GB系统为Windows 10 Enterprise 64位的笔记本上运行, 所用软件版本为MATLAB R2016b.

4.1 IEEE标准两区域模型(The two area model of IEEE)

以IEEE标准两区域的负荷频率控制模型为例, 系统基准容量取5000 MW. 采用MATLAB软件搭建Simulink仿真模型. 仿真中的各学习算法控制器均采用CPS1和ACE实时值作为状态, 即控制器的输入(输

入变量). 控制器的输出为发电机功率指令, 控制算法中的动作离散集均为功率调节集: $A = \{-500, -40, -30, -20, -10, 0, 10, 20, 30, 40, 500\}$, 单位是 MW. 学习步长一般取AGC系统控制周期, 该标准算例中取4 s(所有算法均在MATLAB/Simulink下的S-Function实现, 所有算法的步长均设定为4 s, 特别是PID控制器设计为离散型PID, 每4 s控制一次). 仿真时长为12000 s. 对每个算法进行仿真10次, 统计其仿真时间, 每个算法的仿真时间在 0.00006 s 至 0.01957 s 之间, 完全符合实际情况下的4 s控制指令时间. 仿真模型框图如图5所示, 图中也展示了系统的状态变量, 如 $\Delta f_A, \Delta f_B, \Delta P_{GA}, \Delta P_{GB}, \Delta P_{CA}, \Delta P_{CB}, \Delta P_{tie}$ 等, 系统模型相关参数为

$$\begin{cases} T_g = 0.08 \text{ s}, T_t = 0.3 \text{ s}, T_p = 20 \text{ s}, \\ R = 2.4 \text{ Hz/pu}, K_p = 120 \text{ Hz/pu}, \\ T_{12} = 0.545 \text{ s}. \end{cases} \quad (18)$$

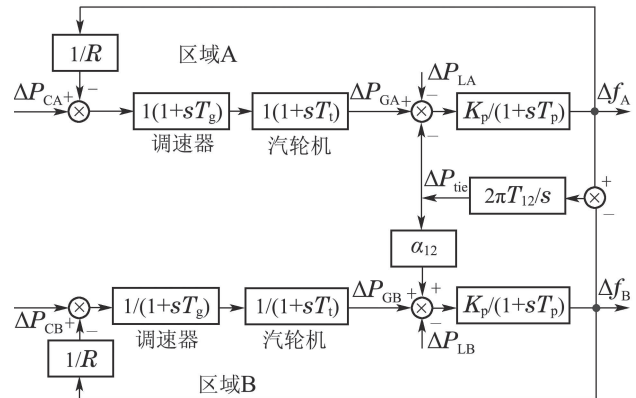


图 5 两区域互联系统负荷频率控制模型

Fig. 5 The two-area power system LFC model

在区域A输入周期1200 s幅值为1000 MW的正弦扰动, 扰动波形如图6所示.

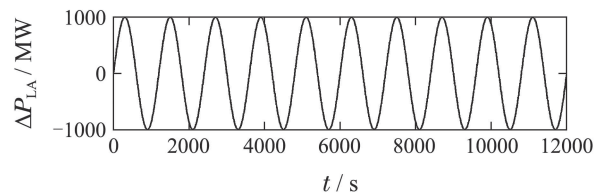
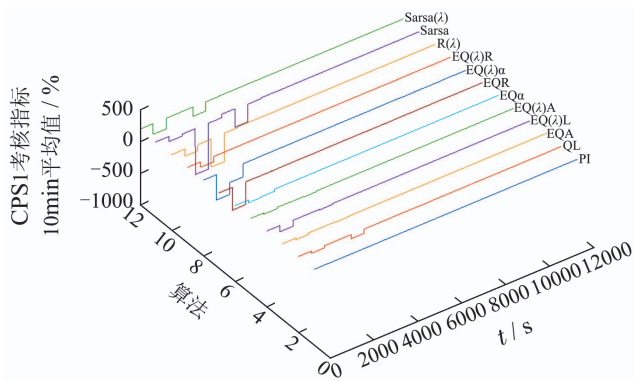


图 6 负荷扰动波形

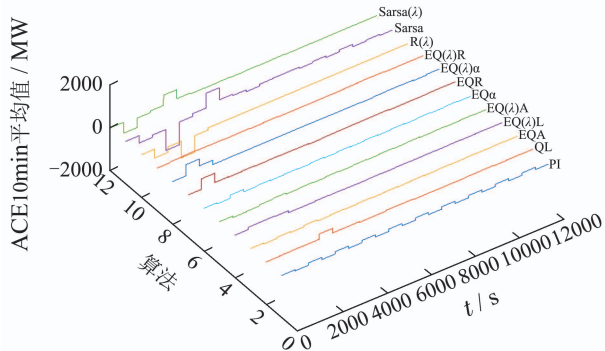
Fig. 6 Load disturbance waveform

在AB两区域都采用PID算法, Q学习、人工情感Q学习(作用于输出动作)、人工情感Q学习(作用于学习率)、人工情感Q学习(作用于奖励函数)、Q(λ)学习、人工情感Q(λ)学习(作用于输出动作)、人工情感Q(λ)学习(作用于学习率)、人工情感Q(λ)学习(作用于奖励函数)、R(λ)、Sarsa和Sarsa(λ)12种算法进行数值仿真, 仿真后得到的CPS1考核指标10分钟平均值、ACE10分钟平均值、CPS2考核指标、频率偏差和功率输出曲线如图7所示.

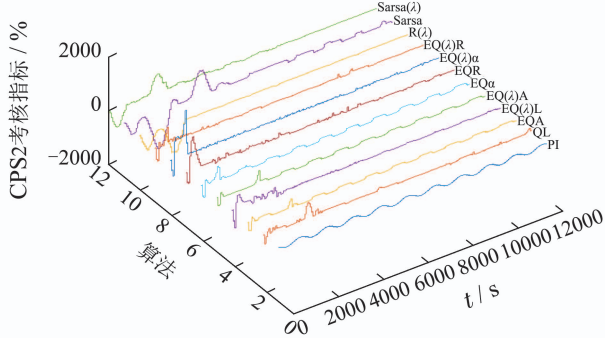
图中, QL为Q学习, EQA为人工情感Q学习(作用于输出动作), EQ α 为人工情感Q学习(作用于学习率), EQR为人工情感Q学习(作用于奖励函数), Q(λ)L为Q(λ)学习, EQ(λ)A为人工情感Q(λ)学习(作用于输出动作), EQ(λ) α 为人工情感Q(λ)学习(作用于学习率), EQ(λ)R为人工情感Q(λ)学习(作用于奖励函数)算法。



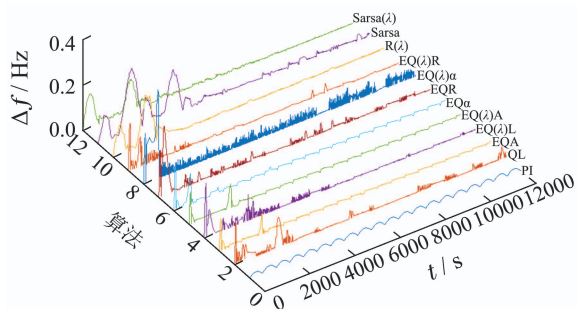
(a) CPS1考核指标10分钟平均值



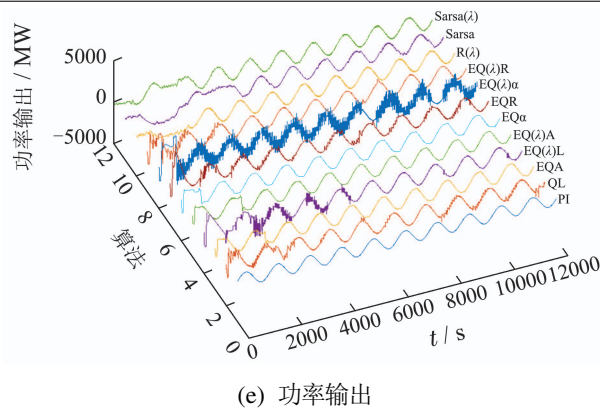
(b) ACE10分钟平均值



(c) CPS2考核指标



(d) 频率偏差



(e) 功率输出

图 7 两区域仿真结果

Fig. 7 The result of the IEEE standard two-area model

从图7可以看出,人工情感Q学习和人工情感Q(λ)学习比Q学习和Q(λ)学习的控制好.作用于输出动作的人工情感学习算法比原算法输出更加光滑,作用于学习率人工情感学习算法比原算法学习速率更快,作用于奖励函数的人工情感学习算法比原有算法更需要经验控制.对于作用于学习率的学习算法而言,人工情感Q学习比人工情感Q(λ)学习算法的效果更加, Q(λ)学习算法中学习率的控制效果弱于Q学习算法.其他学习算中, Sarsa(λ)和R(λ)算法效果与Q(λ)算法效果接近,都强于Sarsa算法的效果.这些算法的效果都优于PID控制算法。

4.2 南方电网四区域模型 (The four-area of the China Southern Power Grid)

为结合实际并验证所提人工情感 Q 学习和人工情感Q(λ)学习算法的可行性,以南方电网广东省电力调度中心实际工程项目搭建的详细全过程动态仿真模型四区域为例.在该算例中,分别采用两区域控制中的12种算法构成控制器,且均以AGC控制周期4 s为仿真步长.对每个算法进行仿真10次,统计其仿真时间,每个算法的仿真时间在0.000003 s至0.017403 s之间,完全符合实际情况下的4 s控制指令时间.控制算法中的动作集为 $A = \{-1000, -600, -300, -100, -50, -20, -10, -5, 0, 5, 10, 20, 50, 100, 300, 600, 1000\}$.

在幅值1000 MW周期为20 min的正弦扰动中加入20%有限带宽白噪声负荷扰动,并进行数值仿真,统计仿真24小时的数据如表2所示. Δf , ACE, CPS1为考核值的24小时内的均值, CPS2, CPS为24小时(仿真时长为86400 s)内考核合格率百分数,取CPS2考核标准阀限值 L_{10} 为288 MW.

从表2中的CPS指标、区域控制误差ACE和频率偏差 Δf 可以看出,人工情感Q学习算法和人工情感Q(λ)学习算法的仿真结果分别优于Q学习算法和Q(λ)学习算法。

表2 具有噪声扰动的仿真结果统计表

Table 2 The statistics of simulation results with noise disturbance

算法	CPS1/%	CPS2/%	ACE/MW	$\Delta f/\text{Hz}$	CPS/%
PID	179.7825	24.3076	413.3252	0.042109	51.5191
QL	227.1068	44.6175	790.0838	0.038711	68.7399
EQA	193.8088	95.3727	94.901	0.020851	95.7605
$\text{EQ}\alpha$	210.1501	85.3358	288.425	0.021626	92.235
EQR	208.3979	81.1379	319.0251	0.02317	89.455
$\text{Q}(\lambda)\text{L}$	195.2192	98.7801	75.0707	0.01838	98.7801
$\text{EQ}(\lambda)\text{A}$	196.2952	99.7107	58.8304	0.015517	99.7107
$\text{EQ}(\lambda)\alpha$	202.9381	85.7328	240.6214	0.022033	92.0394
$\text{EQ}(\lambda)\text{R}$	196.2976	99.5289	58.7561	0.015454	99.5289
$\text{R}(\lambda)$	197.0836	92.0128	175.0015	0.022677	94.6644
Sarsa	199.9805	92.3365	206.268	0.032373	92.3365
$\text{Sarsa}(\lambda)$	193.5893	97.3681	84.4853	0.020744	97.8716

对Q学习和 $\text{Q}(\lambda)$ 学习算法而言,其动作集中的动作作为离散值,而非连续值,这自然会给控制的波形带来一些“毛刺”,可以采集增加动作个数的方式进行改进,但是随着动作数量的增多,会带来“维数灾难”.针对该问题EQA和 $\text{EQ}(\lambda)\text{A}$ 算法则进行了改进,可将原本在Q学习和 $\text{Q}(\lambda)$ 学习算法中动作的非连续值改为连续值,在Q学习和 $\text{Q}(\lambda)$ 学习算法基础上,加入人工情感的控制,在原有动作上考虑人工情感因子,原有离散动作值则变为连续动作值,不仅解决了维数灾难的问题,而且能输出连续动作. $\text{EQ}(\lambda)\text{A}$ 算法的控制效果最优,优于仿真中的其他11种算法.

而 $\text{EQ}\alpha$ 和 $\text{EQ}(\lambda)\alpha$ 则从变学习率的角度来改进算法,EQR和 $\text{EQ}(\lambda)\text{R}$ 则从变奖励函数的角度改进算法.然而变学习率算法的设计中,并非每次都由于原本的Q学习算法.

5 结论(Conclusion)

采用所提算法人工情感的Q学习和人工情感 $\text{Q}(\lambda)$ 学习算法与Q学习, $\text{Q}(\lambda)$,PID, $\text{R}(\lambda)$ 和Sarsa等学习算法在互联网自动发电控制AGC的控制器中进行对比仿真,仿真结果表明所提算法优于原有算法,验证了具有人工情感的智能体比原有仅具有逻辑处理能力的智能体更加智能.

但是具有人工情感的智能体并非绝对智能,该智能体也未能达到人类的智慧,当其具有人工意识、人工认知和人工情绪时,才能够达到真正意义上的“智能”.当智能体达到真正意义上的“智能”时,智能体才能参与人类的带有“情绪”的博弈.

参考文献(References):

[1] ZHENG Wencheng. *Research on multi-agent simulation platform for AGC based on JADE* [D]. Guangzhou: South China University of Technology, 2014.
(郑闻成. 基于JADE的多智能体动态博弈自动发电控制仿真平台研

究 [D]. 广州: 华南理工大学, 2014.)

- [2] PRATAP C P, RABINDRA K S, SIDHARTHA P. Firefly algorithm optimized fuzzy PID controller for AGC of multi-area multi-source power systems with UPFC and SMES [J]. *Journal of Engineering Science and Technology*, 2016, 1(19): 338 – 354.
- [3] YU T, ZHOU B, CHAN K W, et al. $\text{R}(\lambda)$ imitation learning for automatic generation control of interconnected power grids [J]. *Automatica*, 2012, 48(9): 2130 – 2136.
- [4] MITCHELL T M, CARBONELL J G, MICHALSKI R S. *Machine Learning* [M]. New York: Springer, 1996.
- [5] ZHU Yuhong, WEI Jinhai, MAO Junxin. Summary of artificial emotion [J]. *Journal of Jiangnan University(Natural Science Edition)*, 2012, 11(4): 497 – 504.
(祝宇虹, 魏金海, 毛俊鑫. 人工情感研究综述 [J]. 江南大学学报(自然科学版), 2012, 11(4): 497 – 504.)
- [6] WANG W, HUANG X, ZHAO J, et al. Physiological signals based day-dependence analysis with metric multidimensional scaling for sentiment classification in wearable sensors [J]. *Journal of Engineering and Technological Sciences*, 2015, 47(1): 104 – 116.
- [7] CHEN Xin. *Research on facial expression recognition* [D]. Tianjin: Tianjin University, 2014.
(陈鑫. 人脸面部表情的情感识别研究 [D]. 天津: 天津大学, 2014.)
- [8] YAN Jingjie, ZHENG Wenming, XIN Minghai, et al. Bimodal emotion recognition based on body gesture and facial expression [J]. *Journal of image and graphics*, 2013, 18(9): 1101–1106.
(闫静杰, 郑文明, 辛明海, 等. 表情和姿态的双模态情感识别 [J]. 中国图象图形学报, 2013, 18(9): 1101 – 1106.)
- [9] PALO H K, MOHANTY M N, CHANDRA M. Design of neural network model for emotional speech recognition [M] // *Artificial intelligence and evolutionary algorithms in engineering systems*. New Delhi: Springer, 2015: 291 – 300.
- [10] SONG Yixu, JIA Peifa. Artificial emotion and its applications [J]. *Control Theory & Applications*, 2004, 21(2): 315 – 320.
(宋亦旭, 贾培发. 人工情感及其应用 [J]. 控制理论与应用, 2004, 21(2): 315 – 320.)
- [11] HE Haipeng. *Research on a layered model of affect* [D]. Taiyuan: Taiyuan University of Technology, 2010.
(何海鹏. 多层情感模型的研究 [D]. 太原: 太原理工大学, 2010.)
- [12] LIANG Z. Review on virtual human's affective model and related applied algorithm [J]. *Journal of Digital Information Management*, 2012, 10(2): 71 – 78.
- [13] VINCENZO L, SABRINA S. A fuzzy-oriented sentic analysis to capture the human emotion in Web-based content [J]. *Knowledge-Based Systems*, 2014, 58(1): 75 – 85.

- [14] SALICHS M A, MALFAZ M. A new approach to modeling emotions and their use on a decision-making system for artificial agents [J]. *IEEE Transactions on Affective Computing*, 2012, 3(1): 56 – 68.
- [15] ZHU Yuhong, WEI Jinhai. Robot's behavioral learning based on artificial emotion and CMAC network [J]. *Machinery & Electronics*, 2012, 9: 72 – 76.
(祝宇虹, 魏金海. 基于人工情感与CMAC网络的机器人行为学习 [J]. 机械与电子, 2012, 9: 72 – 76.)
- [16] GU Xuejing, GAO Beibei, ZHU Zhaoyue. Application of intelligent decision-making research based on artificial emotion and Q-Learning [J]. *Automation & Instrumentation*, 2015, 30(7): 13 – 16, 48.
(谷学静, 高贝贝, 朱朝月. 基于人工情感的Q-学习算法在机器人行为决策中的应用 [J]. 自动化与仪表, 2015, 30(7): 13 – 16, 48.)
- [17] SHI X, WANG Z, PING A, et al. Artificial emotion model based on reinforcement learning mechanism of neural network [J]. *Journal of China Universities of Posts and Telecommunications*, 2011, 3(18): 105 – 109.
- [18] REZA F, MEHRDAD B, SEYED H H. Load – frequency control of interconnected power system using emotional learning-based intelligent controller [J]. *Journal of Electrical Power and Energy Systems*, 2012, 36(1): 76 – 83.
- [19] ZHAO Dongbing, SHAO Kun, ZHU Yuanheng, et al. Review of deep reinforcement learning and discussions on the development of computer Go [J]. *Control Theory & Applications*, 2016, 33(6): 701 – 717.
(赵冬斌, 邵坤, 朱圆恒, 等. 深度强化学习综述: 兼论计算机围棋的发展 [J]. 控制理论与应用, 2016, 33(6): 701 – 717.)

作者简介:

殷林飞 (1990–), 男, 博士研究生, 目前研究方向为智能电网技术,

E-mail: epyinlinfei@mail.scut.edu.cn;

郑宝敏 (1994–), 男, 硕士研究生, 目前研究方向为智能电网技术,

E-mail: 2420535338@qq.com;

余涛 (1974–), 男, 教授, 博士生导师, 目前研究方向为复杂电力系统的非线性控制理论和仿真, E-mail: taoyu1@scut.edu.cn.