

学习自动机在混杂系统切换控制上的初步应用

董苗波, 吴铁军

(浙江大学 智能系统与决策研究所 工业控制技术国家重点实验室, 浙江 杭州 310027)

摘要: 提出了学习自动机应用于混杂系统切换控制的方法, 其主要思路是首先确定一个作为控制信号的行为集, 利用学习自动机的学习能力, 学习优化行为的发生概率, 最终实现满意的控制. 对一个实际应用中的混杂对象的切换控制的例子作了仿真, 实际结果体现了学习自动机在混杂系统切换控制上的优化能力, 证明了它的可实现性.

关键词: 学习自动机; 混杂系统; 切换控制

中图分类号: TP273 + .22; TP271 + .8 **文献标识码:** A

Applying learning automation on hybrid system switching control

DONG Miao-bo, WU Tie-jun

(National Key Laboratory of Industrial Control Technology, Institute of Intelligent Systems and Decision Making,
Zhejiang University, Zhenjiang Hangzhou 310027, China)

Abstract: An approach of applying learning automation on hybrid system switching control is proposed. A learning automation controller may learn from environment feedback and adjust its actions-control signals, i. e., switching signals leading the hybrid system to satisfactory states. A simulation example is presented to show the good learning result and the feasibility of the approach.

Key words: learning automation; hybrid system; switching control

1 引言(Introduction)

学习自动机以一种自动学习的方式工作, 它的输出由一个行为集以及每一个行为的发生概率决定, 在运行过程中, 系统根据反馈所得的信息进行学习, 调整行为的发生概率, 自动适应周围的环境(如图 1 所示). 在机器人控制、图像识别等智能领域, 学习自动机有很多成功的应用^[1~3].

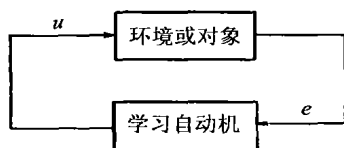


图 1 学习自动机
Fig. 1 Learning automation

很多对象涉及混杂系统的切换控制, 譬如电力系统、交通系统等等^[4,5]. 这些系统都有不止一个的运行模式(子系统), 它们的运行模式可以根据环境输入和系统状态发生切换. 由于它们本身的复杂性, 对它们进行建模, 求解最优切换控制的方法是一个难解的问题, 到目前为止只有一些很简单的例子的控制问题可以通过这样的方法解决.

本文作了将学习自动机方法应用于混杂系统切

换控制的尝试, 利用学习自动机控制系统的学习能力, 自动寻找混杂对象切换的优化控制信号. 比较通过建模求解控制算法的方法, 采用自动机学习系统的方法可以绕过混杂对象的复杂性, 由它的学习机制来保证控制的优化.

2 学习自动机(Learning automation)

一个学习自动机系统具有一个行为集 $\{u_1, u_2, \dots, u_N\}$, 每个行为对应一个发生概率 $p_i(t)$ (对文 [1] 中所述的学习自动机(learning automation)的数学描述作了适合本文的简化).

$$\sum_{i=1}^N p_i(t) = 1.$$

记 $P(t)$ 为行为发生概率 $p_i(t)$ 组成的概率向量.

运行过程中, 学习自动机系统根据环境的变化以及行为产生的反馈调整每个行为的发生概率

$$\begin{cases} p_i(t) = p_i(t-1) + f_i(P(t-1), e(t-1)), \\ i = 1, 2, \dots, N, \\ \sum_{i=1}^N f_i(P(t-1), e(t-1)) = 0. \end{cases}$$

(1)

其中 t 表示学习次数, $f_i(\cdot, \cdot)$ 表示行为发生概率的调整, $e(t-1)$ 表示环境或对象的反馈.

3 混杂系统(Hybrid system)

一个混杂系统可以由下式描述(本文引用了文[6]中关于混杂系统的数学定义,本文中的混杂系统的 $q(t)$ 是可控的,所以对文[6]中的定义作了简化):

$$\dot{x} = f_{q(t)}(x(t), u(t)). \quad (2)$$

其中, $x(t)$ 为系统的状态变量; $q(t) \in Q$, 对应不同的子系统, Q 通常是一个有限的离散集, 在本文中 q 称作切换变量.

混杂系统的切换控制实际上是控制切换变量 $q(t)$, 使得系统的状态 $x(t)$ 趋向稳定或达到满意的范围.

4 学习自动机在混杂系统切换控制上的应用 (Applying learning automation on hybrid system switching control)

应用于控制, 学习自动机的输出是对象的控制信号, 它的一个行为是一个有限时间长度上的控制信号, 对于混杂系统的切换控制, 就是一段切换信号 $q(t): t \in [t_1, t_2]$. 控制过程中, 学习自动机控制系统在前一个行为的结束时刻做出下一个行为, 这样产生的在时间上连续的行为的“串联”构成了切换控制信号, 表示为

$$q(t) = \bigcup_{i=0}^{\infty} q_{t_i, t_{i+1}} = q_{t_0, t_1} \cup q_{t_1, t_2} \cup \dots \cup q_{t_{n-1}, t_n} \cup \dots \quad (3)$$

其中 q_{t_i, t_j} 表示时间段 $[t_i, t_j]$ 上的控制信号, 即 $q(t): t \in [t_i, t_j]$. 在本文中表示时间响应的变量都遵从如下的表示方法:

z_{t_a, t_b} 表示变量 z 在 $[t_a, t_b]$ 上的响应, z_t 表示 z 在 t 时刻的值.

学习自动机控制系统在第 i 个行为结束的时刻根据对象的当前状态以及行为的概率分布决定第 $i+1$ 个行为, 同时根据反馈信息调整与行为相关的概率分布. 控制器内部建立一个从状态对应到行为的控制机制 f 和一个根据反馈调整控制机制 f 的学习机制 S , 控制和学习的机制如下:

$$\begin{cases} f_i: (q_i, x_i) \rightarrow q_{i+1}, \\ g: (x_i, q_{i+1}) \rightarrow x_{i+1}, \\ S: (f_i, q_{i+1}, x_{i+1}) \rightarrow f_{i+1}. \end{cases} \quad (4)$$

其中, f_i 为第 i 次学习后的控制机制, g 表示对象的从切换控制信号对应到状态变化的反馈机制, x 为对象状态.

学习自动机的行为集包含能够有利控制对象的信号集, 它的发生条件以及它本身的参数在学习的过程中不断优化. 在应用到混杂系统的控制时, 行为不限制在一个固定的集合中, 其本身也通过学习不断更新, 以保证行为的不断优化, 学习自动集更新行为的方式为变异和优化选择, 这将在下一节实例的具体实现中得到阐述.

5 一个例子 (An example)

图2所示为一个由3个槽构成的配料系统, 3个槽 T_1, T_2, T_3 装有不同液体, 供料器 S 可以左右移动以固定大小的流量 v 向3个槽供料, 根据 S 的位置不同, 3槽对象所构成的混杂系统可以描述如下:

$$\dot{x} = v p_q - w. \quad (5)$$

其中, $x = (h_1, h_2, h_3)^T$ 为由3槽液位构成的状态向量, $w = (w_1, w_2, w_3)^T$ 为3槽的原料流出向量, $q \in \{0, 1, 2, 3\}$ 表示供料器的不同位置 (0 表示不供料), p_q 为供料器处于不同位置时的加料向量.

$$p_q = \begin{cases} [0, 0, 0]^T, & q = 0, \\ [1, 0, 0]^T, & q = 1, \\ [0, 1, 0]^T, & q = 2, \\ [0, 0, 1]^T, & q = 3. \end{cases}$$

上述的混杂系统涉及切换控制的问题. 在向下一级生产程序以稳定的流量提供原料的条件下, 要求加料器合理地在各个槽之间移动, 避免出现断料和溢出的现象. 这就要控制加料器的位置 q 以及加料状态, 使得液位分布 x 能够维持在一个目标值 x_c 的左右.

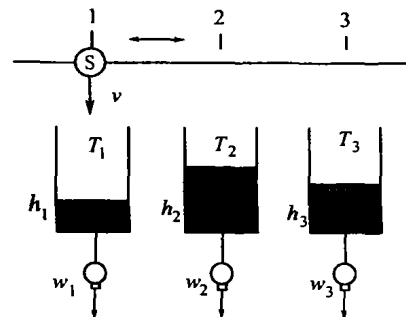


图2 3槽对象
Fig. 2 The 3-tanks object

下面将要给出的就是学习自动机在这样的一个混杂系统切换控制实例上的应用.

如第4节所述,学习自动机的一个行为是一段切换控制信号 $q_{t_i, t_{i+1}}$. 本文确定一个作为控制信号的行为集,它所包含的行为分为图3所示的三种模式. 自动机输出的整个控制信号为这些模式的行为的“串联”,如图4所示的输出的前面部分可以分为是模式(III), (I), (I), (II), (III)的依次串联.

每个模式对应加料器的一个控制动作. 模式(I), 加料器从一个位置 p_0 移到另一个位置 p_1 进行加料, 开始的一段 0 输入表示移动过程中, 供料器处于不供料状态. 起始位置 p_0 和最终位置 p_1 以及移动时间 τ_g 和加料时间 τ_l 是这个模式的参数. 模式(II)表示加料器停止加料, 加料器所处的位置 p_0 和停止加料的维持时间 τ_g 是这个模式的参数. 模式(III)表示加料器持续一段时间的加料, 加料位置 p_0 和持续时间 τ_l 是这个模式的参数.

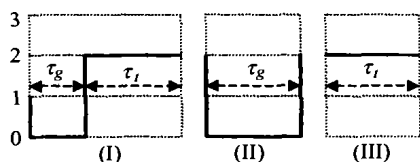


图3 控制输出的三种模式
Fig. 3 Three models of control signal

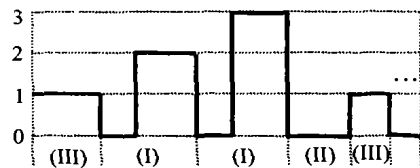


图4 控制信号
Fig. 4 Control signal

学习自动机的控制机制 f_i (见式(4)) 由3槽对象的状态 x_i 及当前供料器位置 q_i 决定下一个行为 $q_{t_i, t_{i+1}}$ 的模式和参数. 在本文给出的方案中, 控制机制采用分区域对应的方式, 把3槽对象的状态变量在三维空间分成一样大小的 K^3 个区域(每一维分成 K 等分). 每一个区域的状态对应控制器的一个行为 $q_{t_i, t_{i+1}}$.

控制系统在一开始建立一个任意的控制机制 f_0 , 由学习机制 S 来实现控制机制的优化. 每一个行为作用于3槽对象后, 学习机制根据对象的反馈给出一个衡量这个行为好坏的优化值, 并与当前记忆的对应该区域的最优行为比较, 选择其中更优的行为记忆下来, 这样的比较选择称之为一次优化选择. 在本文的实现中, 简单的定义行为的优化值为一个行为所经历的优化选择的次数 n , 表示对应该区域的

已发生的 n 次行为中最优的一个.

在学习控制过程中, 控制机制根据3槽对象的当前状态和供料器位置查找对应的最优行为和它的优化值, 在用于控制前, 这个行为以一定的概率发生“变异”, 即发生一些随机的变化, 这样的机制保证行为的不断优化. 行为的优化值越高, 变异的概率越小, 用 γ 表示行为变异的概率, 在一开始 $\gamma = 1$, 随着行为的不断优化, 产生随机行为的概率逐渐降低, $\gamma \rightarrow 0 (t \rightarrow \infty)$.

在本文的控制实现中, 对于每个状态区域对应的行为, 采用了如下的变异概率

$$\gamma = 1/2^n.$$

其中 n 即为前面提到的行为的优化值, 在开始为 0, 每进行一次优化选择 n 加 1.

采用如上所述的机制的学习自动化控制系统应用到三槽对象的切换控制得到很好的仿真结果. 数据如下:

3个槽的高度:

$$(10.0, 10.0, 10.0);$$

3槽对象的初始液位:

$$x_0 = (4.0, 4.0, 4.0);$$

目标液位:

$$x_c = (6.0, 6.0, 6.0);$$

供料器的供料流速:

$$v = 1.0;$$

3个槽的流出速率:

$$w = (0.1, 0.1, 0.2);$$

供料器从一个位置移到相邻位置所需的最少时间 $\tau_{g, \min} = 1.0$.

得到如图5所示的仿真结果, 其中, 灰色的阶跃信号为3槽对象的切换控制信号, 3条曲线分别表示槽 T_1, T_2, T_3 的液位, X 坐标为时间. 3个图分别显示了学习自动机控制系统学习过程3个阶段的控制效果.

由仿真结果可以看到, 学习自动机控制系统在初始阶段对3槽对象的控制是不稳定的, 液位会发生“溢出”的现象(达到最大液位值 10), 随着学习步数的增加, 学习自动机控制系统对3槽对象的切换控制呈现出越来越稳定的趋势. 到最后, 如第三幅图所示, 3槽对象的液位始终被控制在目标液位左右, 说明学习自动机控制系统通过学习自动掌握了3槽对象的控制规律, 体现了学习机制的作用.

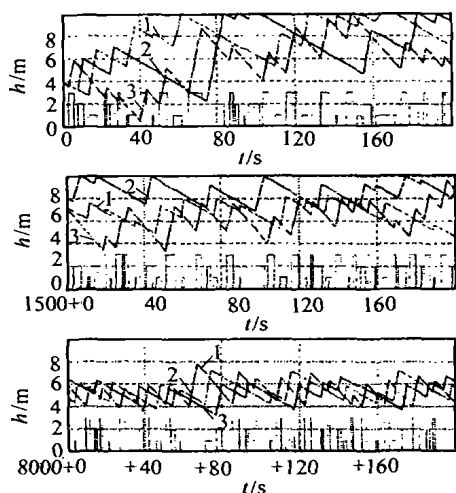


图5 仿真结果

Fig. 5 Simulation results

6 小结 (Summary)

本文在前面的篇幅中已经阐述了学习自动机系统在混杂系统切换控制上的应用方法,并在一个实例中进行了实现,仿真结果显示了好的学习控制效果,证明了它的可行性。

本文所做的工作很大程度上是初步的和尝试性的,文中采用的控制机制和学习机制是较为简单的,在学习自动机的应用方法、具体问题的具体实现、或学习控制的性能研究上都有很大的继续深入的空间。

参考文献 (References):

- [1] NAJIM K, OPPENHEIM G. Learning systems: theory and application [J]. *Computers and Digital Techniques, IEE Proceedings-E*, 1991, 138(4):183-192.
- [2] NARENDRA K S. Learning automation-a survey [J]. *IEEE Trans Systems, Man, and Cybernetics*, 1974, 4:323-334.
- [3] NARENDRA K S, LAKSHMIVARAHAN S. Learning automation - a critique [J]. *J of Cybernetics Information Science*, 1977, 1:53-65.
- [4] ATMADJI A M S, SLOOT J G J. Hybrid switching: a review of current literature [A]. *Proc of IEEE International Conference on Energy Management and Power Delivery* [C]. [s.l.]:[s.n.], 1998, 2: 683-688.
- [5] SHIN K G, DANIEL S W. Analysis and implementation of hybrid switching [J]. *Computers, IEEE Transation*, 1996, 45(6): 684-692.
- [6] HEDLUND S, RANTZER A. Optimal control of hybrid systems [A]. *Proc of the 38th IEEE Conference on Decision and Control* [C]. [s.l.]:[s.n.], 1999, 4:3972-3977.

作者简介:

董苗波 (1976—),男,浙江大学智能系统与决策研究所博士研究生。目前的研究方向为学习系统,神经网络,智能控制。Email:mbdong@mail.sz.tsinghua.edu.cn;

吴铁军 (1950—),男,教授,博士生导师,现为浙江大学智能系统与决策所所长。研究领域为复杂大系统管控一体化,智能控制,工业控制计算机网络。