



浙江工业大学

本科毕业设计论文

题目：人机系统中人的认知偏差的识别方法研究

作者姓名 吴芳

指导教师 赵云波教授

专业班级 自动化 1504

学 院 信息工程学院

提交日期 2019年6月10日

浙江工业大学本科毕业设计论文

人机系统中人的认知偏差的识别方法研究

作者姓名：吴芳

指导教师：赵云波教授

浙江工业大学信息工程学院

2019年6月

**Dissertation Submitted to Zhejiang University of Technology
for the Degree of Bachelor**

**Research on Recognition Method of Human Cognitive
Bias in Human-Machine System**

Student: Wu Fang

Advisor: Professor Zhao Yunbo

College of Information Engineering

Zhejiang University of Technology

June 2019

浙江工业大学

本科生毕业设计(论文、创作)诚信承诺书

本人慎重承诺和声明：

1. 本人在毕业设计（论文、创作）撰写过程中，严格遵守学校有关规定，恪守学术规范，所提交的毕业设计（论文、创作）是在指导教师指导下独立完成的；

2. 毕业设计（论文、创作）中无抄袭、剽窃或不正当引用他人学术观点、思想和学术成果，无虚构、篡改试验结果、统计资料、伪造数据和运算程序等情况；

3. 若有违反学术纪律的行为，本人愿意承担一切责任，并接受学校按有关规定给予的处理。

学生（签名）：

2019 年 6 月 10 日

人机系统中人的认知偏差的识别方法研究

摘 要

随着科学技术的进步以及人们对生产发展的各方面要求越来越高，人机系统的角色越来越丰富。人机系统是由人和机器构成并依赖于人机之间相互作用而完成一定功能的系统，人机系统的性能取决于人和机器两部分，在本文中我们假定机器是完美的，只考虑人的行为对系统性能的影响。当人机系统处于特定的目标、特定的环境下，人可能会比较容易产生某种内部异常或认知偏差，导致错误的决策，影响人机系统的性能。

本文的研究内容是识别人机系统中人的认知偏差，本文将认知偏差从心理学上的定义扩展到包含生理、心理两部分的内部状态异常。由于人机系统中的人产生的内部异常一般会影响到人机系统的输出，因此我们从人机系统的输出数据检测人的内部状态，判断人是否产生了异常。论文的主要工作如下：

1. 综述论文研究背景及意义，介绍对人机系统中的人以及对人和机器协同控制的研究现状。
2. 借鉴了在不确定性条件下做决策的 POMDP 方法，基于 POMDP 对人机系统进行建模，构造了一种从人机系统的输出数据的特点中去推断系统中人的内部状态是否异常的方法。
3. 基于常见的人驾驶汽车场景下的人机系统设置仿真实验，验证本文提出的方法的有效性。

关键词：人机系统，认知偏差，部分可观马尔可夫决策过程，贝叶斯定理

RESEARCH ON RECOGNITION METHOD OF HUMAN COGNITIVE BIAS IN HUMAN-MACHINE SYSTEM

ABSTRACT

With the advancement of science and technology and the increasing demands of people on all aspects of production development, the role of human-machine systems is becoming more and more abundant. Human-machine system (HMS) is a system composed of human and machines and relies on the interaction between human and machine to complete certain functions. The performance of human-machine system depends on two parts: human and machine. In this paper, we assume that the machine is perfect, therefore only consider the impact of human behavior on system performance. Under a specific target and specific environment, people in the HMS may be more likely to have some internal abnormalities or cognitive biases, leading to wrong decisions and affecting the performance of the human-machine system.

The research content of this thesis is to identify the cognitive bias of human beings in HMS. This paper extends the cognitive bias from the definition of psychology to the internal state abnormality including physiological and psychological parts. Since the internal anomalies generated by people in the human-machine system generally affect the output of the HMS, we detect the internal state of the human from the output data of the human-machine system to determine whether the person has an abnormality. The main work of the thesis is as follows:

1. Overview the research background and significance of the thesis. Then introduce the research status of the human-machine system and the collaborative control of people and machines.

2. This thesis draws on the POMDP method for making decisions under uncertainty

conditions. Based on POMDP, the human-machine system is modeled, and a kind of man-machine system's output data is constructed to infer whether the internal state of the person in the system is Anomalous method.

3. The effectiveness of the proposed method is verified based on the simulation experiment of human-machine system setting in the common human driving car scenery.

Key Words: human-machine system, cognitive bias, POMDP, Bayes' theorem

目 录

摘 要	I
ABSTRACT.....	II
第 1 章 绪 论	1
1.1 课题研究背景及意义.....	1
1.1.1 人机系统的定义.....	1
1.1.2 认知偏差的介绍.....	1
1.1.3 人机系统中的认知偏差的定义.....	3
1.1.4 本课题研究的背景.....	3
1.1.5 本课题研究的意义.....	4
1.2 研究现状.....	4
1.3 论文组织结构.....	7
第 2 章 基于 POMDP 的人机系统模型.....	8
2.1 基本方法介绍.....	8
2.1.1 POMDP 框架介绍.....	8
2.1.2 贝叶斯定理介绍.....	10
2.2 基于 POMDP 的人机系统建模.....	11
2.3 人机系统演化的动态贝叶斯网络表示.....	14
2.4 模型的实例化说明.....	16
2.5 本章小结.....	17
第 3 章 基于驾驶系统的仿真实验	18
3.1 实验设置.....	18
3.2 仿真结果.....	20
第 4 章 总结与展望	24
4.1 毕设工作总结.....	24
4.2 未来展望.....	24
参考文献	26
附录	28
致谢	34

第1章 绪 论

1.1 课题研究背景及意义

1.1.1 人机系统的定义

人机系统是指人为了达到某种预定目的，由相互作用、相互依存的人和机器两个子系统构成的一个整体系统^[1]。

机器，广义上是通过变换或传递能量、质量和信息，执行机械运动等达到特定目的的工具、装置或设备的总称，例如控制系统中可以测量受控对象的状态或输出信息的传感器、集成了各种计算设备因而能够进行控制算法运算的控制器和能够按照制定好的控制策略执行特定任务的执行器等。

人机系统中的人具有特殊的地位。一方面，人是控制系统目标的来源，为控制器设定目标、赋予价值，因为任何控制系统的设计总是为人服务的，是为了达到人的某种目的。另一方面，人具有一些现代机器还无法达到的特殊能力，需要借助人的能力达到更好的系统性能。

随着人工智能的发展，具有人机交互过程的人机系统变得越来越日常化，并且逐渐从完全是机器辅助人完成任务的模式转化成人与机器的协同控制、人机融合等等。人与机器合作去完成一个共同的任务，任务完成的质量如何即是人机系统的性能，人机系统的性能取决于人和机器两部分，在本课题中我们关注人的部分，即假设人机系统中的机器都是自动化的、且不会出现问题。

1.1.2 认知偏差的介绍

认知偏差是在某些特定情况下特定的思考、行为倾向，是在判断中偏离规范或理性的系统模式^[2]，会导致理性或判断产生系统性偏差，即个人通过对输入的感知来创造他们自己的“主观社会现实”。这些现象广泛受到心理学与行为经济学研究。本课题一开始希望能从认知偏差的成因、心理机制、研究领域和矫正与应用等方面入手，对人的认知特点进行分析、对认知偏差进行识别。

一些认知偏差是有利的。认知偏差可能会导致在特定环境下更有效的行动^[3]。此外，当及时性比准确性更有价值时，允许认知偏差可以使决策更快，如启发式

[4]。心理学上的启发式是以一种快速的看起来不费力的方式做复杂的决定或进行推论时应用的一些简单规则。其他认知偏差是人类处理局限性的“副产品”^[5]，是由于缺乏适当的心理机制(有限理性)，或仅仅是由于有限的信息处理能力。总的来说启发式是非常有用的，但有时它们也可能导致严重的、系统性的错误，这也正是认知偏差产生的时候了^[4]。

不同的学科领域对认知偏差的定义不尽相同，我们提取出它们的共性：认知偏差产生的前提是人类处于某种特定的情况下，而这种特定的情况可以理解为对人类的下一步思考或者操作产生阻碍的具有不确定性的情况，人类在这样一种不确定性的情况下会有什么样的反应。人们是如何评估一个不确定事件的可能性或是一个不确定的量值的呢？他们依赖有限个数的能降低评估可能性这个任务的难度的启发式原则，以更加简单的方式去预测。认知偏差可以理解为一种信息加工的偏离，它的出现具有普遍性和顽固性，当我们把认知偏差理解为偏离预定的目标时，将有利于对认知加工机制的分析和模型的构建^[6]。

认知偏差已经在各个领域被广泛研究，包括它的成因、分类、影响以及应对策略等等。在不同的研究领域中，认知偏差的分类依据不尽相同，例如是存在去特定群体的偏差还是个人层面的偏差，偏差是否会影响到记忆等等。在本文的研究背景下，我们选择认知偏差的分类依据为是否影响人的决策行为。根据这一分类依据，总结出以下几种比较容易产生的影响人在人机系统中的决策行为的认知偏差。一、锚定效应，指的是为不熟悉事物估值时，会把熟悉的类似事物或不久前接触到的无关数值当作“锚”，估计出来的数值会大大倾向于“锚”。例如人在人机系统中需要做决策并且这个情况是以前没怎么接触过的时候，他做决策的过程可能会比较依赖于之前所经历的熟悉的情况。二、框架效应，指的是同一信息以不同方式呈现会带来不同想法。例如采取不同的措施能达到同样的结果，但是别人可以将这种不同的措施展现在我们面前，面临我们看到的不同措施，我们很可能会忽略到它们实际上会引发相同的结果，而做出了不同的选择决策行为。三、现时偏差，指看重当下利益、低估长远利益，得到利益前的延迟时间越长，对利益的价值评价便给予更多折扣。例如人在做决策的过程中，经常无法从长远的角度考虑，而做出了能获得即时利益的决策。

认知偏差在人机交互领域中的研究并不常见，对人的认知偏差的识别这一方

面的研究也甚少，认知偏差被认为是非常难以检测和覆盖的。我们知悉了常见的认知偏差及其成因等方面的内容，意图使得机器能够在人机交互的过程中识别人的认知偏差，实现必要的或者更好的人机切换策略。

1.1.3 人机系统中的认知偏差的定义

根据上一小节中对认知偏差的描述，可以看出认知偏差通常是心理学意义下的由于不同的内因和外因引起的认知方面的错误。结合人机系统的特殊性，我们可以给出人机系统中的认知偏差的定义：指的是在系统工作的过程中人产生的、会影响人机系统性能的人内部状态的异常。

鉴于人机系统中的人的认知偏差的特殊性——人的认知偏差导致的发生在系统中的决策失误会体现在人机系统的输出数据中，并且实际上在优化人机系统性能的时候，我们并不关心人到底是产生了什么类型的认知偏差，因为它们的本质都是引起了人的内部状态异常，这异常影响了人机系统的性能。我们可以忽略认知偏差本身的一些性质，着重考虑它在人机系统中的表现形式——作为黑箱因素影响系统输出。

1.1.4 本课题研究的背景

根据前面几小节对课题中几个名词的详细定义和描述，可以归纳出本研究课题的背景。在人机系统中，相对于自动化的机器而言，人的生理结构和心理因素的不确定性导致人的认知和行为具有很大的不确定性，这种不确定性可能会导致系统中的人在人机交互的过程中做出错误的决策，进而影响了人机系统的性能。本课题一开始的目标是归类并选择其中的典型认知偏差，刻画偏差的特点，并设计合适的方法对认知偏差进行识别。

本课题的研究过程一开始的着眼点在于从认知偏差入手，对特定的认知偏差进行检测识别，因此对认知偏差相关的文献进行了调研，了解到现有的关于认知偏差的文献大多侧重于对认知偏差的成因、心理机制、影响与矫正这几个方面的内容，并且常用的研究方法是通过进行心理学测试得到一些数据，分析这些数据来得出相关的结论。我们还了解到在诸如心理学、教育学、经济学和管理学等领域对认知偏差的研究和应用比较多，而这些问题的研究背景和研究过程均未涉及

到机器，这与本课题的研究背景是不一致的，我们是希望能够研究出一整套完整的方法可以直接应用于人机系统中去识别人的认知偏差。接下来的研究内容均基于本课题的特殊性，即人不是单独的个体，而是与机器有相互作用的。

1.1.5 本课题研究的意义

人机系统中一般而言机器常常起着传感器、控制器、执行器的作用，而人可能需要在不同的场景下做出决策，人的认知的不确定性会影响决策过程。如果我们能根据系统可获得的信息来监测人的内部状态的变化，对人的行为进行监测，推断识别人的状态、意图等，在影响人机系统性能异常产生的早期将其识别出来并给予反馈，最大程度地减轻人的异常可能对人机系统性能的影响，则能够使得人机系统具有更加良好的表现。

本课题将要识别的认知偏差或异常定义为可逆性的偏差或异常，即若在识别出人的异常时对人进行反馈后人可以主动调整自己的状态至正常状态。并且我们在乎的是会影响人机系统性能的异常，即如果检测到人在短时间内可能产生的异常、但是没有影响人机系统输出的趋势时，我们不会对这种异常进行反馈。

1.2 研究现状

本课题将要识别的认知偏差或异常定义为可逆性的人的内部异常，即若在识别出人的内部异常时对人进行反馈后人可以主动调整自己的状态至正常状态。因此可以将对人的认知偏差或异常进行识别的问题转化成判断何时该对人进行反馈的一个最优决策问题。鉴于面对的人的内部状态是否异常相当于一个黑箱问题，我们将问题定义为在不确定性条件下进行序列决策的问题。

传统的自动控制系统只通过监测机器的状态来防止人为失误和提高安全性，很少对人也进行监测和反馈。而人机系统需要人和机器的相互作用、相互依存，有些情况下二者的联系是比较紧密的，并且人的状态对系统的性能有很大的影响。以人驾驶汽车为例，人和汽车是相互作用的两部分，人要控制车，车也会给予人反馈，例如在某些检测到车陷入异常情况时机器对人做出的警报提醒等。在这个系统中，如果加入对人的监测以及对人的反馈，在一定程度上能够在更早的时候发现异常情况，降低危险系数。

在 DARPA/NSF 关于人机交互的跨学科研究的最终报告中, Burke 和她的同事们发现了人工智能人机交互所面临的巨大挑战。意图识别被认为是第一组技术挑战之一: 捕捉人类的社会线索(注意姿势、肢体语言)和解释人类行为(意图、目标、欲望)^[7]。由于人机系统缺乏与人的可靠和充分的意图沟通, 需要从人的行为甚至从行为引起的环境变化中推断意图或内部状态。为了决定什么时候对人进行反馈, 我们需要估计人的生理状态以及她的意图, 对于人的意图或者状态大多数情况下并不能直接观测到或者是预测到, 实际测量人类的大脑活动很多情况下也不十分现实, 因此我们考虑从人的行为去推断人类的内部状态。

由于人的内部状态是不能直接观察到的, 所以这个场景需要通过一个间接的估计过程来确定人的内部状态。Pentland 等人提出, 许多人类行为可以被精确地描述为一组由马尔可夫链排列在一起的一组动态模型, 其中他们将多个动态模型定义为内部状态, 他们使用这些马尔可夫模型来识别观察数据中的人类行为^[8]。在各种智能车辆控制系统的研究中, 也涉及到对人车系统的整体建模以提升系统安全性能等问题, Takano 等人提出了一种利用 HMM 对汽车进行随机智能认知的方法, 将由环境状态、车辆状态和驾驶员状态组成的驾驶模式原语建模为隐马尔可夫模型(HMM)^[9]。这些年来为了帮助老年人或是其他行动不便以及其他方面有些不方便的人群过上更加舒适的生活, 对辅助机器人的研究也非常广泛。在这个人机交互的系统当中, 对辅助机器人的设计就得比较全面地去考虑对人的识别, 希望不仅能根据一些观察到的因素明确地理解人类意图, 还能从日常生活中推断出人的隐含意图。

大多数研究人员使用一种共享控制的方案来结合人与机器的控制。Chipalkatty 等人直接修改了人的输入, 基于一系列预测的人的输入, 使实际的输入不仅符合人的意图, 而且满足动态约束^[10]。Vasudevan 等人测量了驾驶车辆的安全性, 以确定何时进行干预^[11]。Anderson 等人利用模型预测控制找到安全最优的车辆路径, 然后基于威胁评估, 通过人工输入和控制器输入的加权和对车辆进行控制^[12]。这些方法的共同因素在于, 它们为未来的状态进行规划, 并使用共享的控制方案使未来的状态满足某些标准, 比如未来计划的安全性和动态约束。这些控制器只利用对机器的反馈, 但不考虑将反馈合并到人类, 如对人类提出警告。

开发交互策略即实现对人和机器的统一的监测和建模反馈控制的一种方法

是让机器使用强化学习方法直接从与人的交互中学习得到人类的一些特征和信息^[13]。但是这个学习的过程通常需要大量的数据，而从人类方面收集足够的数据以达到足够的性能水平可能需要非常长的时间或者代价非常昂贵，甚至在某些情况下是无法去收集到的。策略学习的另一种方法是直接从人类交互演示中学习^[14]。但如果要学习的任务很困难，那么数据需求可能很难满足这种技术，因为即使在大型数据集中，也可能缺乏成功交互的例子。通过使用基于专家知识和可用数据的规划模型，可以找到在没有在线学习阶段且交互示例相对较少的情况下执行良好的行动策略^[15]。

[16]中阐述了部分可观测的马尔可夫决策过程（POMDP）能够表示世界的整个状态不是直接可观测的情况，并在规划时能够利用这种隐藏的状态信息，模型中包含了隐藏状态和可以被 agent 观察到的状态的各个方面之间的概率关系。POMDP 策略不是实现状态到操作的映射，而是将信念（状态上的概率分布）映射到操作，这种信念分布根据 agent 在执行过程中接收到的观察值进行维护。状态不能被精确地观察到，随机观察模型将可观测信号与基础状态联系起来。POMDP 可以使用标准贝叶斯跟踪来监视关于系统状态的信念。可以使用多种技术来计算使预期回报最大化（或近似最大化）的策略^[17]。

贝叶斯等概率推理方法具有深刻的优势，能够在信息有限的情况下，尽可能做出好的预测。个体通常在不确定的情况下行动，因为他们永远无法对他们的环境有一个完整和准确的感知^[7]。一般的贝叶斯推理，特别是贝叶斯人工智能都是基于贝叶斯公式：

$$P(h|e) = P(h) \frac{P(e|h)}{P(e)} \quad (1)$$

e 是已知信息， h 是要求解的问题。贝叶斯定理即在已知先验概率 $P(h)$ 和可能性函数 $P(e|h)/P(e)$ 的情况下，可以求得后验概率 $P(h|e)$ 。通常情况下，先验概率 $P(h)$ 是根据已有的经验预估出的 h 事件的概率，可能性函数 $P(e|h)/P(e)$ 是一个调整因子，即新信息 B 带来的调整，作用是使得先验概率更接近真实概率。利用贝叶斯的原理，即通过专家知识或经验给出先验概率，使用观察到的信息来增强或减弱先验概率，得到后验概率^[18]。

然而在大多数实际情况下，要求解的问题和已知信息并不是直接和简单的关

系,因此需要使用包含一些相互关联的变量和状态的表示法,并利用变量的条件独立性。贝叶斯网络是一个好的选择,贝叶斯网络(BN)是一个有向无环图,它对条件独立性的假设进行编码。BN 中的节点表示随机变量,而弧表示因果关系。考虑到动态交互的人机系统的随时间演化性,因此我们可以使用动态贝叶斯网络(DBN)来描述我们所要研究的人机系统的演化过程。DBN 是一种特殊的 BN 体系结构,用于表示变量随时间的演化。

1.3 论文组织结构

本文的研究内容主要是实现对人机系统中人的内部异常的识别,即推断人的内部状态。主要章节内容如下:

第 1 章 绪论。主要介绍了人机系统、认知偏差以及人机系统中的认知偏差的定义,阐述了人机系统中人的认知偏差或内部状态异常研究的背景和意义,总结了相关的研究现状,最后介绍了本文的主要研究内容与行文安排。

第 2 章 基于 POMDP 的人机系统建模。本章提出了对 POMDP 中的信念状态的更新方法做出改变,以适应我们所要研究的内容,根据本文所提出的方法,本章建立了基于 POMDP 的人机系统的模型,并给出了它的动态贝叶斯网络表示。

第 3 章 基于驾驶系统的仿真实验验证。为了验证本文提出的方法的有效性,本章基于驾驶系统在 MATLAB 平台上做了两个不同角度的仿真实验,结果显示本文提出的方法能够有效地识别出人机系统中人的内部状态异常,并且能对其做出良好的反馈。

第 4 章 总结与展望。对本文的工作进行了总结,提出研究中存在的不足,给出未来该课题研究方向的一些建议。

第 2 章 基于 POMDP 的人机系统模型

人机系统中人的内部状态是隐藏状态，我们希望通过对人进行反馈来使得系统的演化往最好的方向进行，即可看作是一个在不确定性环境下做序列决策的问题，因此本文采用 POMDP 框架作为基本框架。本章分 4 小节，2.1 节主要介绍了本文所采用的 POMDP 框架和贝叶斯定理的相关概念和计算；2.2 节建立了 POMDP 框架下的人机系统模型、给出了相关变量的定义；2.3 节给出了 POMDP 框架下的系统模型演化的动态贝叶斯网络表示；2.4 节用实际例子对模型进行了更加形象的说明；2.5 节对本章内容进行了小结。

2.1 基本方法介绍

2.1.1 POMDP 框架介绍

部分可观马尔可夫决策过程(POMDP)是在具有不完全可观的状态的系统模型中使用 MDP 做最优决策的过程。在本文中我们假设是一个独立于人和机器本身的控制器来决定是否对人作出提醒，实际上这个控制器也可能是集成在机器上的。控制器会在每一个采样的时间点上，根据对系统状态的判断，以最大化回报的方式作出行动决策，这里的回报跟系统状态和控制器采取的行动有关。由于系统是有隐藏状态的，因此控制器需要通过传感器返回的观察数据来推测隐藏的状态当下正处于何种状态。这种对隐藏状态可能处于何种状态的描述被称为信念状态，用隐藏状态的概率分布来表示。

POMDP 框架是包含隐藏状态和动作效果的不确定性的、能用于各种实际的顺序决策问题建模的通用框架，例如在机器人导航问题、机器维护和其他情况下的不确定性规划等方面的应用。离散时间的 POMDP 可以用以下七元组来表示 $(S, A, T, R, \Omega, O, \gamma)$ ，其中：

- S 代表一组有限的离散状态；
- A 代表一组离散的行动；
- T 是状态转移函数，表示状态之间的一组转移概率， $T(s, a, s') := Pr(s_{t+1} = s' | s_t = s, a_t = a)$ 表示在状态 s 执行动作 a 达到状态 s' 的概率。

- Ω 是一组观察，用以提供关于不确定环境的信息。
- O 表示一组条件观察概率， $O(s', a, o') := \Pr(o_{t+1} = o' | a_t = a, s_{t+1} = s')$ 表示执行动作 a 达到状态 s' 观察到 o' 的概率。
- 回报函数 $R(s, a)$ 表示在 s 上执行行动 a 所得到的即时回报。
- $\gamma \in [0,1]$ 是折扣因子。

在每一个时间步上，系统处于某些状态 $s \in S$ ，控制器选择其中一个动作 $a \in A$ ，这个动作会使得系统状态以 $T(s' | s, a)$ 的概率转移到状态 $s' \in S$ ，同时控制器以 $O(o | s', a)$ 的概率接收到一个依赖于系统所处的新状态的观察 $o \in \Omega$ ，最后控制器得到一个回报 $R(s, a)$ ，上述过程对每个时间步进行迭代。

引入信念状态 b 作为一个包含所有所需要的历史信息的统计量，即用来描述隐藏的内部状态的概率分布，用 $b(s)$ 来表示环境处于状态 s 的概率。 t 时刻的信念状态描述如下：

$$b_t = P(s_t | a_t, o_t, a_{t-1}, o_{t-1}, \dots, a_0, o_0, s_0) \quad (2)$$

t 时刻的信念状态 b' 可以由 $t-1$ 时刻的信念状态 b 来更新，如下：

$$b'(s') = \Pr(s' | b, a, o') = \frac{O(s', a, o') \sum_{s \in S} T(s, a, s') b(s)}{\Pr(o' | a, b)} \quad (3)$$

信念状态由最近的行动和观察来更新， $\Pr(o' | a, b)$ 是一个归一化参数。

对于上述信念状态的更新过程，放在不确定环境下的机器人导航问题中会更容易理解。已知上一时刻机器人在环境中所处的信念状态 b ，机器人采取一个动作 a ，在这个过程中它会获得观察数据 o' ，根据这个观察数据和状态转移概率，可以推断机器人下一时刻所处环境的信念状态。

POMDP 的目标是找到每一个时间步上的最优动作以最大化累积回报。POMDP 可以视为具有信念状态的马尔可夫决策过程(MDP)，MDP 的求解算法主要分为值迭代和策略迭代两类，因此值迭代算法也可以扩展用于求解有限时间步的 POMDP 最优策略

- 初始化： $t = 0$ ；对于所有的 $b \in B$ ，都有 $V_0(b) = 0$ 。
- 若 $|V_{t+1}(b) - V_t(b)| > \varepsilon$ ，对于所有的 $b \in B$ 按下式计算 $V_{t+1}(b)$ ， t 更新为 $t + 1$

$$V_{t+1}(b) = \max_{a \in A} \left[R^b(b, a) + \gamma \sum_{b' \in B} T^b(b, a, b') V_{t-1}^*(b') \right] \quad (4)$$

在实际应用中，精确求解 POMDP 问题通常会由于受到两大诅咒而非常难进行，“维数诅咒”和“历史诅咒”。维数诅咒，指在一个状态数为 n 的 POMDP 规划问题中，规划所需的计算须在 n 维状态空间中完成。历史诅咒，指的是不同历史的数目会随着规划步数的增加呈指数级别的增长。解决 POMDP 问题通常在计算上难以处理，也因此发展了一些离线或在线的近似规划算法，由于大规模问题用离线规划算法还是工作量很大，因此在线规划算法的各种延伸非常多。在线规划算法的主要思想包括分支限界裁剪法、蒙特卡罗采样法和启发式搜索法^[19]。

2.1.2 贝叶斯定理介绍

贝叶斯定理基于以下基本思想：人们一开始对事件或变量有一个事先的认知，借助的可能是自身的经验或者其他内容，我们将这个事先的认知称为先验分布或先验概率；如果观察到新的样本信息，那么我们可以利用这些观察信息来更新先验知识，得到一个后验分布或后验概率。

在正式介绍贝叶斯定理之前，先介绍几个相关的概念定义：

(1) 边缘概率（又称先验概率）

边缘概率指的是某个事件发生的概率。边缘概率是这样得到的：在联合概率中，把最终结果中那些不必要的事件通过合并成它们的全概率，而消去它们（对离散随机变量用求和得全概率，对连续随机变量用积分得全概率），这称为边缘化，比如 A 的边缘概率表示为 $P(A)$ ， B 的边缘概率表示为 $P(B)$ 。

(2) 联合概率

联合概率表示两个事件共同发生的概率。事件 A 与 B 的联合概率表示为 $P(A \cap B)$ 或 $P(A, B)$ 。

(3) 条件概率（又称后验概率）

条件概率指的是事件 A 在另外一个事件 B 已经发生条件下的发生概率。条件概率表示为 $P(A|B)$ ，读作“在 B 发生的条件下 A 发生的概率”。

比如，在同一个样本空间 Ω 中的事件或者子集 A 与 B ，如果随机从 Ω 中选出的一个元素属于 B ，那么这个随机选择的元素还属于 A 的概率就定义为在 B

的前提下 A 的条件概率，所以： $P(A|B) = |A \cap B|/|B|$ ，接着分子、分母都除以 $|\Omega|$ 得到：

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (5)$$

本文主要会用到如下贝叶斯公式：

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^n P(A_j)P(B|A_j)} \quad (6)$$

其中，事件 B 发生之前，我们对事件 A 的发生有一个基本的概率判断，称为 A 的先验概率，用 $P(A)$ 表示；事件 B 发生之后，我们对事件 A 的发生概率重新评估，称为 A 的后验概率，用 $P(A|B)$ 表示。

根据公式可以看出，若以事件 B 发生为条件求事件 A 发生的概率不能直接求，而我们可以知道在事件 A 发生为条件时事件 B 发生的概率、以及事件 A 发生的先验概率已知的话，则可以根据已知的信息求未知的条件概率。

2.2 基于 POMDP 的人机系统建模

通过对人机系统中各个部分进行抽象并用变量表示，根据各个变量之间的关系，可以得出如下图所示的人机系统框图：

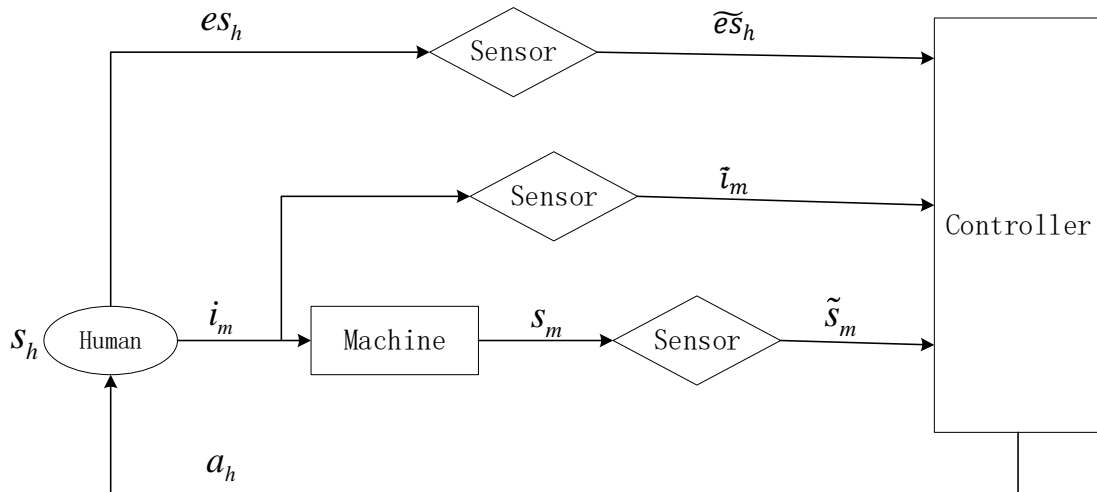


图 1 人机系统框图

图 1 是人机系统框图，其中变量定义如下：

$s_h \in S_h$ ，人类内部状态的集合，可以是人的意图或生理状态。

$es_h \in ES_h$ ，人类外部状态或行为的集合。

$\tilde{es}_h \in O_{es_h}$ ，人类外部状态的一组观察。

$i_m \in I_m$ ，人类施加的对机器的控制输入。

$\tilde{i}_m \in O_{i_m}$ ，人类对机器施加的控制输入的一组观察。

$s_m \in S_m$ ，机器状态的集合，也可视为人机系统的输出（例如位置、速度等）。

$\tilde{s}_m \in O_{s_m}$ ，机器状态/系统输出的一组观察。

$a_h \in A_h$ ，对人类的控制反馈集合（例如提醒或警告信号）。

我们假设所有上述集合都是有限的。在图 1 中，人类具有内部状态 s_h ，这可能是人的意图或是人类想要达到的目标，或者是人的生理状态，如疲劳、愤怒、醉酒等。根据人类的内部状态，其会有外部状态表现或行为 es_h ，也会根据自己感知到的机器状态或系统输出对机器施加控制输入 i_m 来实现人的意图。例如，驾驶系统中驾驶员的下一阶段的目标是左转，如果驾驶员是清醒的，则其大概率会有外部状态表现为眼睛睁开，以及为了左转而对车施加的控制输入——向左转动汽车方向盘。一些人为动作是机器的控制输入，因此机器的状态或系统的输出 s_m 将随时间改变。通过传感器来测量人的外部状态表现或行为、人对机器施加的控制输入和机器的状态或系统输出。我们将人的外部状态或行为的测量表示为 \tilde{es}_h ，将人类对机器施加的控制输入的测量表示为 \tilde{i}_m ，并将机器状态或系统输出的测量表示为 \tilde{s}_m 。

在本课题的研究背景下，信念状态指的是人的内部状态即隐藏状态的概率分布。根据上一节中对 POMDP 的基本内容的介绍，可知要更新信念状态需要先得到状态转移概率和观察函数。下面分别介绍如何获得状态转移概率和观察模型。

人机系统中的控制器使用观测结果及其各个变量之间的关系作为输入，根据是否加入对人的反馈 a_h 情况下各自的回报，根据两个值的相对大小，估计隐藏状态 s_h 概率分布，判断人的内部状态是否产生了偏差或异常。

上述过程针对每个时间步进行迭代，因此整个过程可以被视为具有一组隐藏状态的马尔可夫决策过程 (MDP)，它具有隐藏状态集合 $S_h * I_m * S_m$ （我们关心的是 S_h ），行动集或控制集 A_h ，观测集 $O_{es_h} * O_{i_m} * O_{sm}$ 和转移概率：

$$P(s'_h, i'_m, s'_m | s_h, i_m, s_m, a_h)$$

$$= P(s'_h | s_h, i_m, s_m, a_h) \times P(i'_m | s'_h, s_h, i_m, s_m, a_h) \times P(s'_m | s'_h, i'_m, s_h, i_m, s_m, a_h) \quad (7)$$

上述分解基于概率链式规则。如公式所示，转移概率看起来很复杂，但是我们可以通过做出一些合理的条件独立假设来简化转移概率。

第一个条件独立假设是人的内部状态只取决于其之前的内部状态以及控制器对人的反馈，即：

$$P(s'_h | s_h, i_m, s_m, a_h) = P(s'_h | s_h, a_h) \quad (8)$$

我们将其称为人类内部状态模型，本课题的主要目的是识别会影响人机系统性能的人的内部状态异常。若识别出 t 时刻人的内部状态 s_h 产生了影响人机系统性能的异常，则控制器对人进行反馈、即选择执行动作 a_h ，则人的内部状态会以一定的概率从 t 时刻的 s_h 转移到 $t+1$ 时刻的 s'_h 。

第二个假设是人对机器的控制输入只取决于其下一时刻的内部状态以及机器当前时刻的状态，即：

$$P(i'_m | s'_h, s_h, i_m, s_m, a_h) = P(i'_m | s'_h, s_m) \quad (9)$$

我们将其称为人类行为模型，在获得当前时刻的机器状态 s_m 的情况下，人类依赖于下一时刻自身的内部状态 s'_h 以一定的概率施加对机器的控制输入 i'_m 以实现人机系统的目标。

关于转移概率的最后一个假设是机器的状态只取决于当前时刻的机器状态和人下一时刻对机器的控制输入，我们将其称为机器动态模型，即：

$$P(s'_m | s'_h, i'_m, s_h, i_m, s_m, a_h) = P(s'_m | i'_m, s_m) \quad (10)$$

在当前机器状态为 s_m 的情况下，若下一时刻人对机器的控制输入为 i'_m ，则相应的机器状态在下一时刻会以一定的概率转移到 s'_m 。

综上所述，POMDP 框架下的人机系统的转移概率转换为：

$$P(s'_h, i'_m, s'_m | s_h, i_m, s_m, a_h) = P(s'_h | s_h, a_h) P(i'_m | s'_h, s_m) P(s'_m | i'_m, s_m) \quad (11)$$

获得最终的转移概率模型表达的过程也是理清模型中各个变量之间的关系的过程。在观察模型中，我们假设对人类外部状态或行为的观察只依赖于真实的人类外部状态，对人类施加给机器的控制输入的观察只依赖于人类对机器的实际控制输入，对机器状态的观察只依赖于真实的机器状态，即：

$$P(\tilde{e}_{s_h}, \tilde{i}_m, \tilde{s}_m | s_h, e_{s_h}, i_m, s_m, a_h) = P(\tilde{e}_{s_h} | e_{s_h})P(\tilde{i}_m | i_m)P(\tilde{s}_m | s_m) \quad (12)$$

得到上述转移概率和观察模型的表达式之后，根据前文介绍的 POMDP 问题中对信念状态的更新的过程，我们可以更新信念状态 \mathbf{b} ，即更新隐藏状态的概率分布，本课题中隐藏状态即人机系统中人的内部状态。由于我们的目标是识别人机系统中人的认知偏差、即内部状态的异常，我们将基于 POMDP 的人机系统模型中的动作定义为对人的反馈，相应的我们在本课题中研究识别认知偏差的方法的时候对 POMDP 中原先的信念状态更新方法做出改变，采用贝叶斯定理来更新本课题所要研究的信念状态——人机系统中人的内部状态的概率分布。这一步改变主要是基于 POMDP 的信念状态更新的过程与贝叶斯定理的原理相同，都是利用观察数据即一些证据来对增强对原本不确定的变量或事件的置信度，并且在每一步更新过程中，贝叶斯公式涉及到的变量种类更少，相对而言计算更为简单。

奖励函数 $R(s_h, e_{s_h}, i_m, s_m, a_h)$ 的设定取决于系统的目标。例如，如果系统目标是提高安全性，那么安全状态下的奖励应该很高，而不安全状态下的奖励应该很小。如果系统同时有多个目标，则根据各个目标的相对重要性设定各自的奖励或惩罚。例如，在下一章的仿真验证实验中，我们希望既能提高安全性又能最大限度地减少干扰，因此我们会在安全状态下给予高回报的同时，惩罚来自反馈的动作干扰。

为了更清晰地描述上述人机系统模型中各个变量之间的相互关系以及动态演化过程，我们将在下一节中将 POMDP 框架下的人机系统模型表示为动态贝叶斯网络的形式。

2.3 人机系统演化的动态贝叶斯网络表示

贝叶斯网络(BN)也称作概率网络或信念网络，是基于有向无环图来刻画变量之间的依赖关系的一种网络结构，并使用条件概率表来描述联合概率分布，是目前不确定知识表达和推理领域最有效的理论模型之一。动态贝叶斯网络(DBN)是通过相邻的时间步长将变量相互关联的贝叶斯网络。这通常被称为双时间片被也是网络(2TBN)，因为它表示在任何时间点 T ，变量的值可以从内部回归量和直接先验值(时间 $T-1$) 计算^[20]。

动态贝叶斯网络可以较好地表示相邻时间点上状态变量之间的转移关系，根

据前文所介绍的人机系统在 POMDP 框架下的变量及其条件独立关系, 我们可以将上述对人机系统模型的描述表示为如下图所示的动态贝叶斯网络的形式:

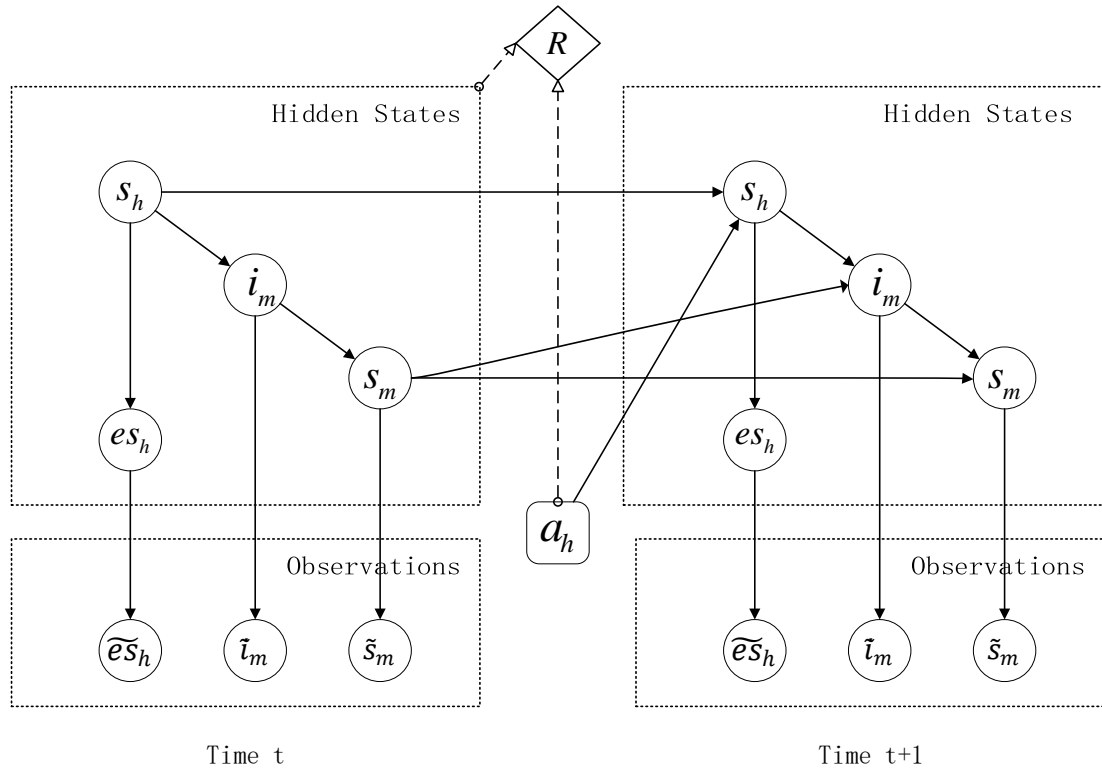


图 2 基于 POMDP 的人机系统模型的动态贝叶斯网络表示

根据贝叶斯定理, 我们可以根据传感器测得的 $\tilde{e}s_h$ 来得到 s_h 的先验概率分布。通过收集到的观测信息 \tilde{s}_m 和人面对相应的 \tilde{s}_m 时的反应或对机器的控制输入 \tilde{i}_m 的相互关系, 来进一步更新 s_h 的概率分布, 得到它的后验概率分布。

总结前文内容, 本文提出的方法可以概括为以下几个步骤:

- (1) 根据经验或合理估计, 由观察数据集合 $\tilde{e}s_h$ 推断人的内部状态 s_h 的先验概率分布。
- (2) 通过观察数据集合 \tilde{s}_m 和 \tilde{i}_m 的相互关系, 来进一步更新 s_h 的概率分布, 得到它的后验概率分布, 即得到更新之后的信念状态。
- (3) 结合获得的 s_h 的后验概率分布与观察数据 \tilde{s}_m , 分别计算不同反馈条件 a_h 下的回报函数。
- (4) 基于转移函数和观察函数决定最优策略, 选择回报较大的 a_h , 推动系统演化, 即改变系统中人内部状态 s_h 的概率分布, 以改善人机系统的性能。

2.4 模型的实例化说明

本小节以日常生活中常见的人-计算机系统对上一节中所描述的模型进行较为形象的说明。人-计算机系统中主要涉及到的有操控计算机的人、键盘鼠标和电脑屏幕以及一些外接的传感器设备。假设在这个人机系统中，我们要识别的人的认知偏差或异常定为人的情绪，要推断出人的情绪是轻松欢快的还是激动愤怒的。对上一节中描述的基于 POMDP 的人机系统模型中的变量进行实例化：

$$s_h = \{\text{轻松欢快, 激动愤怒}\}$$

$$es_h = \{\text{笑脸, 愤怒脸}\}$$

$$i_m = \{\text{正常按键, 按键加重}\}$$

$$s_m = \{\text{输出正确文字, 输出错误文字}\}, \text{ 给 } s_m \text{ 附相应的回报值。}$$

$$a_h = \{\text{对识别出愤怒的人进行调节, 不进行调节}\}, \text{ 给 } a_h \text{ 附相应的代价值。}$$

相应的还有几个表示各个传感器获得的观测变量不做过多描述，将上述对人机系统模型中的各个变量的说明在系统模型的动态贝叶斯网络描述中表示出来：

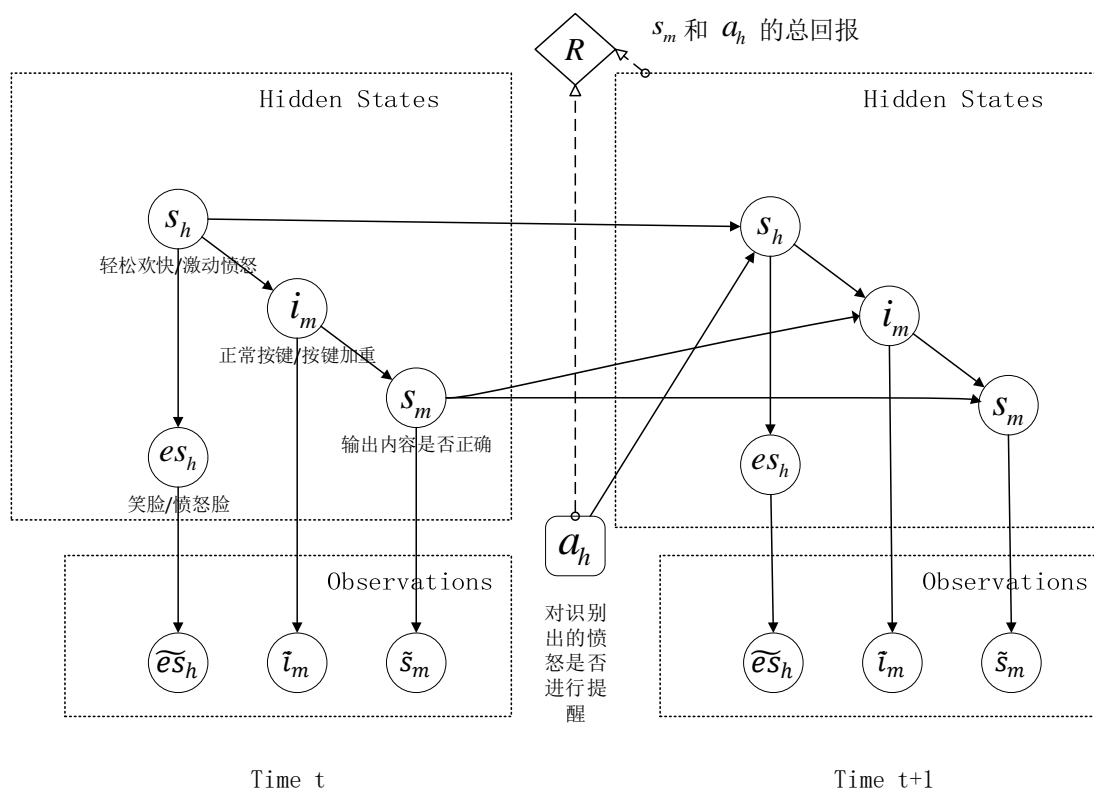


图3 基于 POMDP 的人-计算机系统模型的动态贝叶斯网络表示

在这个具体的人机系统中我们的方法具体化为：首先利用监测人脸得到的图像与笑脸数据集、愤怒脸数据集进行匹配，根据相应的匹配程度给出对于系统中人的情绪的先验概率。假设我们能获得屏幕输出正确与否的信息，根据系统中人在面对屏幕输出正确与否时对按键操纵的轻重程度来更新人的情绪的先验概率，得到情绪后验概率。根据总体回报最大化的原理采取相应的动作 a_h ，以推动系统向更好的方向演化。

2.5 本章小结

本章先介绍了本课题提出的方法所基于的基本方法，包括 POMDP 的基本内容和贝叶斯定理的相关概念。主要介绍了基于 POMDP 的人机系统建模过程，并将其表示为动态贝叶斯网络的形式，以更加直观地说明模型。在第 4 小节中以人-计算机系统为例，更加形象地辅助说明了该模型的合理性。

第3章 基于驾驶系统的仿真实验

3.1 实验设置

驾驶系统是比较常见、并且容易由于驾驶员的内部状态异常而引发事故的人机系统，因此在本课题中我们基于驾驶系统来做仿真实验，以验证本文提出的 POMDP 在识别人机系统中人的内部异常方面的有效性。对仿真实验设置做一定的假设：驾驶员的目标是保持汽车单线行驶，即保持在单车道上；为了简单起见，假设影响驾驶员的正常行为的内部状态仅考虑人可能会犯困。仿真实验的具体人机系统模型如下：

司机的内部状态有两种：

$$S_h = \{\text{清醒}, \text{犯困}\}$$

其中如果司机的内部状态是犯困的，即是司机的内部状态产生了异常。根据这两种内部状态，司机的外部状态或行为表现为眼睛可能是睁着或闭着的。因此：

$$ES_h = \{\text{睁眼}, \text{闭眼}\}$$

司机在驾驶车辆以保持车单线行驶的过程中对车施加控制输入，主要表现为对车的方向盘的控制：

$$I_m = \{\text{向左}, \text{直行}, \text{向右}\}$$

机器状态（或人机系统的输出）即汽车所处车道的位置，由于 POMDP 只能处理离散状态，因此我们将汽车所处车道的水平位置离散化：

$$S_m = \{-2, -1, 0, +1, +2, \text{偏离}\}$$

其中-2 表示在车道最左边，0 表示在中间，+2 在最右边，偏离即表示已经偏离了目标车道。框架中的行动集合即对人的反馈，定义为一个提醒信号，本课题假设提醒总是有效的，即若判断出司机犯困的概率后若选择提醒，则提醒之后司机的内部状态会从原来某一概率的犯困转为清醒。

$$A_h = \{\text{提醒}, \text{不提醒}\}$$

通过传感器来检测人的外部状态、人对机器的控制输入和机器状态或系统输出以获得观察数据，因此观察信息为：

$$O_{esh} = \{\text{睁眼, 闭眼}\}$$

$$O_{im} = \{\text{向左, 向前, 向右}\}$$

$$O_{sm} = \{-2, -1, 0, +1, +2, \text{偏离}\}$$

如转移概率方程(11)所示，转移概率取决于人体内部状态模型、人类行为模型和机器动态模型。在实际问题的应用中，最好是能够从数据中学习得到相关的模型，但是囿于实验设备有限、且为了方便起见，该仿真实验中我们根据自身对驾驶系统的经验或其他先验知识手动选择仿真中的概率，得到人体内部状态模型和人类行为模型如下图所示：

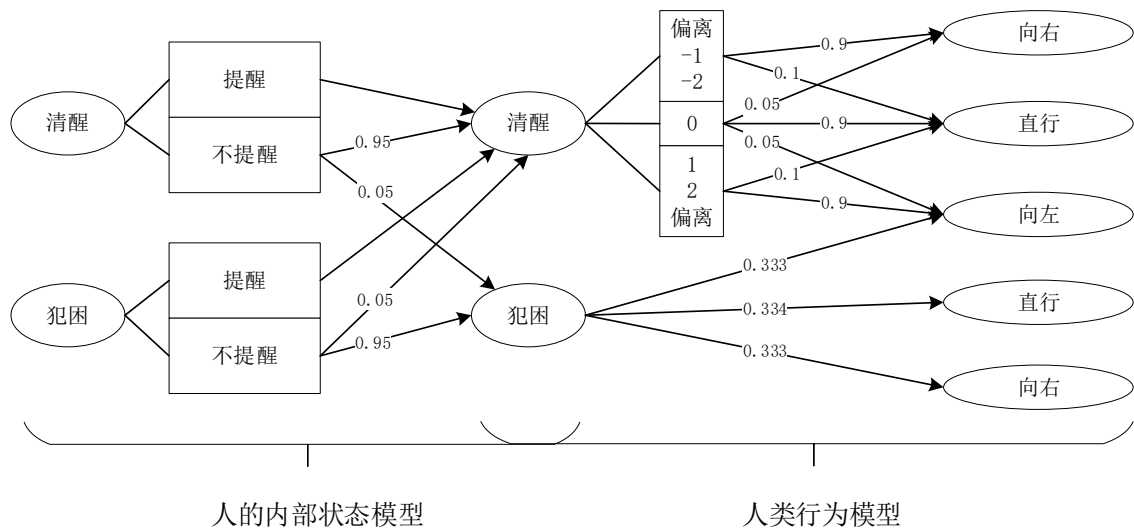


图 4 人的内部状态模型和人类行为模型框图

根据上述人的内部状态模型和人类行为模型，我们可以从人的内部状态的先验概率分布推导出它的后验概率分布，推断出人的内部状态发生异常的概率。其中得到先验概率的分布需要能够观测到想要识别的人的内部异常有其对应的外部状态表现，根据经验或专家知识或学习的过程来确定某一异常的先验概率分布。

根据系统的安全性和效率要求，我们定义总的回报函数包括对系统输出即车辆位置 s_m 的回报以及对反馈动作 a_h 的代价，总的回报函数如下：

$$R(s_m, a_h) = R_1(s_m) + R_2(a_h)$$

其中对系统输出即车辆位置的回报 $R_1(s_m)$ 如下：

s_m	-2	-1	0	1	2	偏离
$R_1(s_m)$	5	10	20	10	5	0

$R_2(a_h)$ 是对 POMDP 框架中的反馈动作的惩罚，本实验中的动作 a_h 是控制器对人的反馈。 $R_2(a_h)$ 如下：

a_h	提醒	不提醒
$R_2(a_h)$	-5	0

基于在满足系统性能要求的前提下使得累积回报最大的考虑，可以在每个采样的时间步上选择执行能够使总的回报最大的动作，使得系统朝着更好的方向演化。当控制器推测出系统中的人以某个一定概率发生异常时，如果对人进行反馈使得系统的输出达到另一状态时获得的回报与执行了这个反馈动作付出的代价之总和并不比不对人进行反馈时导致的系统的输出状态所能获得的回报更大的话，则控制器不会对人进行反馈。

3.2 仿真结果

仿真部分从两个角度验证本文提出的方法的有效性。一方面，对能够获得到的已经完成演化的系统数据进行分析处理以判断人的内部状态的概率分布，验证本文提出的方法是否能有效地识别出人机系统演化过程中人的内部状态异常，即先不将识别出的系统中人的异常概率分布用于对系统的改变。另一个角度的验证实验则将对识别出一定概率的异常进行反馈动作的选择，以期改善系统的演化进程。在本课题的仿真实验中，对于每一个时间步上获得的可量测数据，基于转移函数和观察函数、利用贝叶斯公式对信念状态进行更新，获得人的内部状态的后验概率分布，只推断人的内部异常概率分布而不施加反馈动作的仿真结果如下图 5 所示：

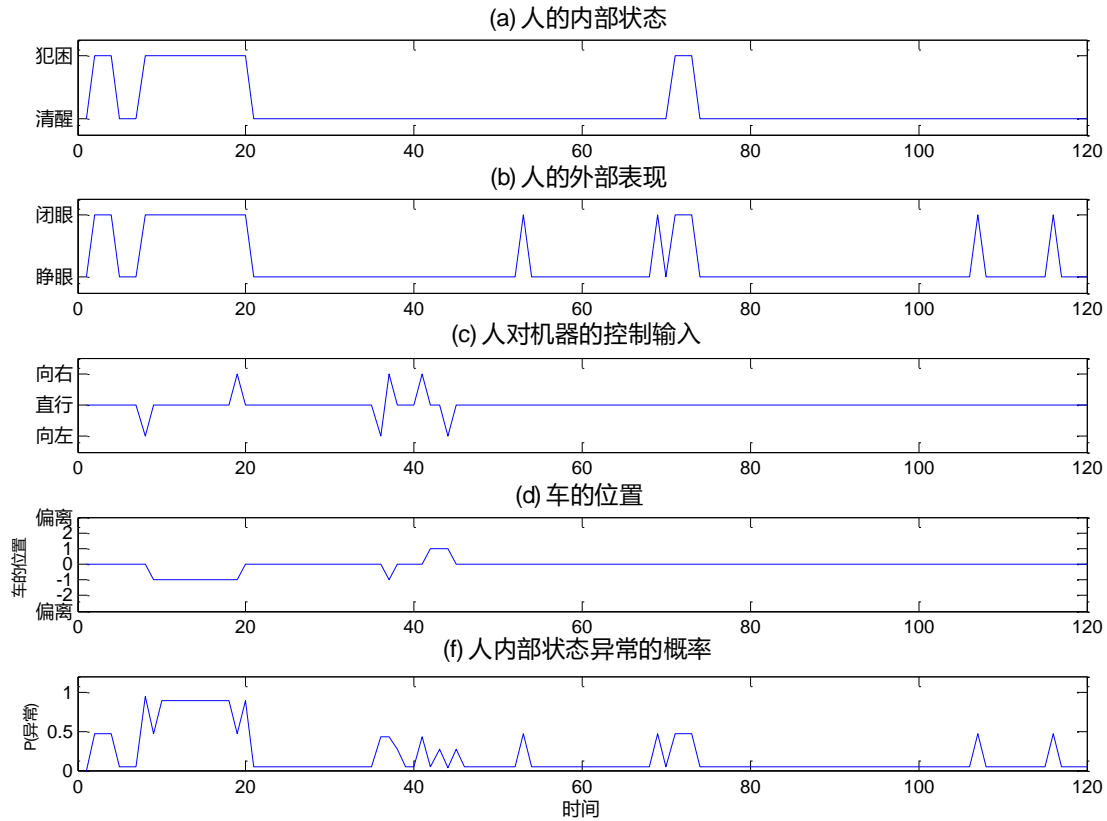


图 5 演化完成的驾驶系统中人驾驶员的内部状态异常检测仿真结果

图 5(a)是驾驶系统演化的一段时间内驾驶员实际的隐藏的内部状态，包括正常的清醒和异常的犯困；图 5(b)是传感器得到的与驾驶员的实际内部状态相对应的外部状态表现，即睁眼或闭眼；图 5(c)是驾驶员对车施加的控制输入，控制汽车的方向盘来改变车在车道上的水平位置；图 5(d)是相应的一段时间内车的水平位置，是系统的输出，也是系统的控制目标；图 5(e)是根据上面几个变量之间的相互关系以及观察数据求得的驾驶员的内部状态的概率分布，图中体现的是驾驶员内部状态异常的概率。如图 5(e)和(a)、(d)所示，当图 5(a)所示驾驶员内部状态发生异常、且图 5(d)所示系统的控制性能受到了影响时，如图 5(e)所示我们可以看到在这些时间点上 $P(\text{异常})$ 的值比较大，这说明系统能够比较好地识别出影响驾驶系统性能的驾驶员的内部状态异常情况。

另一方面，加入判断出驾驶员的内部状态概率分布之后的根据使得累积回报最大的最优策略执行相应的动作，以改变系统的演化进程，验证本文提出的方法在动态演化的系统中是否有效。仿真结果如下图所示(横坐标均为时间 t):

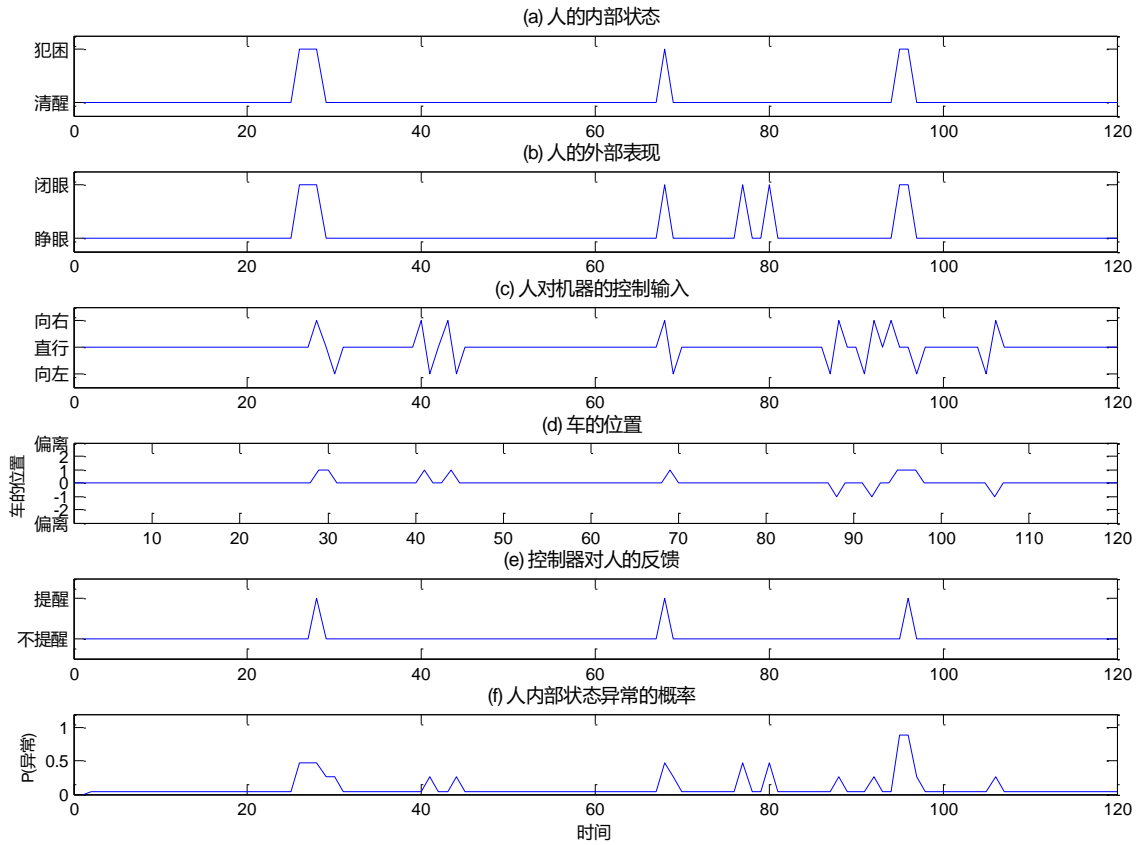


图 6 动态演化的人机系统中人的内部状态异常检测仿真结果

跟上一个不同角度的仿真实验相比，这个实验是在每一个时间步上推断驾驶员的内部状态，并利用 POMDP 的最优策略决定是否对人做出反馈，以改变系统演化进程的仿真实验。如图 6(e)所示，系统在推断出人的内部状态异常的概率较大且对人进行反馈能够得到更大回报的时间点上对人做出了反馈，即能够在关键的地方对驾驶员做出反馈，相比于直接根据传感器观察到的驾驶员的外部状态表现做出提醒的策略减少了较多次的误提醒。

根据上述仿真实验结果我们可以看出控制器能够对基于 POMDP 的人机系统中的人进行较好的内部状态检测，改善系统的性能。

本文提出的方法主要分两部分，第一部分是对人机系统中人的异常的概率的推断，即用贝叶斯定理从先验推后验的方法代替 POMDP 的信念状态的更新过程；第二部分是 POMDP 的根据达到状态获得的回报与采取动作付出的代价的综合回报做出决策。因此对状态的回报和动作的代价的赋值偏好会直接影响决策过

程，接下来用仿真实验来证明这一点。为了清楚地显示实验的效果，我们给予动作代价较大幅度的变化，将提醒这一动作的代价从原来的 5 提升为 10，实验效果如下图 7 所示。

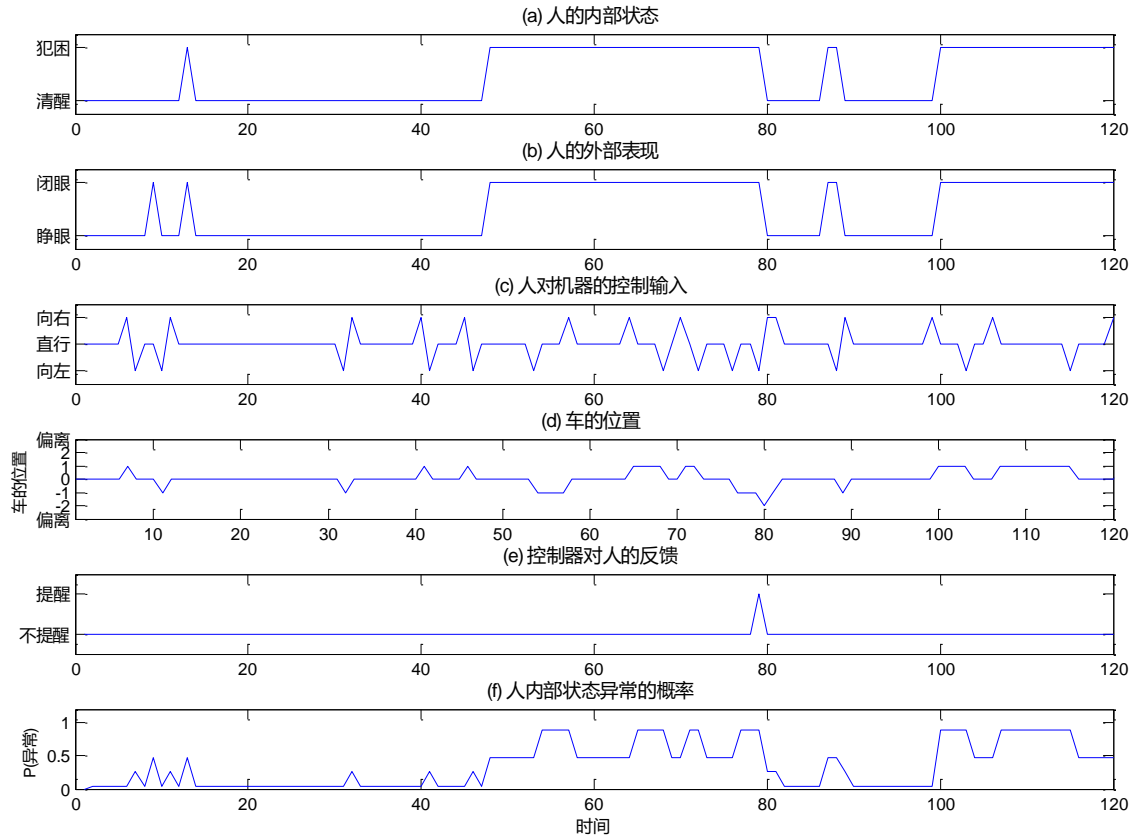


图 7 动作代价过大时的人机系统中人的内部状态异常检测仿真结果

如图 7 所示，当动作代价扩大到不合理的范围时，控制器只在人机系统的输出已经非常偏离正常值时对人进行反馈，对人机系统的性能有较大的影响。因此，从图 6 和图 7 的对比中我们可以看出，设置合适的动作代价与状态回报是决策成功的关键一环。

第4章 总结与展望

4.1 毕设工作总结

由于人和机器有各自所擅长以及弱势的领域，人与机器合作去完成一个任务往往能比单独由人或者机器去完成有更加良好的效果，人机系统的角色越来越丰富。人机系统是由人和机器构成并依赖于人机之间相互作用而完成一定功能的系统，人机系统的性能取决于人和机器两部分，在本文中我们假定机器是自动化的、且不考虑机器会出问题，只考虑人的行为对系统性能的影响。传统的对认知偏差的研究很少是放在人机系统背景下的，本文从人的认知偏差对人机系统的影响入手对认知偏差进行研究。

当人机系统处于特定的目标、特定的环境下，人可能会比较容易产生某种内部异常或认知偏差，导致错误的决策，影响人机系统的性能。

本文的研究内容是识别人机系统中人的认知偏差，本文将认知偏差从心理学上的定义扩展到包含生理、心理两部分的内部状态异常。由于人机系统中的人产生的内部异常一般会影响到人机系统的输出，因此我们从对人机系统的输出数据的分析中去识别人员的内部状态，判断人是否产生了异常。论文的主要工作如下：

- 1) 综述论文研究背景及意义，介绍对人机系统中的人以及对人和机器协同控制的研究现状。
- 2) 本文借鉴了在不确定性条件下做决策的 POMDP 方法，基于 POMDP 框架对人机系统进行建模，构造了一种从人机系统的输出数据的特点中去推断系统中人的内部状态是否异常的方法。
- 3) 基于常见的人驾驶汽车场景下的人机系统设置不同角度的仿真实验，验证本文提出的方法的有效性。

4.2 未来展望

经过查阅文献和分析总结，本文提出了从人机系统的输出数据的特点中去推断系统中人的内部状态是否异常的方法，方法的有效性在目标简单的驾驶系统中

也得到了仿真验证。但是存在以下几个方面的局限性：

第一，要求我们对系统有比较高的了解程度。信念状态的更新过程依赖于对人的行为如何影响系统输出的了解。

第二，所研究的人的内部异常状态种类有限。如果要识别的人产生的内部状态异常可能有多种情况，工作量又会增大，方法没能很好地适应状态空间中状态数量的增加。

另外根据本文提出的方法所使用的数据以及所能够解决的问题类型，提出以下两点展望。

第一，本文现在的方法主要使用两部分数据，对人的观察得到的人的外部状态信息和系统的输出数据。这说明该方法适用于能够监测到人的人机系统，而无法识别监测不到人的人机系统中人的异常，因此希望之后的工作能够只从系统的输出数据去判断人产生异常的概率。

第二，本文提出的方法识别的是当前时刻人产生异常的概率，之后的工作方向应该是能够预测下一时刻人产生异常的概率。

参 考 文 献

- [1] 李玲, 解洪成, 陈圻. 复杂人机系统人机协作模型的探讨[J]. 人类工效学, 2007, 13(4):36-38.
- [2] Haselton M G, Nettle D & Andrews P W. The evolution of cognitive bias[M]. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology*. John Wiley & Sons Inc. 2005, pp. 724–746.
- [3] Gigerenzer G, Goldstein D G. Reasoning the fast and frugal way: Models of bounded rationality[J]. *Psychological Review*. 1996, 103 (4): 650–669.
- [4] Tversky A & Kahneman D. Judgement under uncertainty: heuristics and biases[J]. *Science*. 1974, **185** (4157): 1124–1131.
- [5] Cialdini R B, Martin S J, & Goldstein N J. Small behavioral science-informed changes can produce large policyrelevant effects[J]. *Behavioral Science & Policy*, 2015,1, 21-27.
- [6] 周鹏生. 认知偏差的产生及其与认知闭合需要的关系[J]. *心理学研究*, 2017, 10(5):11-18.
- [7] Tahboub K A. Intelligent human-machine interaction based on dynamic bayesian networks probabilistic intention recognition[J]. *Journal of Intelligent and Robotic Systems*, 2006, 45(1):31-52.
- [8] Pentland A and Liu A. Modeling and prediction of human behavior[J]. *Neural Computation*, 1999, 11:229–242.
- [9] Takano W, Matsushita A, Iwao K, et al. Recognition of human driving behaviors based on stochastic symbolization of time series signal[C]. *IEEE/RSJ International Conference on Intelligent Robots & Systems*. IEEE, 2008.
- [10] Chipalkatty R, Daep H, Egerstedt M, et al. Human-in-the-loop: MPC for shared control of a quadruped rescue robot[C]. *IEEE/RSJ International Conference on Intelligent Robots & Systems*. IEEE, 2011.
- [11] Munir S, Stankovic J A, Liang C J M, et al. Cyber physical system challenges for human-in-the-loop control[C]. Presented as part of the 8th International Workshop

- on Feedback Computing, 2013.
- [12] Anderson S J, Peters S C, Pilutti T E, et al. An optimal-control-based framework for trajectory planning, threat assessment, and semi-autonomous control of passenger vehicles in hazard avoidance scenarios[J]. *International Journal of Vehicle Autonomous Systems*, 2010, 8(2/3/4):190.
- [13] Sutton RS, Barto AG. Reinforcement learning: an introduction[J]. *IEEE Trans Neural Netw.* 1998, 9(5):1054.
- [14] Colman AM. Cooperation, psychological game theory, and limitations of rationality in social interaction[J]. *Behav Brain Sci.* 2003, 26(2):139–153.
- [15] Broz F, Nourbakhsh I, Simmons R. Planning for human–robot interaction in socially situated tasks[J]. *International Journal of Social Robotics*, 2013, 5(2):193–214.
- [16] Kaelbling LP, Littman ML, Cassandra AR. Planning for human-robot interaction in partially observable stochastic domains[J]. *Artif Intell.* 1998, 101(1–2):99–134.
- [17] Bosch Lt, Oostdijk N, Ruiters JPd. Durational aspects of turn-taking in spontaneous face-to-face and telephone dialogues[C]. In: *Proc of text, speech, and dialogue, 7th international conference (TSD)*. 2004, pp 563–570.
- [18] 王丽. 浅析贝叶斯公式及其在概率推理中的应用[J]. *科技创新导报*, 2010, (24):136.
- [19] 章宗长, 陈小平. 杂合启发式在线 POMDP 规划[J]. *软件学报*, 2013, 24(7):1589-1600.
- [20] Paul Dagum, Adam Galper, Eric Horvitz. Dynamic Network Models for Forecasting[C]. *Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence*. 1992, AUAI Press: 41–48.

附录：仿真部分主要程序

```

%%      演化完成的驾驶系统中人驾驶员的内部状态异常检测
clear all;clc;
t = 1:1:120;
ESh = zeros(1,120); %眼睛的睁/闭
Im = zeros(1,120); %人对车的控制输入 (-1 向左、0 直行、1 向右)
Sm = zeros(1,120); %车的状态/位置 (-2、-1、0、+1、+2)
P = zeros(1,120);
P_A1 = 0;P_A2 = 0;% A1=Sleepy  A2=Awake
%% 随机生成一段 Sh(人的内部状态)
Sh =[-1,1,1,1,-1,-1,-1,1,1,1,1,1,1,1,1,1,1,1,1,1,...
      -1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,...
      -1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,...
      -1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,-1,-1,-1,-1,-1,-1,-1,...
      -1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,...
      -1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1];%Human internal state 自然演化的
      系统中人的内部状态
%% 相应生成一段 Ah1(人眼睛的状态)
for i = 1:1:120
    content = [0 1]; prob = [0.02 0.98];
    tt = randsrc(1,1,[content;prob]);
    if tt == 1
        ESh(i) = Sh(i);
    else
        ESh(i) = Sh(i)*(-1);
    end
end
%% 生成相应的 Ah2(人对车的控制输入)和 Sm(车的位置)      通过此时刻车的状态加上人
      的操作得到下一时刻的车位置
for i = 2:1:120
    j = i+1;
    if Sh(i) == -1          %Awake
        if Sm(i) < 0
            content = [0 1]; prob = [0.1 0.9];
            Im(i) = randsrc(1,1,[content;prob]);
        elseif Sm(i) > 0
            content = [-1 0]; prob = [0.9 0.1];
            Im(i) = randsrc(1,1,[content;prob]);
        else
            content = [-1 0 1]; prob = [0.01 0.98 0.01];
            Im(i) = randsrc(1,1,[content;prob]);
        end
    end
end

```

```

        end
    else
        content = [-1 0 1]; prob = [0.1 0.8 0.1];
        Im(i) = randsrc(1,1,[content;prob]);
    end
    Sm(j) = Sm(i) + Im(i);
end

for i = 2:1:120
    if ESh(i) == -1 % 眼睛睁着
        P_A1 = 0.1;
        P_A2 = 0.9;
    else % 眼睛闭着
        P_A1 = 0.7;
        P_A2 = 0.3;
    end
    if (Sm(i-1) == -2) || (Sm(i-1) == -1) % 车位置偏左
        switch Im(i)
            case 0 % 控制输入为“保持直行”
                P_B2 = 0.1; % B 表示“在 Sm=-2 时, 采取 Im=0” P_B2= P(B|A2)
                P_B1 = 0.334;
            case 1
                P_B2 = 0.9;
                P_B1 = 0.333;
            case -1
                P_B2 = 0;
                P_B1 = 0.333;
        end
    elseif Sm(i-1) == 0 % 车在中间
        switch Im(i)
            case 0
                P_B2 = 0.9;
                P_B1 = 0.334;
            case 1
                P_B2 = 0.05;
                P_B1 = 0.333;
            case -1
                P_B2 = 0.05;
                P_B1 = 0.333;
        end
    else % 车位置偏右
        switch Im(i)
            case 0
                P_B2 = 0.1;

```

```

        P_B1 = 0.334;
    case 1
        P_B2 = 0;
        P_B1 = 0.333;
    case -1
        P_B2 = 0.9;
        P_B1 = 0.333;
    end
end
a = P_A1*P_B1+P_A2*P_B2;
b = P_A1*P_B1;
P(i) = b/a;
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%      动态演化的驾驶系统中人驾驶员的内部状态异常检测
clear all;clc;

t = 1:1:120; t1 = 1:1:121;
Sh = - ones(1,120); %人的内部状态
ESh = - ones(1,120); %眼睛的睁/闭(0/1)
Im = zeros(1,120); %人对车的控制输入 (-1 向左、0 向右、1 直行)
Sm = zeros(1,120); %车的状态/位置 (-2、-1、0、+1、+2)
Warning = zeros(1,120);
P = zeros(1,120);
Reward1 = 0;Reward0 = 0;

for i = 2:1:120
    j = i+1;

    content = [0 1]; prob = [0.02 0.98];
    tt = randsrc(1,1,[content;prob]);
    if tt == 1
        ESh(i) = Sh(i);
    else
        ESh(i) = Sh(i)*(-1);
    end

    if ESh(i) == -1 % 眼睛睁着
        P_A1 = 0.1;
        P_A2 = 0.9;
    else % 眼睛闭着
        P_A1 = 0.7;
        P_A2 = 0.3;
    end
end

```

```

if (Sm(i) == -2) || (Sm(i) == -1)    % 车位置偏左
    switch Im(i)
        case 0                % 控制输入为“保持直行”
            P_B2 = 0.1; % B 表示“在 Sm=-2 时, 采取 Im=0” P_B2 = P(B|A2)
            P_B1 = 0.334;
        case 1
            P_B2 = 0.9;
            P_B1 = 0.333;
        case -1
            P_B2 = 0;
            P_B1 = 0.333;
    end
elseif Sm(i) == 0                % 车在中间
    switch Im(i)
        case 0
            P_B2 = 0.9;
            P_B1 = 0.334;
        case 1
            P_B2 = 0.05;
            P_B1 = 0.333;
        case -1
            P_B2 = 0.05;
            P_B1 = 0.333;
    end
else                                % 车位置偏右
    switch Im(i)
        case 0
            P_B2 = 0.1;
            P_B1 = 0.334;
        case 1
            P_B2 = 0;
            P_B1 = 0.333;
        case -1
            P_B2 = 0.9;
            P_B1 = 0.333;
    end
end

a = P_A1*P_B1+P_A2*P_B2;
b = P_A1*P_B1;
P(i) = b/a;

x = 1;

```

```
while(x>=1)
    Sh(j) = -1;

    if Sm(i) < 0
        content = [0 1]; prob = [0.1 0.9];
        Im(i) = randsrc(1,1,[content;prob]);
    elseif Sm(i) > 0
        content = [-1 0]; prob = [0.9 0.1];
        Im(i) = randsrc(1,1,[content;prob]);
    else
        content = [-1 0 1]; prob = [0.01 0.98 0.01];
        Im(i) = randsrc(1,1,[content;prob]);
    end
    Sm(j) = Sm(i) + Im(i);

    switch Sm(j)
        case -2
            Rsm = 5;
        case -1
            Rsm = 10;
        case 0
            Rsm = 20;
        case 1
            Rsm = 10;
        case 2
            Rsm = 5;
        otherwise
            Rsm = 0;
    end
    Reward1 = Rsm - 5;

    Sh(j) = Sh(i);

    if Sh(j) == -1;
        if Sm(i) < 0
            content = [0 1]; prob = [0.1 0.9];
            Im(i) = randsrc(1,1,[content;prob]);
        elseif Sm(i) > 0
            content = [-1 0]; prob = [0.9 0.1];
            Im(i) = randsrc(1,1,[content;prob]);
        else
            content = [-1 0 1]; prob = [0.05 0.9 0.05];
            Im(i) = randsrc(1,1,[content;prob]);
        end
    end
end
```



```
else
    content = [-1 0 1]; prob = [0.1 0.8 0.1];
    Im(i) = randsrc(1,1,[content;prob]);
end

Sm(j) = Sm(i) + Im(i);

switch Sm(j)
    case -2
        Rsm = 5;
    case -1
        Rsm = 10;
    case 0
        Rsm = 20;
    case 1
        Rsm = 10;
    case 2
        Rsm = 5;
    otherwise
        Rsm = 0;
end
Reward0 = Rsm;
x = x-1;
end

if (Reward1 > Reward0) && (P(i)>0.4)
    Warning(i) = 1;
    Sh(j) = -1;
else
    Warning(i) = 0;
    if Sh(i) == -1    %醒着的
        content = [-1 1]; prob = [0.95 0.05];
        Sh(j) = randsrc(1,1,[content; prob]);    %下一时刻大概率清醒
    else    %睡着的
        content = [-1 1]; prob = [0.05 0.95];
        Sh(j) = randsrc(1,1,[content; prob]);    %下一时刻大概率睡着
    end
end
end
end
```

致 谢

大学四年果然比想象中更快地溜走了，作为人生中不可或缺的四年，一路走来充满了喜怒哀乐。有失望的泪水、奋斗的汗水，也有收获的喜悦。志同道合的同学伙伴，更有循循善诱的师长。在这个接近四年尾声的时刻，对曾经帮助过我的老师同学和默默付出的父母表达深深的感谢。

首先，应当感谢的是母校浙江工业大学。是她给我提供了一个优越的学习环境，使我可以安心沉浸于学习和研究当中，而不会收到外界的干扰。

然后，非常需要感谢的是我的指导教师赵云波教授。赵老师严谨的治学态度深深吸引着我。从最初的选题、开题答辩，到后面的中期答辩、毕业论文的撰写，赵老师给予了我非常大的帮助。每次遇到瓶颈觉得无从下手的时候，与赵老师的交流总能让我豁然开朗，帮助我明确了前进的方面。没有赵老师的指导，在整个毕业设计工作和毕业论文的撰写的过程中应该会走很多弯路。另外也非常非常感谢身边的同学朋友和实验室的师兄师姐，在整个做毕业设计的过程中给我提供了很及时的帮助，给予了我很多次的柳暗花明。

最后，我想特别感谢我的父母。是你们给我提供了良好且来之不易的学习条件，使我不用担心除学习以外的事情。每次与你们的交流都会成为我继续乐观开朗、继续前进的动力。我们怀揣着同样的心、以不同的方式朝着同样的目标前进——为了更好的自己、为了更好的家，再次感谢我的父母。

大学四年，收获的不只是知识、不仅是情谊，更是面对未知挑战的勇气。尽管我可能还是不那么机灵、不那么擅长做很多事情，尽管未来的人生还是会遇到很多的困难和失败，坚持下去，找到属于自己的道路！在未来可能会遇到更多挑战的日子里，不抛弃、不放弃，相信自己会取得成功！