(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115461767 A
(43)申请公布日 2022.12.09

(21)申请号 202180029984.4

(22)申请日 2021.01.15

(30)优先权数据
2020-078883 2020.04.28 JP
2020-116497 2020.07.06 JP

(85)PCT国际申请进入国家阶段日
2022.10.21

(86)PCT国际申请的申请数据
PCT/JP2021/001234 2021.01.15

(87)PCT国际申请的公布数据
WO2021/220556 JA 2021.11.04

(71)申请人 索尼集团公司
地址 日本东京

(72)发明人 滝裕子 高塚进 铁川弘树

(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
专利代理师 郑宗玉

(51)Int.Cl.
G06N 20/00(2006.01)

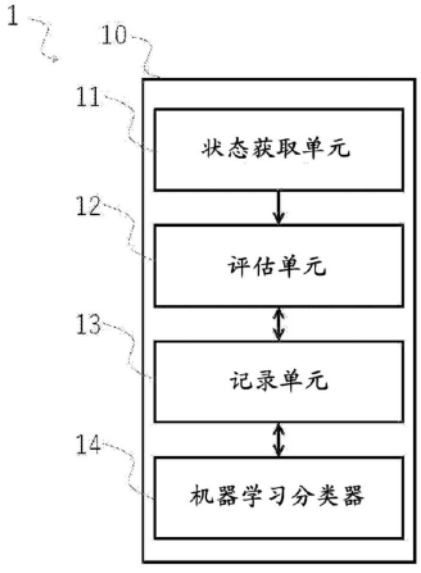
权利要求书2页 说明书21页 附图13页

(54)发明名称

机器学习系统、机器学习设备和机器学习方法

(57)摘要

提供了一种机器学习系统、机器学习设备和机器学习方法，用于对人的行为和人周围的环境之间的相关性进行机器学习，以便促进从人的行为到目标行为的转变。本技术提供了一种机器学习系统，至少设置有：状态获取单元，至少获取关于人的行为的状态信息；评估单元，评估关于当获取状态信息在人周围的环境的环境信息和状态信息时，以便获得价值函数；以及机器学习分类器，对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便鼓励行为到目标行为的转变。



CN 115461767 A

1.一种机器学习系统,至少包括:

状态获取单元,所述状态获取单元被配置为至少获取关于人的行为的状态信息;

评估单元,所述评估单元被配置为通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数;以及

机器学习分类器,所述机器学习分类器对价值函数执行强化学习,并且选择当价值函数最高时的环境信息,以便促进所述人来参与目标行为。

2.根据权利要求1所述的机器学习系统,

其中,所述评估单元被配置为:

基于状态信息和关于目标行为的目标状态信息之间的差异来计算奖励,以及

基于奖励、环境信息和状态信息来计算价值函数。

3.根据权利要求1所述的机器学习系统,

其中,所述系统保持包括多个目标行为信息的目标状态相关信息。

4.根据权利要求3所述的机器学习系统,

其中,目标状态相关信息包括特定于时间的目标状态信息和/或特定于阶段的目标状态信息。

5.根据权利要求1所述的机器学习系统,

其中,环境信息包括关于气味、照明、温度、湿度、视频或声音的信息。

6.根据权利要求1所述的机器学习系统,

还包括气味控制单元,

其中,所述气味控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制所生成的气味。

7.根据权利要求6所述的机器学习系统,

还包括芳香化单元,

其中,所述芳香化单元被配置为基于由所述机器学习分类器选择的环境信息使物品带有气味,以及

所述机器学习分类器基于环境信息来确定所述气味控制单元和所述芳香化单元中的哪一个将生成气味。

8.根据权利要求1所述的机器学习系统,

还包括照明控制单元,

其中,所述照明控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制要发射的光。

9.根据权利要求1所述的机器学习系统,

还包括空气调节单元,

其中,所述空气调节单元被配置为基于由所述机器学习分类器选择的环境信息来控制温度和/或湿度。

10.根据权利要求1所述的机器学习系统,

还包括视频控制单元,

其中,所述视频控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制要显示的视频。

11.根据权利要求1所述的机器学习系统，

还包括声音控制单元，

其中，所述声音控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制要播放的声音。

12.根据权利要求1所述的机器学习系统，

其中，价值函数被划分为多个价值组，以及

所述机器学习分类器使用由所述多个价值组中的每个价值组保持的价值函数。

13.根据权利要求1所述的机器学习系统，还包括：

多个状态获取单元；以及

实现难度等级计算单元，

其中，所述实现难度等级计算单元被配置为基于由所述多个状态获取单元中的每个状态获取单元获取的状态信息来计算目标行为的实现难度等级。

14.根据权利要求13所述的机器学习系统，

其中，实现难度等级包括实现率，所述实现率表示促进到目标行为的程度。

15.根据权利要求13所述的机器学习系统，

其中，实现难度等级包括标准实现时间，所述标准实现时间表示促进到目标行为的标准时间。

16.根据权利要求13所述的机器学习系统，

其中，实现难度等级包括平均环境项目数量，所述平均环境项目数量表示当促进到目标行为时的环境信息中的项目的平均数量。

17.一种机器学习设备，至少包括：

状态获取单元，所述状态获取单元被配置为至少获取关于人的行为的状态信息；

评估单元，所述评估单元被配置为通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及

机器学习分类器，所述机器学习分类器对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

18.一种机器学习方法，至少包括：

至少获取关于人的行为的状态信息；

通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及

对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

机器学习系统、机器学习设备和机器学习方法

技术领域

[0001] 本技术涉及机器学习系统、机器学习设备和机器学习方法。

背景技术

[0002] 已经使用了如下技术，其中计算机对关于人的行为的信息执行机器学习，以便促进(prompt)人来参与目标行为。

[0003] 例如，专利文献1公开了“用于使用计算机网络向消费者提供促销信息以诱导消费并且促进(promote)销售的促销系统”。专利文献1描述了将基于在提供促销信息之后的消费者的行为而执行的机器学习。

[0004] 引用列表

[0005] 专利文献

[0006] 专利文献1:日本专利申请特许公开第2019 028899号

发明内容

[0007] 本发明要解决的问题

[0008] 传统上，向人提供信息的直接表示(诸如专利文献1中所述的促销信息)，以便促进该人来参与目标行为。

[0009] 然而，当人改变他们的行为时，他们可能只是在不加思考的情况下响应于他们周围的环境的变化而改变该行为。

[0010] 因此，本技术主要旨在提供机器学习系统、机器学习设备和机器学习方法，其分别通过对人的行为和他们周围的环境之间的相关性进行机器学习来促进这样的人参与目标行为。

[0011] 问题的解决方案

[0012] 本技术提供一种机器学习系统，至少包括:状态获取单元，所述状态获取单元至少获取关于人的行为的状态信息;评估单元，所述评估单元通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数;以及机器学习分类器，所述机器学习分类器对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

[0013] 所述评估单元可以被配置为:基于状态信息和关于目标行为的目标状态信息之间的差异来计算奖励，以及基于奖励、环境信息和状态信息来计算价值函数。

[0014] 所述机器学习系统可以保持包括多个目标行为信息的目标状态相关信息。

[0015] 目标状态相关信息可以包括特定于时间的目标状态信息和/或特定于阶段的目标状态信息。

[0016] 环境信息可以包括关于气味、照明、温度、湿度、视频或声音的信息。

[0017] 机器学习系统还可以包括气味控制单元，并且所述气味控制单元可以被配置为基于由所述机器学习分类器选择的环境信息来控制所生成的气味。

[0018] 所述机器学习系统还可以包括芳香化单元，并且所述芳香化单元可以被配置为基于由所述机器学习分类器选择的环境信息使物品带有气味，以及

[0019] 所述机器学习分类器可以基于环境信息来确定所述气味控制单元和所述芳香化单元中的哪一个将生成气味。

[0020] 所述机器学习系统还可以包括照明控制单元，并且所述照明控制单元可以被配置为基于由所述机器学习分类器选择的环境信息来控制要发射的光。

[0021] 所述机器学习系统还可以包括空气调节单元，并且所述空气调节单元可以被配置为基于由所述机器学习分类器选择的环境信息来控制温度和/或湿度。

[0022] 所述机器学习系统还可以包括视频控制单元，并且所述视频控制单元可以被配置为基于由所述机器学习分类器选择的环境信息来控制要显示的视频。

[0023] 所述机器学习系统还可以包括声音控制单元，并且所述声音控制单元可以被配置为基于由所述机器学习分类器选择的环境信息来控制要播放的声音。

[0024] 价值函数可以被划分为多个价值组，以及

[0025] 所述机器学习分类器可以使用由所述多个价值组中的每个价值组保持的价值函数。

[0026] 所述机器学习系统还可以包括多个状态获取单元；以及实现难度等级计算单元，并且所述实现难度等级计算单元可以被配置为基于由所述多个状态获取单元中的每个状态获取单元获取的状态信息来计算目标行为的实现难度等级。

[0027] 实现难度等级可以包括实现率，所述实现率表示促进到目标行为的程度。

[0028] 实现难度等级可以包括标准实现时间，所述标准实现时间表示促进到目标行为的标准时间。

[0029] 实现难度等级可以包括平均环境项目数量，所述平均环境项目数量表示当促进到目标行为时的环境信息中的项目的平均数量。

[0030] 此外，本技术还提供一种机器学习设备，至少包括：状态获取单元，所述状态获取单元被配置为至少获取关于人的行为的状态信息；评估单元，所述评估单元被配置为通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及机器学习分类器，所述机器学习分类器对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

[0031] 此外，本技术还提供了一种机器学习方法，至少包括：至少获取关于人的行为的状态信息；通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

附图说明

[0032] 图1是示出根据本技术的一个实施例的机器学习系统1的配置的框图。

[0033] 图2是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0034] 图3是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0035] 图4是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0036] 图5是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0037] 图6是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0038] 图7是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0039] 图8是示出根据本技术的一个实施例的行为变化的一个示例的概念图。

[0040] 图9是示出根据本技术的一个实施例的目标状态相关信息的一个示例的数据库。

[0041] 图10是示出根据本技术的一个实施例的机器学习设备10的硬件配置的框图。

[0042] 图11是示出根据本技术的一个实施例的机器学习系统1的配置的框图。

[0043] 图12是示出根据本技术的一个实施例的气味控制单元23的配置的框图。

[0044] 图13是示出根据本技术的一个实施例的机器学习系统1的配置的框图。

[0045] 图14是示出根据本技术的一个实施例的机器学习系统1的过程的流程图。

[0046] 图15是示出根据本技术的一个实施例的照明控制单元24的配置的框图。

[0047] 图16是示出根据本技术的一个实施例的空气调节单元25的配置的框图。

[0048] 图17是示出根据本技术的一个实施例的视频控制单元26的配置的框图。

[0049] 图18是示出根据本技术的一个实施例的声音控制单元27的配置的框图。

[0050] 图19示出由根据本技术的一个实施例的机器学习设备10使用的数据库的一个示例。

[0051] 图20是示出根据本技术的一个实施例的机器学习设备10的示例性过程的流程图。

[0052] 图21是示出根据本技术的一个实施例的机器学习系统1的配置的框图。

[0053] 图22是示出由根据本技术的一个实施例的实现难度等级计算单元54计算的实现难度等级的图。

[0054] 图23是示出由根据本技术的一个实施例的实现难度等级计算单元54计算的实现难度等级的图。

[0055] 图24是示出根据本技术的一个实施例的机器学习方法的过程的流程图。

具体实施方式

[0056] 下文中将描述用于实现本技术的优选实施例。下面描述的实施例分别示出了本技术的代表性实施例的示例，并且本技术的范围不限于此。此外，每个图是示意图，并且不一定精确地图示。

[0057] 将按以下顺序来描述本技术。

[0058] 1、本技术的第一实施例(机器学习系统的示例1)

[0059] (1)概要

[0060] (2)评估单元

[0061] (3)机器学习分类器

[0062] (4)行为变化流程

[0063] (5)目标状态信息

[0064] (6)硬件配置

[0065] 2、本技术的第二实施例(机器学习系统的示例2)

[0066] (1)概要

[0067] (2)气味控制单元

[0068] (3)芳香化单元

[0069] (4)照明控制单元

[0070] (5)空气调节单元

[0071] (6)视频控制单元

[0072] (7)声音控制单元

[0073] 3、本技术的第三实施例(机器学习系统的示例3)

[0074] 4、本技术的第四实施例(机器学习系统的示例4)

[0075] 5、本技术的第五实施例(机器学习系统的示例5)

[0076] (1)概要

[0077] (2)实现难度等级

[0078] 6、本技术的第六实施例(机器学习方法)

[0079] [1.本技术的第一实施例(机器学习系统的示例1)]

[0080] [(1)概要]

[0081] 根据本技术的一个实施例的机器学习系统可以通过对关于人的行为的信息和关于这样的人周围的环境的信息进行评估并且执行机器学习来获取行为和环境之间的相关性。因此,可以通过控制环境来促进人参与目标行为。

[0082] 将参考图1来描述根据本技术的一个实施例的机器学习系统的配置。图1是示出根据本技术的一个实施例的机器学习系统1的配置的框图。

[0083] 如图1中所示,机器学习系统1可以包括例如机器学习设备10。机器学习设备10可以包括例如状态获取单元11、评估单元12、记录单元13和机器学习分类器14。

[0084] 状态获取单元11至少获取关于人的行为的状态信息。因此,机器学习系统1可以识别人如何响应于环境的变化而改变其行为。

[0085] 状态信息的示例包括用于在用户访问网站时识别该用户的“小甜饼”(Cookie)、电子商务(e-commerce)购买历史、例如通过GPS(全球定位系统)获取的位置信息、聊天对话历史以及通过感测技术获取的其他信息。

[0086] 此外,状态信息可以包括关于人所在地区的天气或温度的信息。在这样的情况下,机器学习系统1可以学习特定于人所在地区的天气或温度的独特行为变化。

[0087] 评估单元12通过评估关于当获取状态信息时在人周围的环境的环境信息和状态信息来获得价值函数。因此,获取了环境信息和状态信息之间的相关性。稍后将描述具体的评估处理。

[0088] 环境信息的示例包括关于气味、照明、温度、湿度、视频或声音的信息。稍后将描述环境信息的具体示例。

[0089] 记录单元13记录例如状态信息和环境信息。此外,机器学习系统1获取状态信息,但是使用所记录的而尚未获取的环境信息。

[0090] 机器学习分类器14对价值函数执行强化学习,并且选择当价值函数最高时的环境信息,以便促进人来参与目标行为。因此,获取了行为和环境之间的相关性。

[0091] 机器学习方法没有特别限制,但是例如,可以使用强化学习。强化学习是这样的机器学习训练方法,其中软件能够感知并且解释当前状态(本技术中的状态信息),以及确定代理器(agent)应该参与的行为(本技术中的环境信息的变化)。代理器(本技术中的机器学习分类器14)可以利用通过试错法进行的强化学习来确定当价值最高时的行为。

[0092] 用于实现强化学习的传统方法的示例包括蒙特卡洛学习、动态规划、状态行为奖励状态行为(SARSA)和Q学习。本技术将参考Q学习进行描述，该Q学习是强化学习的示例。此外，可以在本技术中使用除了Q学习以外的强化学习算法。

[0093] 此外，尽管未示出，但是机器学习设备10可以配备有控制每个组件的控制单元、经由网络建立通信的通信接口等。

[0094] [(2)评估单元]

[0095] 如上所述，评估单元12通过评估关于当获取状态信息时在人周围的环境的环境信息和状态信息来获得价值函数。

[0096] 尽管评估单元12的实现方式没有特别限制，但是评估单元12可以配备有奖励计算单元(未示出)和价值计算单元(未示出)。

[0097] 关于人的行为的状态信息可以根据环境信息的变化而变化。当机器学习系统1改变了环境信息时，奖励计算单元基于关于目标行为的目标状态信息和状态信息之间的差异来计算奖励R。奖励R的值越大，表示目标状态信息和状态信息之间的差异越小。也就是说，奖励R的值越大，人的行为越接近于目标行为。

[0098] 奖励R可以通过例如以下等式(1)来表示，该等式(1)使用与目标状态信息对应的得分Pt和与状态信息对应的得分Pm。

[0099] [数学式1]

[0100] $$R = 1 - \frac{|Pt - Pm|}{Pt} \quad \cdots (1)$$

[0101] 下面将描述具体示例。“使用电子商务网站购买产品A”被设定为目标行为。然后，当人由于环境信息变化而参与目标行为时，给予5分。

[0102] 此外，当人由于环境信息变化而参与接近于目标行为的行为(例如“访问包括产品A的网站”)时，给予2分。

[0103] 将数字拟合到等式(1)，与目标状态信息对应的得分Pt为“5”。当人由于环境信息变化而参与目标行为时，与状态信息对应的得分Pm也为“5”。此时，奖励R为“1”。

[0104] 当人由于环境信息变化而参与接近于目标行为的行为时，与状态信息对应的得分Pm为“2”。此时，奖励R为“0.4”。

[0105] 当人由于环境信息变化而参与除这两个行为以外的行为时，与状态信息对应的得分Pm为“0”。此时，奖励R也为“0”。

[0106] 换言之，奖励R的值随着由于环境信息的变化而导致的人的行为更接近于目标行为而增加。奖励计算单元计算当环境信息变化时的最高奖励R。

[0107] 价值计算单元基于奖励R、环境信息和状态信息来计算价值函数Q。价值计算单元基于当环境信息以最高奖励R变化时的状态信息来计算价值函数Q。例如，当对在时间t的状态信息s_t进行环境信息的变化a_t时的价值函数由$Q(s_t, a_t)$来表示。

[0108] 例如，价值函数Q可以由记录单元13记录。更具体地，记录单元13可以针对每个状态信息或环境信息将价值函数Q记录在表上。

[0109] [(3)机器学习分类器]

[0110] 如上所述，机器学习分类器14对价值函数Q执行强化学习，并且选择当价值函数Q最高时的环境信息。

[0111] 下面将描述该强化学习。机器学习分类器14自动通过试错法进行学习，使得价值函数Q变得最高。随着价值函数Q越高，人的行为越接近于目标行为。通过执行强化学习以使得价值函数Q变得最高，机器学习分类器14可以促进人针对当前行为参与目标行为。

[0112] 在选择当价值函数Q最高时的环境信息时，机器学习分类器14更新价值函数Q。例如，当对在时间t的状态信息s_t进行环境信息的变化a_t并且转换到在时间t+1的状态信息s_{t+1}时，利用以下等式(2)来更新价值函数$Q(s_t,a_t)$。

[0113] [数学式2]

[0114] $$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[R_{t+1} + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t)\right] \quad \cdots (2)$$

[0115] α表示学习系数。学习系数α的值落在0<α≤1的范围内；最常用的值为大约0.1。

[0116] R_{t+1}表示通过状态信息的转换获得的奖励。

[0117] γ表示贴现率(discount rate)。贴现率γ的值落在0<γ≤1的范围内；最常用的值为大约0.9到0.99。

[0118] $\max Q(s_{t+1},a)$表示未来的理想的价值函数。$\max Q(s_{t+1},a)$是当在时间t+1的状态s_{t+1}中选择具有最高价值函数Q的行为a时的价值函数。将价值函数$\max Q(s_{t+1},a)$乘以贴现率γ。

[0119] 机器学习分类器14使用上述等式(2)不断更新价值函数Q，并且选择当价值函数Q最高时的环境信息。因此，机器学习分类器14可以选择可以促进人来参与目标行为的环境信息。

[0120] [(4)行为变化流程]

[0121] 假设在达到目标行为之前经历了多个行为变化。将参考图2对此进行描述。图2是示出根据本实施例的行为变化的一个示例的概念图。如图2中所示，“于在线流媒体平台处观看视频”被设定为目标行为。然后，设定用于促进人来参与目标行为的多个行为变化流程。机器学习分类器14可以利用重复的强化学习来配置该行为变化流程。机器学习分类器14可以通过遵循该流程来促进人来参与目标行为。

[0122] 此外，根据人的行为与目标行为的接近程度，将人的行为分类为多个级别。例如，第一级别行为可以是最接近于目标行为的行为。第二级别行为可以是次接近于目标行为的行为。与第一级别行为有关的价值函数Q高于与第二级别行为有关的价值函数Q。

[0123] 在该示例中，第一级别行为包括“去洗手间”和“坐在沙发上”。第二级别行为包括“孩子上床睡觉”、“回家”、“离开桌子”和“喝酒”。然后，通过连接多个行为中的每个行为来配置行为变化流程。例如，该个人的行为变化特征表明，当“孩子上床睡觉”发生时，他们倾向于参与“去洗手间”的行为。

[0124] 其他示例性的行为变化流程如图3到图7中所示。图3到图7各自是示出根据本实施例的行为变化的一个示例的概念图。图3示出了用于促进人来参与目标行为(即“使用电子商务网站”)的行为变化流的一个示例。图4示出了用于促进人来参与目标行为(即“访问SNS(社交网络服务)账号”)的行为变化流的一个示例。图5示出了用于促进人来参与目标行为(即“进行大额购买”)的行为变化流程的一个示例。图6示出了用于促进人来参与目标行为(即“喝啤酒”)的行为变化流程的一个示例。图7示出了用于促进人来参与目标行为(即“去睡觉”)的行为变化流程的一个示例。如上所示，可以根据目标行为来配置各种行为变化流程。

[0125] 此外,即使目标行为相同,用于促进人来参与目标行为的行为变化流程也可以根据个人而不同。将参考图8对此进行描述。图8是示出根据本实施例的行为变化的一个示例的概念图。图8A示出了某个个人的行为变化的一个示例。如图8A中所示,“购买衣服”被设定为目标行为。

[0126] 另一方面,图8B示出了除图8A中所示的个体之外的个体的行为变化的一个示例。如图8B中所示,目标行为与图8A相同,但是用于促进人来参与目标行为的行为变化流程不同。

[0127] [(5)目标状态信息]

[0128] 根据本技术的一个实施例的机器学习设备10可以保持关于一个目标行为的目标状态信息,但是可以保持关于多个目标行为的多个目标状态信息。例如,可以按时间和/或按阶段设定多个目标行为中的一部分或全部。

[0129] 例如,可以按时间设定多个目标行为中的一部分或全部。更具体地,例如,多个目标行为中的一部分或全部可以被分类为一天中的第一时间段(time zone)(例如,从12:00AM到6:00AM)中的目标行为、第二时间段(例如,从7:00AM到7:00PM)中的目标行为和第三时间段(例如,从8:00PM到11:00PM)中的目标行为。

[0130] 例如,第一时间段(例如,从12:00AM到6:00AM)中的目标行为可以是“去睡觉”。例如,第二时间段(例如,从7:00AM到7:00PM)中的目标行为可以是“吃食物S”。例如,第三时间段(例如,从8:00PM到11:00PM)中的目标行为可以是“喝饮料T”。

[0131] 按时间设定多个目标行为中的一部分或全部,由此可以根据例如时间段灵活地设定目标行为。例如,机器学习设备10可以促进人来参与目标行为,即“在3:00PM吃食物S”。

[0132] 替选地,可以例如按阶段设定多个目标行为中的一部分或全部。更具体地,例如,多个目标行为中的一部分或全部可以被分类为第一阶段中的目标行为和第二阶段中的目标行为。例如,第一阶段中的目标行为可以是“去商店U”。例如,第二阶段中的目标行为可以是“吃食物S”。

[0133] 按阶段设定目标行为中的一部分或全部,由此可以设定具有一系列流程的多个目标行为。例如,机器学习设备10可以促进人来参与目标行为,即“去商店U并且吃食物S”。

[0134] 替选地,可以例如按时间并且按阶段设定多个目标行为中的一部分或全部。更具体地,可以在第三时间段内设定第一阶段和第二阶段中的目标行为。

[0135] 按时间并且按阶段设定目标行为中的一部分或全部,由此可以根据例如时间段灵活地设定具有一系列流程的多个目标行为。例如,机器学习设备10可以促进人来参与目标行为,即“早上去商店U并且吃食物S”。

[0136] 为了实现上述方案,根据本技术的一个实施例的机器学习设备10可以保持包括多个目标状态信息的目标状态相关信息。

[0137] 将参考图9描述目标状态相关信息。图9是示出根据本实施例的目标状态相关信息的一个示例的数据库。如图9中所示,时间段与目标行为相关联。第一时间段(从12:00AM到6:00AM)中的目标行为是“去睡觉”。第二时间段(从7:00AM到7:00PM)中的目标行为是“吃食物S”。例如,第三时间段(从8:00PM到11:00PM)中的目标行为是“喝饮料T”。此外,这同样适用于对于每个阶段设定的目标状态相关信息。

[0138] 目标状态相关信息可以记录在例如包括在机器学习设备10中的记录单元13中。此

外,目标状态相关信息可以由除机器学习设备10以外的计算机设备保持。例如,目标状态相关信息可以保持在云上的服务器中。在这样的情况下,机器学习设备10可以经由信息通信网络从服务器接收目标状态相关信息。

[0139] [(6)硬件配置]

[0140] 将参考图10描述机器学习设备10的硬件配置。图10是示出根据本实施例的机器学习设备10的硬件配置的框图。如图10中所示,机器学习设备10可以包括CPU 101、储存器102、随机存取存储器(RAM)103和通信接口104作为组件。各个组件例如通过作为数据传输路径的总线来连接。

[0141] CPU 101例如由微计算机来实现,并且控制机器学习设备10的每个组件。CPU 101例如可以用作评估单元12或机器学习分类器14。机器学习分类器14可以通过例如程序来实现。该程序可以通过被CPU 101读取来工作。

[0142] 储存器102存储诸如由CPU 101使用的程序和操作参数之类的控制数据。储存器102可以使用例如硬盘驱动器(HDD)或固态驱动器(SSD)来实现。储存器102例如可以用作记录单元13。

[0143] RAM 103临时存储例如由CPU 101执行的程序。

[0144] 通信接口104具有使用诸如Wi-Fi、蓝牙(注册商标)或长期演进(LTE)之类的通信协议经由信息通信网络建立通信的功能。

[0145] 实现机器学习分类器14等的程序可以存储在除机器学习系统1以外的计算机设备或计算机系统中。在这种情况下,机器学习系统1可以采用提供程序的功能的云服务。云服务的示例包括软件即服务(SaaS)、基础设施即服务(IaaS)和平台即服务(PaaS)。

[0146] 此外,可以使用各种非瞬态计算机可读介质来存储程序并且将该程序提供给计算机。非瞬态计算机可读介质包括各种有形存储介质。非瞬态计算机可读介质的示例包括磁记录介质(例如软盘、磁带或硬盘驱动器)、磁光记录介质(例如磁光盘)、紧凑盘只读存储器(CD-ROM)、CD-R、CD-R/W和半导体存储器(例如掩模ROM、可编程ROM(PROM)、可擦除PROM(EPROM)、闪存ROM或随机存取存储器(RAM))。此外,程序可以通过各种瞬态计算机可读介质提供给计算机。瞬态计算机可读介质的示例包括电信号、光信号和电磁波。瞬态计算机可读介质可以经由有线通信路径(例如电线和光纤)或无线通信路径向计算机提供程序。

[0147] [2、本技术的第二实施例(机器学习系统的示例2)]

[0148] [(1)概要]

[0149] 根据本技术的一个实施例的机器学习系统1可以包括用于控制人周围的环境的环境控制设备。环境控制设备基于由机器学习分类器14选择的环境信息来控制人周围的环境。因此,机器学习系统1可以促进人来参与目标行为。

[0150] 例如,机器学习系统1可以通过促进人参与目标行为来促进销售。机器学习系统1可以控制商店的内部或外部环境,以便使客户购买产品。此外,采用机器学习系统1的商店不限于线下商店;它可以用于电子商务网站,即在线购物中心。替选地,机器学习系统1还用于下载或流式传输的网站或内容,并且可以促进对这样的网站或内容的访问。

[0151] 替选地,例如,机器学习系统1可以通过促进人参与目标行为来改善人的日常习惯。具体地,机器学习系统1可以控制环境,以便促使人戒烟或戒酒。此外,机器学习系统1可以帮助克服例如购物依赖性、睡眠不足、缺乏锻炼等。

[0152] 替选地，机器学习系统1可以通过促进人参与目标行为来使该人在选举中进行投票。

[0153] 替选地，机器学习系统1可以通过促进人参与目标行为来提高公众健康意识或道德意识。具体地，机器学习系统1可以用来提高废物管理意识，改善对乱扔垃圾、拥挤乘车和排队的态度，遵循建议(诸如咳嗽遮挡和保持双手清洁)，以及提高对欺凌的意识。

[0154] 替选地，例如，机器学习系统1可以通过促进人参与目标行为来提高工作效率。更具体地，机器学习系统1可以用于提高注意力、学习注意力和工作-休息平衡。

[0155] 将参考图11描述根据本实施例的机器学习系统1的配置。图11是示出根据本实施例的机器学习系统1的配置的框图。此外，与第一实施例中的组件类似的组件由类似的附图标记来表示，并且将省略其详细描述。

[0156] 如图11中所示，根据本实施例的机器学习系统1还可以包括环境控制设备20。环境控制设备20和机器学习设备10经由信息通信网络40连接。环境控制设备20基于由机器学习分类器14选择的环境信息来控制环境，诸如气味、照明、温度、湿度、视频或声音。

[0157] 环境控制设备20可以包括例如通信控制单元21、存储器22、气味控制单元23、照明控制单元24、空气调节单元25、视频控制单元26和声音控制单元27。

[0158] 此外，环境控制设备20可以不具有气味控制单元23、照明控制单元24、空气调节单元25、视频控制单元26和声音控制单元27的全部；换句话说，它可以具有这些组件中的至少有一个组件。

[0159] 此外，机器学习系统1可以包括多个环境控制设备20。例如，当机器学习系统1包括两个环境控制设备20时，一个环境控制设备20可以包括气味控制单元23，而另一个环境控制设备20可以包括照明控制单元24。

[0160] 通信控制单元21可以经由信息通信网络40与机器学习设备10进行信息通信。此外，通信控制单元21可以控制每个组件。

[0161] 存储器22可以记录由环境控制设备20使用的信息，例如环境信息。

[0162] 注意，机器学习设备10中设置的机器学习分类器14可以被包括在例如环境控制设备20中，或者可以被包括在其他计算机设备中。

[0163] [(2)气味控制单元]

[0164] 环境控制设备20可以包括例如气味控制单元23，以控制人周围的气味。气味控制单元23基于由机器学习分类器14选择的环境信息来控制所生成的气味。包括气味控制单元23的环境控制装置20可以使用例如芳香扩散器(aroma diffuser)来实现。

[0165] 此外，气味包括可以被人感知为气味的气味，以及不能被人感知为气味但被吸入以对这样的人施加某种作用的气味。例如，吸入的镇静剂或通过吸入而作用于人的身体状况的无味气体(例如氧气或二氧化碳)也被包括在气味中。

[0166] 通过吸入由气味控制单元23控制的并且针对人优化的气味，促进该人不知不觉地参与目标行为。

[0167] 将参考图12描述气味控制单元23的配置。图12是示出根据本实施例的气味控制单元23的配置的框图。如图12中所示，气味控制单元23可以包括例如添加剂盒231、气味控制单元232和气味输出单元233。

[0168] 添加剂盒231是存放气味的添加剂的组件。添加剂盒231可以是可更换的。添加剂

盒231例如可以是:容纳添加剂的容器,诸如圆筒、瓶子或罐;吸附添加剂的材料,诸如纸、无纺布或石头;或与添加剂混合的固体,诸如蜡或肥皂。

[0169] 例如,添加剂可以是包括粉末和凝胶的固体、液体或气体,或者其混合物。例如,添加剂可以是天然提取的香料、从化学合成获得的合成香料或通过混合这些香料而制备的制备香料。替选地,添加剂可以不包含香料。

[0170] 气味控制单元232基于环境信息来控制用于生成气味的添加剂。例如,气味控制单元232可以确定在混合添加剂时每个添加剂的比率。替选地,气味控制单元232可以确定稀释率。根据由机器学习分类器14选择的环境信息来确定比率或稀释率。

[0171] 替选地,气味控制单元233可以控制例如用于输出气味的参数,例如喷射压力和喷射次数。根据由机器学习分类器14选择的环境信息来确定喷射压力或喷射次数。

[0172] 气味输出单元233基于由气味控制单元232确定的信息来输出气味。

[0173] 替选地,设置有气味控制单元23的机器学习系统1可以通过使人吸入特定气味来促进人参与目标行为。例如,机器学习系统1可以生成这样的气味,该气味在生理上诱导人购买特定产品,从而他们通过邮件订购或去商店购买这样的产品。

[0174] 替选地,气味可以与特定内容相关联。因此,机器学习系统1可以在执行强化学习之前使人不知不觉地学习气味和内容之间的关联。

[0175] 例如,机器学习系统1可以使人在观看特定视频时吸入特定气味。例如,视频包括与特定产品有关的广告。因此,机器学习系统1可以使人不知不觉地学习特定气味和特定产品之间的关联。当用户从机器学习系统1吸入特定气味时,它将诱导用户通过邮件订购特定产品或去商店查找特定产品。

[0176] 替选地,气味可以与特定环境相关联。该环境与人实际经历的地方或事物有关。该地方的示例包括商店、公共交通、电影院、剧院和主题公园。因此,机器学习系统1可以在执行强化学习之前使人不知不觉地学习气味和环境之间的关联。

[0177] 例如,机器学习系统1可以使人在他们参观特定商店时吸入特定气味。因此,机器学习系统1可以使人不知不觉地学习特定气味和特定商店之间的关联。当该人在与商店不同的地方从机器学习系统1吸入特定气味时,它将诱导该人去商店或通过邮件订购商店中展示的产品。

[0178] 例如,机器学习系统1可以通过体验(例如喝咖啡)使人吸入从特定产品生成的气味。因此,机器学习系统1可以使人不知不觉地学习特定气味和特定产品之间的关联。当气味控制单元23生成该特定气味时,机器学习系统1将诱导该人通过邮件订购特定产品或去商店查找特定产品。

[0179] [(3)芳香化单元]

[0180] 气味可以粘附在某个物品上。物品的示例可以包括交付给要被促进来参与目标行为的人的衣服、书籍、杂货、促销物品或包装材料。通过吸入粘附在物品上并且针对人进行优化的气味,促进该人不知不觉地参与目标行为。

[0181] 为了实现上述方案,机器学习系统1可以设置有芳香化单元。将参考图13对此进行描述。图13是示出根据本实施例的机器学习系统1的配置的框图。此外,与第一实施例中的组件类似的组件由类似的附图标记来表示,并且将省略其详细描述。如图13中所示,机器学习系统1还包括芳香化单元30。芳香化单元30和机器学习设备10经由信息通信网络40连接。

此外,芳香化单元30可以具有与气味控制单元23类似的配置。

[0182] 将气味控制单元23布置在要被促进来参与目标行为的人周围。另一方面,例如,将芳香化单元30布置在物品被装运的工厂中。芳香化单元30基于由机器学习分类器14选择的环境信息使物品带有气味。

[0183] 机器学习分类器14基于环境信息确定气味控制单元23和芳香化单元30中的哪一个生成气味。

[0184] 将参考图14描述此时的机器学习系统1的过程。图14是示出根据本实施例的机器学习系统1的过程的流程图。

[0185] 如图14中所示,机器学习分类器14在机器学习的初始阶段对气味和行为之间的相关性执行强化学习(步骤S11)。机器学习分类器14选择环境信息(步骤S12),并且确定气味控制单元23生成气味(步骤S13)。气味控制单元23控制人周围的气味(步骤S14)。重复步骤S11至S14,直到气味和行为之间的相关性被充分训练(步骤S15:否)。

[0186] 在气味和行为之间的相关性已被充分训练的阶段(步骤S15:是),机器学习分类器14确定芳香化单元30生成气味(步骤S16),并且芳香化单元30使物品带有气味(步骤S17)。

[0187] 因此,机器学习系统1可以更灵活地控制人周围的气味。例如,在机器学习的初始阶段(例如,从学习的开始起大约1个月到3个月),在布置在人周围的气味控制单元23在短时间段内改变气味的同时,机器学习分类器14高效地执行气味和行为之间的相关性的强化学习。机器学习分类器14确定针对目标行为进行优化的气味。

[0188] 此后,可以通过例如交付带有最适合于目标行为的气味的物品来持续地促进该目标行为。在芳香化单元30在很长时间段内改变气味的同时,机器学习分类器14持续进行气味和行为之间的相关性的强化学习。

[0189] [(4)照明控制单元]

[0190] 描述返回到图11。环境控制装置20可以包括例如照明控制单元24,以控制人周围的照明。照明控制单元24基于由机器学习分类器14选择的环境信息来控制要发射的光。包括照明控制单元24的环境控制设备20可以使用例如灯泡(包括所谓的智能灯泡)来实现。

[0191] 通过在视觉上识别针对人进行优化并且由照明控制单元24照射的光,促进该人不知不觉地参与目标行为。

[0192] 将参考图15描述照明控制单元24的配置。图15是示出根据本实施例的照明控制单元24的配置的框图。如图15中所示,照明控制单元24可以包括例如光控制单元241和光输出单元242。

[0193] 光控制单元241控制要输出的光的表现。更具体地,例如,光控制单元241可以确定光的色温和亮度。根据由机器学习分类器14选择的环境信息来确定色温或亮度。例如,色温可以被确定为3500K到3900K,并且亮度可以被确定为3000lm到4000lm。此外,在如上所述确定范围的情况下,光控制单元241可以随机确定落在该范围内的值。机器学习设备10可以在重复强化学习时缩小该范围。此外,这同样适用于下文所述的其他组件。

[0194] 光输出单元242基于由光控制单元241确定的信息来输出光。

[0195] [(5)空气调节单元]

[0196] 环境控制设备20可以包括例如空气调节单元25,以控制人周围的空气。空气调节单元25基于由机器学习分类器14选择的环境信息来控制温度和/或湿度。包括空气调节单

元25的环境控制设备20可以使用例如空调来实现。

[0197] 通过由空气调节单元25针对人进行优化并且控制的温度和/或湿度，促进该人不知不觉地参与目标行为。

[0198] 将参考图16描述空气调节单元25的配置。图16是示出根据本实施例的空气调节单元25的配置的框图。如图16中所示，空气调节单元25可以包括例如空气控制单元251和空气输出单元252。

[0199] 空气控制单元251可以确定空气的温度和/或湿度。根据由机器学习分类器14选择的环境信息来确定温度和/或湿度。例如，温度可以被确定为25.5℃至27.5℃，并且湿度可以被确定为45％至50％。

[0200] 空气输出单元252基于由空气控制单元251确定的信息来输出空气。

[0201] [(6)视频控制单元]

[0202] 环境控制设备20可以包括例如视频控制单元26，以控制要向人显示的视频。视频控制单元26基于由机器学习分类器14选择的环境信息来控制要显示的视频。包括视频控制单元26的环境控制设备20可以使用例如电视、便携式游戏机、PC、平板计算机、智能电话、头戴式显示器(HMD)、可穿戴设备或汽车导航系统来实现。

[0203] 注意，视频包括运动图像和静止图像两者。此外，视频可以包括声音。

[0204] 通过在视觉上识别由视频控制单元26显示并且针对人进行优化的视频，促进该人不知不觉地参与目标行为。

[0205] 将参考图17描述视频控制单元26的配置。图17是示出根据本实施例的视频控制单元26的配置的框图。如图17中所示，视频控制单元26可以包括例如视频选择单元261和视频显示单元262。

[0206] 视频选择单元261选择要输出的视频。选择处理没有特别限制，但是例如，视频选择单元261可以使用例如视频文件被记录的地址或广告横幅的代码来确定。根据由机器学习分类器14选择的环境信息来确定地址或代码。此外，视频选择单元261可以合成或编辑多个视频文件。此外，视频选择单元261可以例如调整视频的色温或亮度。

[0207] 此外，视频文件可以记录在视频控制单元26中，或者可以记录在视频控制单元26外部。

[0208] 视频显示单元262基于由视频选择单元261确定的信息来输出视频。

[0209] [(7)声音控制单元]

[0210] 环境控制设备20可以包括例如声音控制单元27，以控制针对人播放的声音。声音控制单元27基于由机器学习分类器14选择的环境信息来控制要播放的声音。包括声音控制单元27的环境控制设备20可以使用例如扬声器(包括所谓的智能扬声器和具有流传输功能的扬声器)、平板设备、智能电话、耳机、可穿戴设备或汽车立体声来实现。

[0211] 通过聆听由声音控制单元27播放并且针对人进行优化的声音，促进该人不知不觉地参与目标行为。

[0212] 将参考图18描述声音控制单元27的配置。图18是示出根据本实施例的声音控制单元27的配置的框图。如图18中所示，声音控制单元27可以包括例如声音选择单元271和声音输出单元272。

[0213] 声音选择单元271选择要播放的声音。选择处理没有特别限制，但是例如，声音选

择单元271可以使用例如音频文件被记录的地址或广告横幅的代码来确定。根据由机器学习分类器14选择的环境信息来确定地址或代码。此外，声音选择单元271可以合成或编辑多个音频文件。此外，声音选择单元271可以调整例如音高和音量。

[0214] 此外，音频文件可以记录在声音控制单元27中，或者可以记录在声音控制单元27外部。

[0215] 声音输出单元272基于由声音选择单元271确定的信息来输出声音。

[0216] [3、本技术的第三实施例(机器学习系统的示例3)]

[0217] 根据本技术的一个实施例的机器学习设备10可以记录每个目标行为的价值函数Q、状态信息s和环境信息的变化a。然后，机器学习设备10可以通过对人的行为和人周围的环境之间的相关性执行强化学习来选择可以促进人参与目标行为的环境信息。

[0218] 此时，可以将具有类似的在行为和环境之间的相关性的多个人放在同一价值组中。例如，可以将当感觉到气味和温度变化时可能被促进参与特定目标行为的多个人放在同一组中。

[0219] 将参考图19对此进行描述。图19示出了由根据本实施例的机器学习设备10使用的数据库的一个示例。如图19中所示，数据库保持每个目标行为的价值函数Q、状态信息s和环境信息的变化a。根据状态信息s和环境信息的变化a之间的相关性的相似度级别，将价值函数Q划分为多个价值组Q1到Q8。人h01到h32分别属于多个价值组Q1到Q8中的每个价值组，并且与每个价值组相关联。指示每个人的特征的属性信息A到C可以与每个人相关联。

[0220] 当已知强化学习的对象所属的价值组时，机器学习设备10可以使用诸如与该价值组有关的价值函数Q之类的信息。例如，通过使用诸如已经经过强化学习的价值函数Q之类的信息，机器学习设备10可以部分地省略强化学习过程，并且减少执行强化学习所花费的时间。

[0221] 下面将描述具体示例。假设某个人的目标行为被设定为“玩视频游戏”。还假设通过强化学习发现，该人在受到气味和温度变化的影响时倾向于玩视频游戏。

[0222] 接下来，目标行为从“玩视频游戏”变为“喝啤酒”。之前的强化学习表明，该人很容易受到气味和温度变化的影响。因此，诸如易受气味和温度变化影响的价值组的价值函数Q以及已获得更高奖励的环境信息的变化a之类的信息可以用作用于促使该人参与新的目标行为的强化学习的初始值。机器学习设备10可以使用已经经历强化学习的信息作为初始值来启动强化学习。

[0223] 此外，诸如已经经历强化学习的价值函数Q之类的信息可以用于对属于同一价值组的其他人的行为的强化学习。参考图19，例如，诸如已经经历对属于价值组Q4的人h13的行为的强化学习的价值函数Q之类的信息可以用于对属于同一价值组Q4的人h14的行为的强化学习。

[0224] 将参考图20描述根据本实施例的机器学习设备10的过程。图20是示出根据本实施例的机器学习设备10的示例性过程的流程图。

[0225] 如图20中所示，机器学习设备10中包括的状态获取单元11获取状态信息(步骤S21)。

[0226] 接下来，机器学习设备10中包括的评估单元12基于状态信息计算奖励和价值函数(步骤S22)。

[0227] 接下来，机器学习设备10中包括的机器学习分类器14更新价值函数(步骤S23)。

[0228] 接下来，为了学习进一步的行为变化，机器学习分类器14选择环境信息(步骤S24)。

[0229] 接下来，机器学习分类器14确定是否满足预定条件(步骤S25)。该确定条件没有特别限制，但是可以通过例如价值函数的更新次数是否超过预定阈值来确定。

[0230] 当满足预定条件时(步骤S25：是)，机器学习分类器14参考数据库并且获取诸如相似组的价值函数Q和已获得更高奖励的环境信息的变化a之类的信息(步骤S26)。该数据库可以包括在机器学习设备10中，或者可以包括在除机器学习设备10以外的计算机设备中。机器学习设备10可以使用已经经历强化学习的信息来执行强化学习。

[0231] 另一方面，当不满足预定条件时(步骤S25：否)，不获取相似组的价值函数。

[0232] 接下来，机器学习分类器14确定是否应终止强化学习(步骤S27)。该确定条件没有特别限制，但是可以通过例如价值函数是否大于预定阈值来确定。

[0233] 当确定不应终止机器学习时(步骤S27：否)，重复步骤S21至S26的过程。

[0234] 当确定应该终止机器学习时(步骤S27：是)，机器学习分类器14选择环境信息(步骤S28)。

[0235] [4、本技术的第四实施例(机器学习系统的示例4)]

[0236] 可以随机设定目标行为。通过在不限于特定目标行为的情况下促进各种目标行为，机器学习分类器14可以执行行为和环境之间的相关性的强化学习。利用强化学习，例如，即使在被认为与行为的关系较低的环境的变化中，机器学习分类器14也可以发现诸如行为的迹象和连续性之类的规律性。

[0237] 还将参考图19对此进行描述。在数据库中登记的多个人中的每个人与指示该人的特征的属性信息相关联。例如，属性A可以是具有以下特征的人：“当照明的色温为3650K、照明的亮度为3000lm并且温度为26.5℃时，倾向于喝啤酒”。例如，属性C可以是具有以下特征的人：“当温度为25℃、湿度为48％并且气味包含添加剂T时，倾向于访问电子商务网站”。

[0238] 为了实现上述该方案，可以获取随机选择的组的价值函数，而不是获取相似组的价值函数(步骤S26)，如图20中所示的流程图中所示。

[0239] [5、本技术的第五实施例(机器学习系统的示例5)]

[0240] [(1)概要]

[0241] 根据本技术的一个实施例的机器学习系统1可以包括多个机器学习设备。将参考图21对此进行描述。图21是示出根据本技术的一个实施例的机器学习系统1的配置的框图。

[0242] 如图21中所示，机器学习系统1可以包括例如多个机器学习设备10a到10d。多个机器学习设备10a至10d中的每个可以分别包括例如状态获取单元11a至11d、评估单元12a至12d、记录单元13a至13d和机器学习分类器14a至14d。此外，环境控制设备(未示出)可以连接到多个机器学习设备10a到10d中的每个。此外，机器学习设备的数量没有特别限制。

[0243] 此外，机器学习系统1可以包括实现难度等级计算设备50。实现难度等级计算设备50可以具有如图10中所示的硬件配置。实现难度等级计算设备50经由例如信息通信网络40连接到多个机器学习设备10a到10d中的每个，并且可以聚合从多个机器学习设备10a到10d中的每个获得的信息，由此获得状态信息和环境信息之间的相关性的趋势。更具体地，实现难度等级计算设备50可以计算目标行为的实现难度等级。当被定义为成功实现了所促进的

目标行为时，实现难度等级表示该实现的难度的程度。将在稍后详细描述实现难度等级。

[0244] 实现难度等级计算设备50可以包括例如信息获取单元51、对象信息记录单元52、行为信息记录单元53和实现难度等级计算单元54。

[0245] 信息获取单元51获取由多个机器学习设备10a到10d中的每个获得的状态信息。信息获取单元51可以使用例如通信接口104来实现。

[0246] 多个机器学习设备10a到10d中的每个可以针对不同的对象。对象信息记录单元52保持关于由多个机器学习设备10a到10d中的每个所针对的对象的信息。例如，该信息包括对象的识别号码、性别或年龄。可以使用例如储存器102来实现对象信息记录单元52。

[0247] 行为信息记录单元53保持关于针对多个机器学习设备10a到10d中的每个设定的目标行为的信息。例如，该信息包括关于目标行为的信息、状态信息和关于状态信息的历史信息。可以使用例如储存器102来实现行为信息记录单元53。

[0248] 实现难度等级计算单元54可以基于由多个状态获取单元11a到11d中的每个获取的状态信息来计算目标行为的实现难度等级。实现难度等级计算单元54可以使用例如CPU 101和程序来实现。

[0249] 此外，实现难度等级计算单元54可以包括在实现难度等级计算设备50中，可以包括在多个机器学习设备10a到10d中的每个中，或者可以包括在多个环境控制设备(未示出)中的每个中。

[0250] 此外，尽管未示出，但是机器学习系统1可以包括多个实现难度等级计算设备。在多个实现难度等级计算设备当中，可以存在用于中继的实现难度等级计算设备，该用于中继的实现难度等级计算设备聚合从多个机器学习设备当中的特定机器学习设备获得的信息。

[0251] [(2)实现难度等级]

[0252] 如上所述，实现难度等级表示在促进目标行为时的难度。例如，通过计算实现难度等级，机器学习系统1可以推导出很可能或不太可能被促进参与目标行为的对象，或者推导出对象很可能或不太可能被促进参与目标行为的环境信息。

[0253] 很可能被促进参与目标行为的对象组被定义为适应性组，并且不太可能被促进参与目标行为的对象组被定义为挑战组。例如，在产品开发或广告宣传中，机器学习系统1可以通过推导出适应性组来推导出产品的目标受众。例如，目标受众包括年龄和性别。通过推导出产品的目标受众，可以更有效地进行产品开发和广告宣传。

[0254] 例如，设置在例如便利店中的销售点(POS)系统可以与当地活动(例如运动会或烟花节)相关联。因此，机器学习系统1可以推导出例如很可能在活动期间被购买的产品和该产品的目标受众。

[0255] 替选地，产品的示例包括热销产品和长销产品。适应性组可用于前者的开发和广告活动，并且挑战组可用于后者的开发和广告活动。

[0256] 此外，目标行为可以被分类为基本目标行为和与基本目标行为相关联的应用目标行为。基本目标行为包括大致按类型分类的行为，例如，“外出”、“吃喝”和“购买”。应用目标行为更具体地表示基本目标行为；例如，“在黑色星期五去特定商店”、“去特定地方”和“参加当地节日”。

[0257] 机器学习系统1首先推导出适应性组，该适应性组很可能被促进参与应用目标行

为。机器学习系统1可以通过推导出与多个应用目标行为中的每个有关的适应性组并且请求关于适应性组的信息来推导出与基本目标行为有关的适应性组。也就是说，机器学习系统1可以获得多个适应性组共同的倾向。因此，例如，可以针对产品开发推导出到目前为止尚未被注意到的新目标受众。

[0258] 此外，如第二实施例中所述，适应性组的推导也可以用于改善日常习惯。

[0259] 实现难度等级可以包括例如实现率r，该实现率r表示促进到目标行为的程度。具有较高实现率r的对象被分类到适应性组中。

[0260] 例如，实现率r可以由以下等式(3)来表示，该等式(3)使用促进到目标行为的状态信息的数量n和包括没有促进到目标行为的状态信息的所有状态信息的数量n_{all}。

[0261] [数学式3]

[0262] $$r = \frac{100n}{n_{all}} \quad \cdots (3)$$

[0263] 实现难度等级可以包括例如标准实现时间s，该标准实现时间s表示促进到目标行为的标准时间。具有较短标准实现时间s的对象被分类到适应性组中。

[0264] 例如，标准实现时间s可以由以下等式(4)来表示，该等式(4)使用实现时间x和平均实现时间p，该实现时间x表示将对象促进到参与目标行为所花费的时间，该平均实现时间p表示将对象促进到参与目标行为所花费的平均时间。平均实现时间p可以通过将实现时间x的总和除以所有状态信息的数量n_{all}来计算。

[0265] [数学4]

[0266] $$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - p)^2} \quad \cdots (4)$$

[0267] 此外，尽管使用标准偏差来计算标准实现时间s，以便不受具有超长实现时间的对象的影响，但是使用平均值而不是标准偏差的平均实现时间p可以包括在实现难度等级中。

[0268] 替选地，实现难度等级可以包括例如平均环境项目数量q，该平均环境项目数量q表示当促进到目标行为时的环境信息中的项目的平均数量。环境信息中的项目的示例包括气味、照明、温度、湿度、视频或声音。具有较少平均环境项目数量q的对象被分类到适应性组中。例如，与不受气味和温度两者影响的对象相比，只受气味影响的对象更有可能被促进参与目标行为。

[0269] 平均环境项目数量q可以由例如以下等式(5)来表示，该等式(5)使用促进到目标行为的状态信息的数量n和当促进到目标行为时的环境信息中的项目的数量e。此外，实现难度等级计算单元54可以计算标准偏差而不是平均值，如等式(4)中那样。

[0270] [数学式5]

[0271] $$q = \frac{1}{n}\sum_{i=1}^{n} e_i \quad \cdots (5)$$

[0272] 可以与平均环境项目数量q的计算一起记录环境信息中的项目的名称。例如，行为信息记录单元53可以记录环境信息中的项目的名称。因此，机器学习系统1可以推导出对于

特定环境信息很容易促进行为变化的适应性组。例如，机器学习系统1可以推导出对于气味很容易促进行为变化的适应性组。

[0273] 此外，实现难度等级可以包括实现率r、标准实现时间s和平均环境项目数量q中的至少一个。然而，例如，与在仅实现率r被包括在实现难度等级中的情况下相比，在实现率r和标准实现时间s两者被包括在实现难度等级中的情况下，可以更容易地推导出适应性组。

[0274] 将参考图22描述实现难度等级。图22是示出由根据本实施例的实现难度等级计算单元54计算的实现难度等级的图。如图22中所示，示出了实现难度中包括的实现率(AR)、标准实现时间(SAT)和平均环境项目数量(NKV)。

[0275] 图22的A示出了当将不是每天锻炼的对象促进到目标行为“进行锻炼”时的实现难度等级。如类别中所示，被促进参与该目标行为的对象是20岁至59岁的男性和女性。所有对象的实现率为30％，标准实现时间为54个小时，并且平均环境项目数量为2个。

[0276] 子类别是类别的细分。在本文中，作为示例，基于性别对类别进行细分。所有男性的实现率为31％，标准实现时间为55个小时，并且平均环境项目数量为1个。另一方面，所有女性的实现率为29％，标准实现时间为53个小时，并且平均环境项目数量为3个。这表明男性对象具有比女性对象高的实现率。换句话说，当关注实现率时，男性对象对应于适应性组。

[0277] 子子类别(sub-subcategory)是子类别的细分。在本文中，作为示例，基于年龄对子类别进行细分。20岁到39岁的所有男性的实现率为34％，标准实现时间为38个小时，并且平均环境项目数量为1个。在该子子类别下的四个组当中，该组具有最高的实现率、最短的标准实现时间和最少的平均环境项目数量。也就是说，该组对应于适应性组。机器学习系统1可以以该方式推导出适应性组。关于与目标行为“进行锻炼”有关的产品和服务的宣传活动可以针对该适应性组。

[0278] 图22的B示出了当将过去5年内从未投票的对象促进到目标行为“去投票”时的实现难度等级。在子子类别下的四个组当中，50岁到79岁的女性对象的组具有最高的实现率、最短的标准实现时间和最少的平均环境项目数量。也就是说，该组对应于适应性组。

[0279] 如上所述，机器学习系统1可以通过计算实现率、标准实现时间或平均环境项目数量来推导出适应性组。例如，机器学习系统1可以针对目标行为“买啤酒”推导出具有实现率≥80％和标准实现时间≤3个小时的适应性组。当啤酒公司推出新产品时，啤酒公司可以针对适应性组进行新产品的广告和宣传活动。

[0280] 此外，例如，机器学习系统1可以针对目标行为“于在线流媒体平台处观看电视节目或视频”推导出具有实现率≥90％和平均环境项目数量≤2个的适应性组。视频流媒体服务提供商可以针对适应性组进行订阅其服务的广告和宣传活动。即使在适应性群体成为订户后，服务提供商也可以针对该适应性组进行鼓励续订的广告和宣传活动。

[0281] 该实现难度等级的值可以随着反复促进到目标行为而改变。将参考图23对此进行描述。图23是示出由根据本实施例的实现难度等级计算单元54计算的实现难度等级的图。

[0282] 图23的A示出了当促进到目标行为“购买产品S至少两次”时的实现难度等级。此外，在该情况下，当购买产品S两次时，实现率为100％，而当购买产品S一次时，实现率为50％。

[0283] 在图23的A中，基于实现难度等级将对象分类为多个组。例如，第一组G1具有396个

对象，其中实现率为86％，标准实现时间为67个小时，并且平均环境项目数量为2个。此外，第二组G2具有283个对象，其中实现率为62％，标准实现时间为120个小时，并且平均环境项目数量为3个。在四个组当中，第一组G1和第二组G2具有较高的实现率、较短的标准实现时间和较少的平均环境项目数量。即，第一组G1和第二组G2对应于适应性组。产品S的销售方可以促进适应性组参与与产品S有关的目标行为。

[0284] 随后，机器学习系统1促进第一组G1和第二组G2参与与产品S有关的目标行为，并且促进与挑战组对应的第三组G3和第四组G4参与与作为另一产品的产品T有关的目标行为。

[0285] 图23的B示出了当促进到目标行为“购买产品T至少两次”时的实现难度等级。如图23的B中所示，第三组G3的实现率为68％，标准实现时间为258个小时，并且平均环境项目数量为3个。在四个组当中，第三组G3具有较高的完成率、较短的标准完成时间和较少的平均环境项目数量。即，第三组G3对应于适应性组。

[0286] 即，例如，通过将第一组G1和第二组G2设定为产品S的目标受众，并且将第三组G3设定为产品T的目标受众，可以提高产品的销售或宣传活动的效率。

[0287] [6、本技术的第六实施例(机器学习方法)]

[0288] 根据本技术的一个实施例的机器学习方法是用于使用计算机设备训练人的行为和人周围的环境之间的相关性的机器学习方法。将参考图24描述根据本实施例的机器学习方法。图24是示出根据本实施例的机器学习方法的过程的流程图。如图24中所示，根据本实施例的机器学习方法至少包括：至少获取关于人的行为的状态信息(步骤S1)；通过评估关于当获取状态信息时在人周围的环境的环境信息和状态信息来获得价值函数(步骤S2)；以及对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进人来参与目标行为(步骤S3)。

[0289] 根据本实施例的机器学习方法可以使用根据第一实施例到第四实施例的技术。因此，将省略描述。

[0290] 此外，本说明书中描述的有益效果仅为示例，而不是限制，并且可以预期其他效果。

[0291] 此外，本技术还可以具有以下配置。

[0292] [1]一种机器学习系统，至少包括：

[0293] 状态获取单元，所述状态获取单元被配置为至少获取关于人的行为的状态信息；

[0294] 评估单元，所述评估单元被配置为通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及

[0295] 机器学习分类器，所述机器学习分类器对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

[0296] [2]根据[1]中所述的机器学习系统，

[0297] 其中，所述评估单元被配置为：

[0298] 基于状态信息和关于目标行为的目标状态信息之间的差异来计算奖励，以及

[0299] 基于奖励、环境信息和状态信息来计算价值函数。

[0300] [3]根据[1]或[2]中所述的机器学习系统，

[0301] 其中，所述系统保持包括多个目标行为信息的目标状态相关信息。

[0302] [4]根据[3]中所述的机器学习系统，

[0303] 其中，目标状态相关信息包括特定于时间的目标状态信息和/或特定于阶段的目标状态信息。

[0304] [5]根据[1]至[4]中任一项所述的机器学习系统，

[0305] 其中，环境信息包括关于气味、照明、温度、湿度、视频或声音的信息。

[0306] [6]根据[1]至[5]中任一项所述的机器学习系统，

[0307] 还包括气味控制单元，

[0308] 其中，所述气味控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制所生成的气味。

[0309] [7]根据[6]中所述的机器学习系统，

[0310] 还包括芳香化单元，

[0311] 其中，所述芳香化单元被配置为基于由所述机器学习分类器选择的环境信息使物品带有气味，以及

[0312] 所述机器学习分类器基于环境信息来确定所述气味控制单元和所述芳香化单元中的哪一个将生成气味。

[0313] [8]根据[1]至[7]中任一项所述的机器学习系统，

[0314] 还包括照明控制单元，

[0315] 其中，所述照明控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制要发射的光。

[0316] [9]根据[1]至[8]中任一项所述的机器学习系统，

[0317] 还包括空气调节单元，

[0318] 其中，所述空气调节单元被配置为基于由所述机器学习分类器选择的环境信息来控制温度和/或湿度。

[0319] [10]根据[1]至[9]中任一项所述的机器学习系统，

[0320] 还包括视频控制单元，

[0321] 其中，所述视频控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制要显示的视频。

[0322] [11]根据[1]至[10]中任一项所述的机器学习系统，

[0323] 还包括声音控制单元，

[0324] 其中，所述声音控制单元被配置为基于由所述机器学习分类器选择的环境信息来控制要播放的声音。

[0325] [12]根据[1]至[11]中任一项所述的机器学习系统，

[0326] 其中，价值函数被划分为多个价值组，以及

[0327] 所述机器学习分类器使用由所述多个价值组中的每个价值组保持的价值函数。

[0328] [13]根据[1]至[12]中任一项所述的机器学习系统，还包括：

[0329] 多个状态获取单元；以及

[0330] 实现难度等级计算单元，

[0331] 其中，所述实现难度等级计算单元被配置为基于由所述多个状态获取单元中的每个状态获取单元获取的状态信息来计算目标行为的实现难度等级。

[0332] [14]根据[13]中所述的机器学习系统，

[0333] 其中，实现难度等级包括实现率，所述实现率表示促进到目标行为的程度。

[0334] [15]根据[13]或[14]中所述的机器学习系统，

[0335] 其中，实现难度等级包括标准实现时间，所述标准实现时间表示促进到目标行为的标准时间。

[0336] [16]根据[13]至[15]中任一项所述的机器学习系统，

[0337] 其中，实现难度等级包括平均环境项目数量，所述平均环境项目数量表示当促进到目标行为时的环境信息中的项目的平均数量。

[0338] [17]一种机器学习设备，至少包括：

[0339] 状态获取单元，所述状态获取单元被配置为至少获取关于人的行为的状态信息；

[0340] 评估单元，所述评估单元被配置为通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及

[0341] 机器学习分类器，所述机器学习分类器对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

[0342] [18]一种机器学习方法，至少包括：

[0343] 至少获取关于人的行为的状态信息；

[0344] 通过评估关于当获取状态信息时在所述人周围的环境的环境信息和状态信息来获得价值函数；以及

[0345] 对价值函数执行强化学习，并且选择当价值函数最高时的环境信息，以便促进所述人来参与目标行为。

[0346] 参考标记列表

[0347] 1 机器学习系统

[0348] 10 机器学习设备

[0349] 11 状态获取单元

[0350] 12 评估单元

[0351] 13 记录单元

[0352] 14 机器学习分类器

[0353] 20 环境控制设备

[0354] 23 气味控制单元

[0355] 24 照明控制单元

[0356] 25 空气调节单元

[0357] 26 视频控制单元

[0358] 27 声音控制单元

[0359] 30 芳香化单元

[0360] 40 信息通信网络

[0361] 50 实现难度等级计算设备

[0362] 51 信息获取单元

[0363] 52 对象信息记录单元

[0364] 53 行为信息记录单元

[0365] 54 实现难度等级计算单元

[0366] S1 至少获取状态信息

[0367] S2 通过评估环境信息和状态信息来获得价值函数

[0368] S3 选择环境信息。

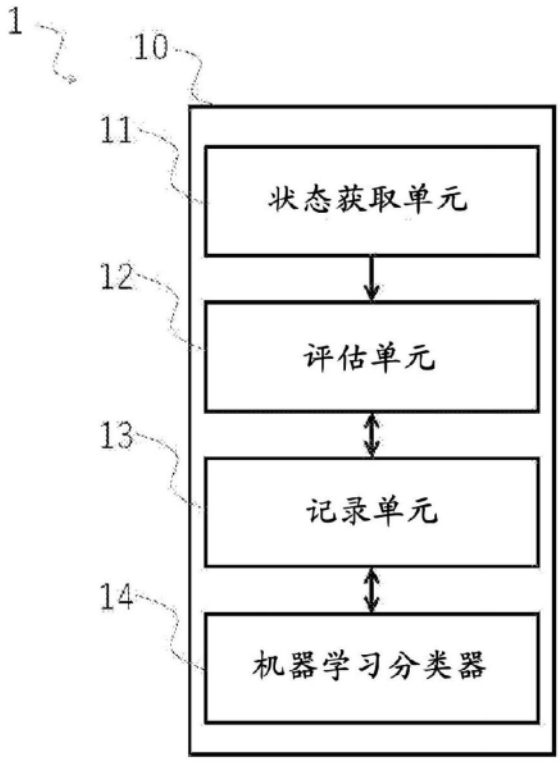

1
10
11
状态获取单元
12
评估单元
13
记录单元
14
机器学习分类器

图1

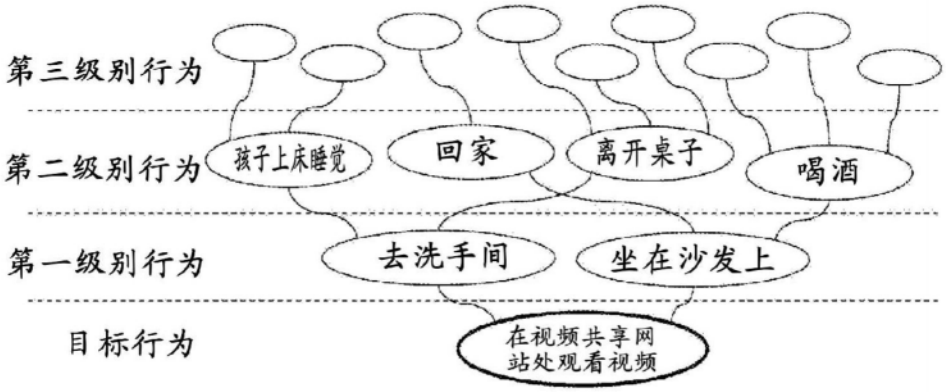

第三级别行为
第二级别行为
第一级别行为
目标行为
孩子上床睡觉
回家
离开桌子
喝酒
去洗手间
坐在沙发上
在视频共享网站处观看视频

图2

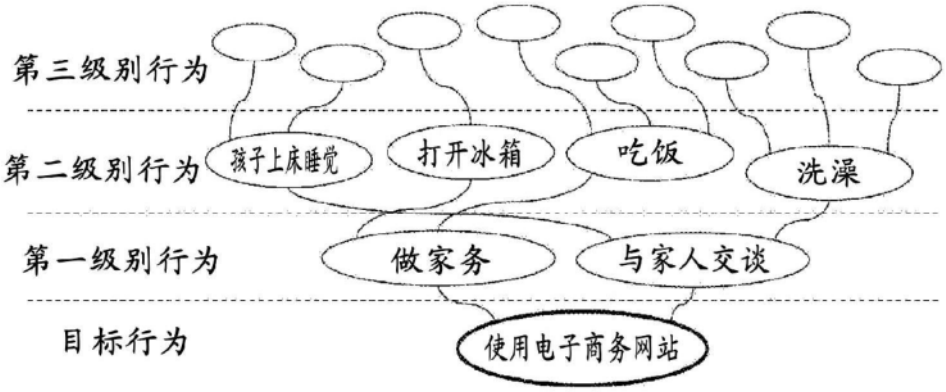

第三级别行为
第二级别行为
孩子上床睡觉
打开冰箱
吃饭
洗澡
第一级别行为
做家务
与家人交谈
目标行为
使用电子商务网站

图3

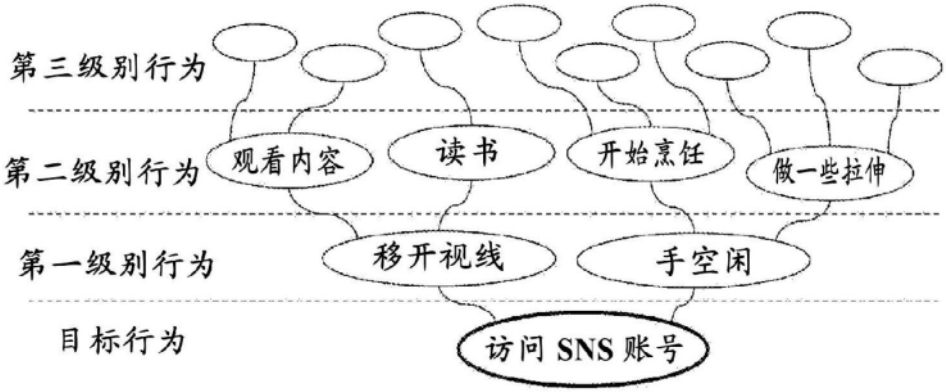

第三级别行为
第二级别行为
观看内容
读书
开始烹饪
做一些拉伸
第一级别行为
移开视线
手空闲
目标行为
访问 SNS 账号

图4

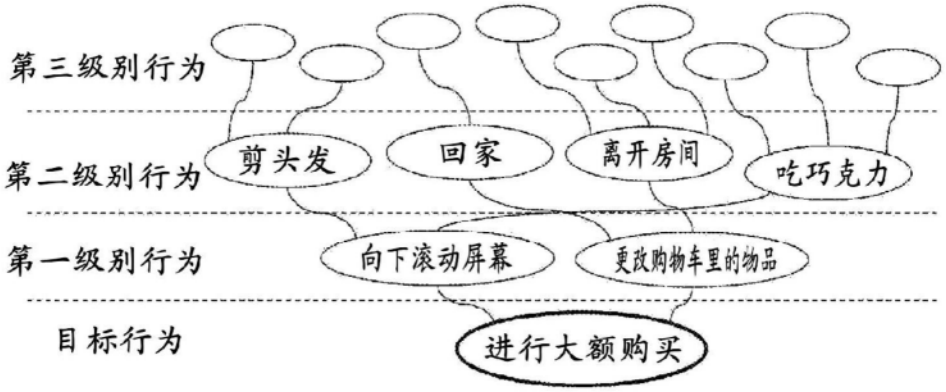

第三级别行为
第二级别行为
剪头发
回家
离开房间
吃巧克力
第一级别行为
向下滚动屏幕
更改购物车里的物品
目标行为
进行大额购买

图5

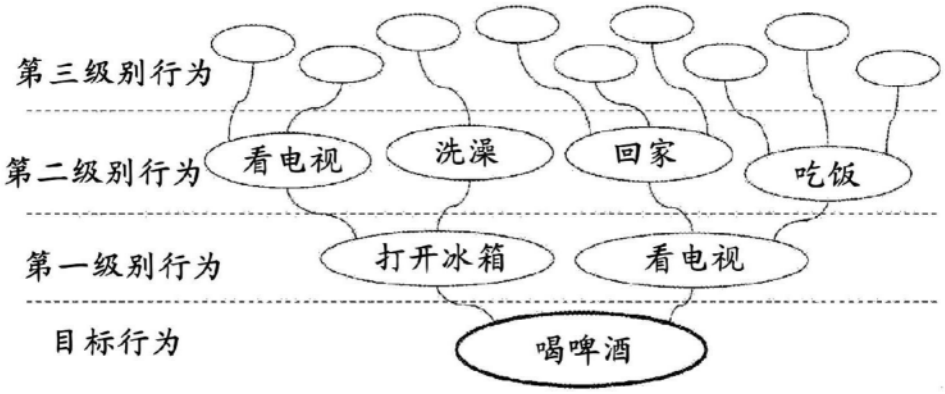

第三级别行为
第二级别行为
看电视
洗澡
回家
吃饭
第一级别行为
打开冰箱
看电视
目标行为
喝啤酒

图6

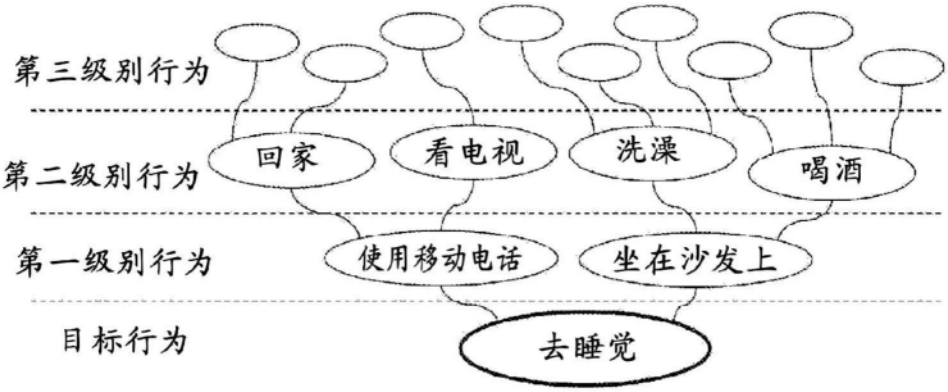

第三级别行为
第二级别行为
回家
看电视
洗澡
喝酒
第一级别行为
使用移动电话
坐在沙发上
目标行为
去睡觉

图7

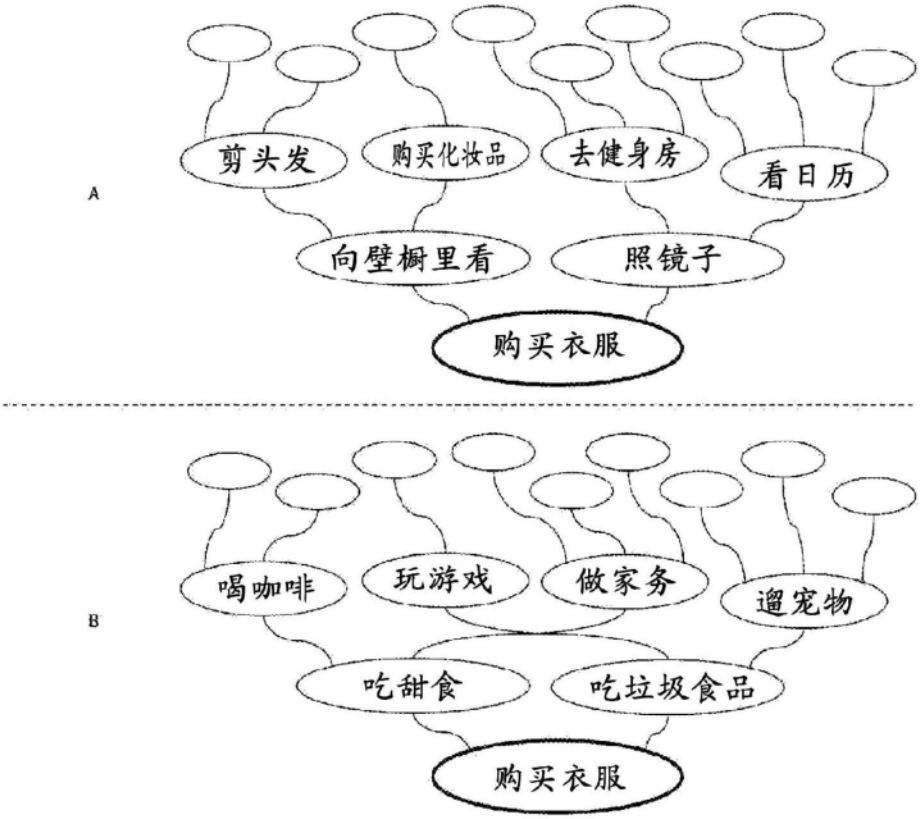


图8

时间段	目标行为
第一时间段（大约 12:00 AM-6:00 AM）	去睡觉
第二时间段（大约 7:00 AM-7:00 PM）	吃食物 S
第三时间段（大约 8:00 PM-11:00 PM）	喝饮料 T

图9

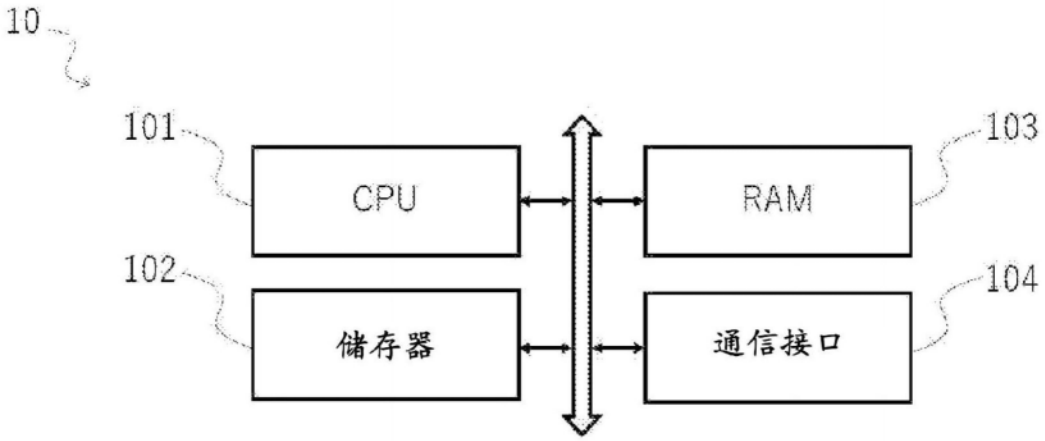

10
101
CPU
103
RAM
102
储存器
104
通信接口

图10

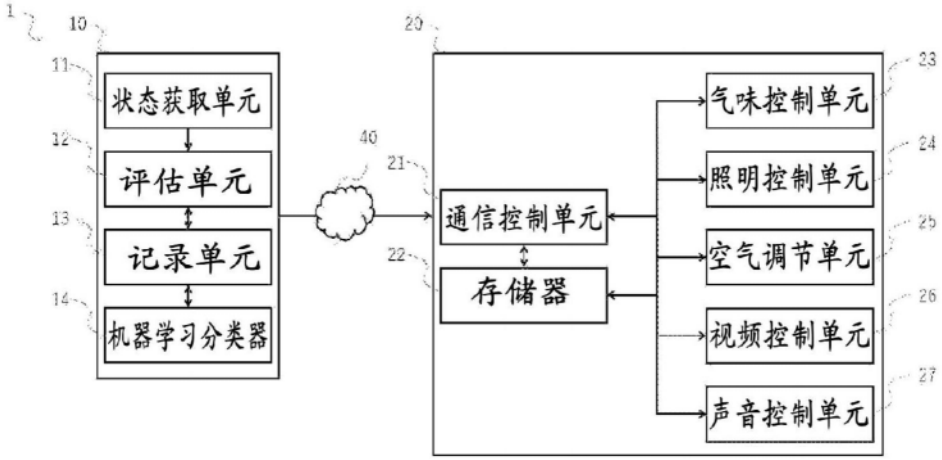

1
10
11
状态获取单元
12
评估单元
13
记录单元
14
机器学习分类器
40
20
21
通信控制单元
22
存储器
23
气味控制单元
24
照明控制单元
25
空气调节单元
26
视频控制单元
27
声音控制单元

图11

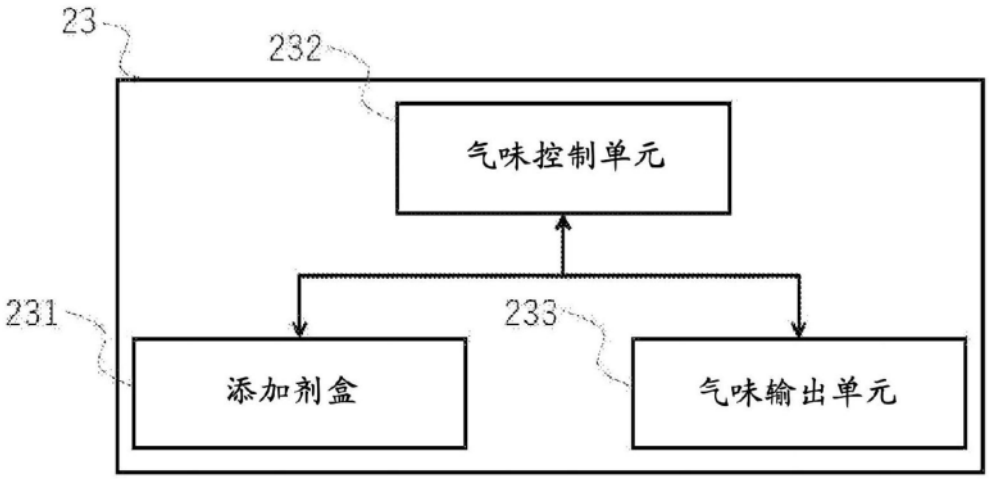

23
232
气味控制单元
231
添加剂盒
233
气味输出单元

图12

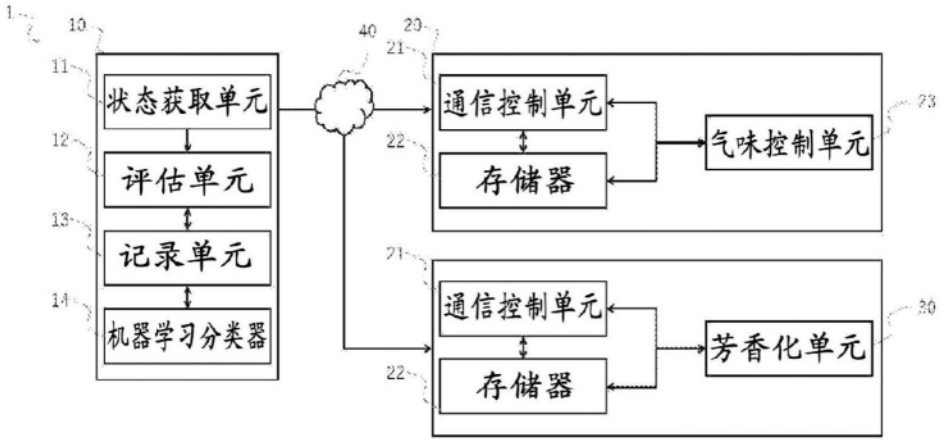

1
10
11
状态获取单元
12
评估单元
13
记录单元
14
机器学习分类器
40
20
21
通信控制单元
22
存储器
23
气味控制单元
21
通信控制单元
22
存储器
30
芳香化单元

图13

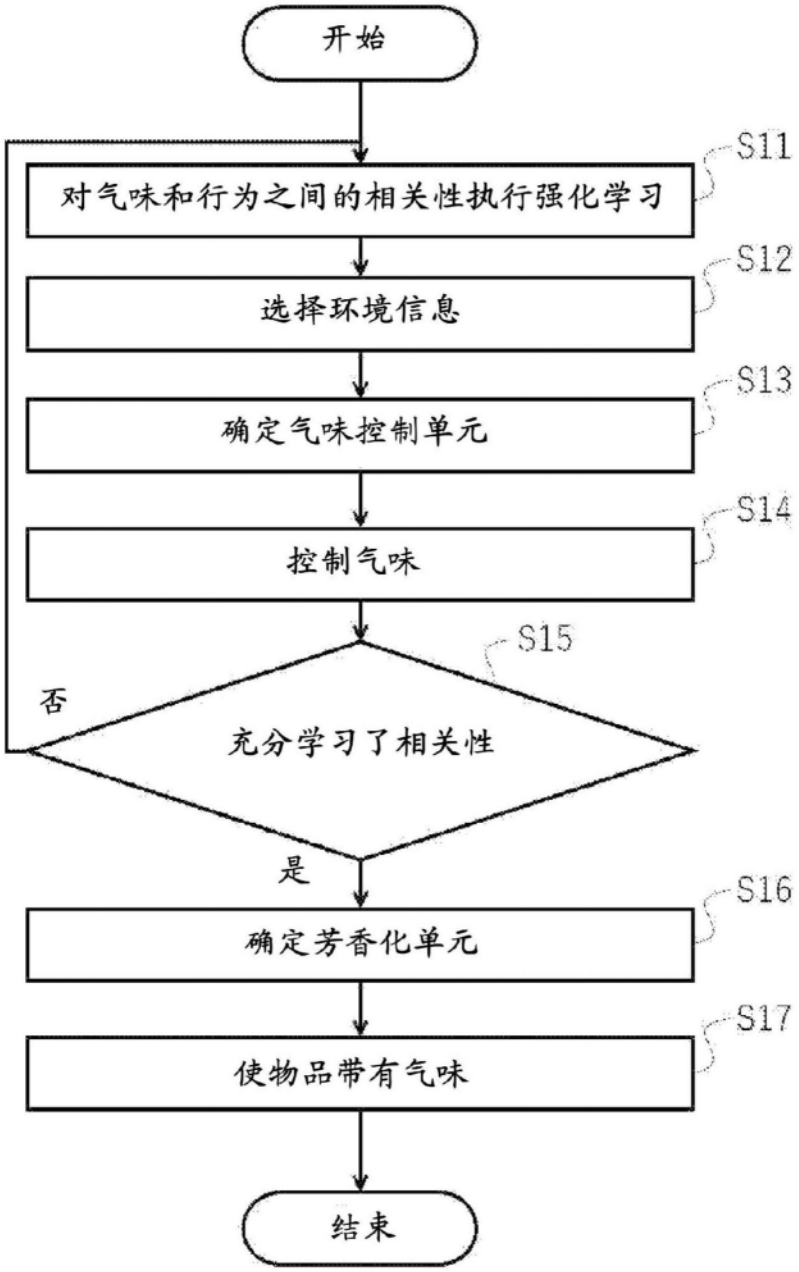

开始
S11
对气味和行为之间的相关性执行强化学习
S12
选择环境信息
S13
确定气味控制单元
S14
控制气味
S15
否
充分学习了相关性
是
S16
确定芳香化单元
S17
使物品带有气味
结束

图14

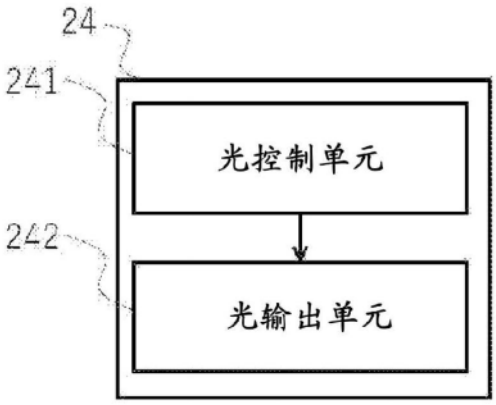

24
241
光控制单元
242
光输出单元

图15

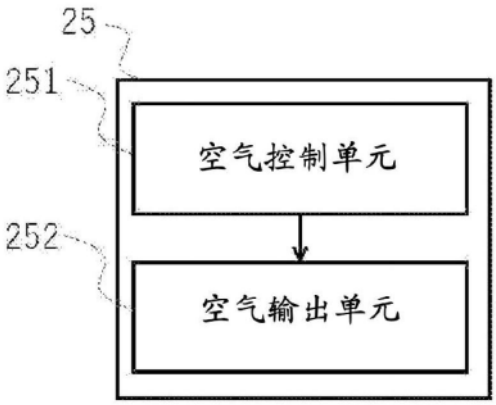

25
251
空气控制单元
252
空气输出单元

图16

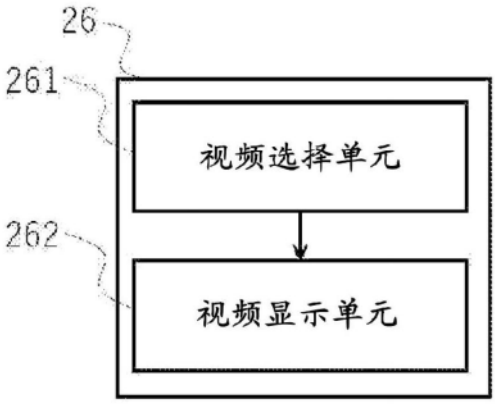

26
261
视频选择单元
262
视频显示单元

图17

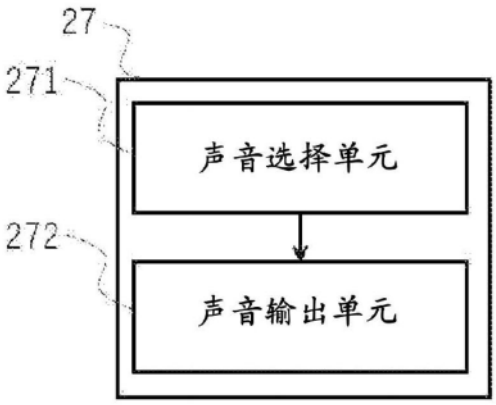


图18

目标行为	第一目标行为															
价值组	Q1				Q2				Q3				Q4			
状态信息	s1				s2				s3				s4			
环境信息	a1				a2				a3				a4			
属性	A	B	A	A	C	B	C	A	C	B	C	A	C	B	C	A
人	h01	h02	h03	h04	h05	h06	h07	h08	h09	h10	h11	h12	h13	h14	h15	h16

目标行为	第二目标行为															
价值组	Q5				Q6				Q7				Q8			
状态信息	s5				s6				s7				s8			
环境信息	a5				a6				a7				a8			
属性	A	B	A	A	C	B	C	A	C	B	C	A	C	B	C	A
人	h17	h18	h19	h20	h21	h22	h23	h24	h25	h26	h27	h28	h29	h30	h31	h32

图19

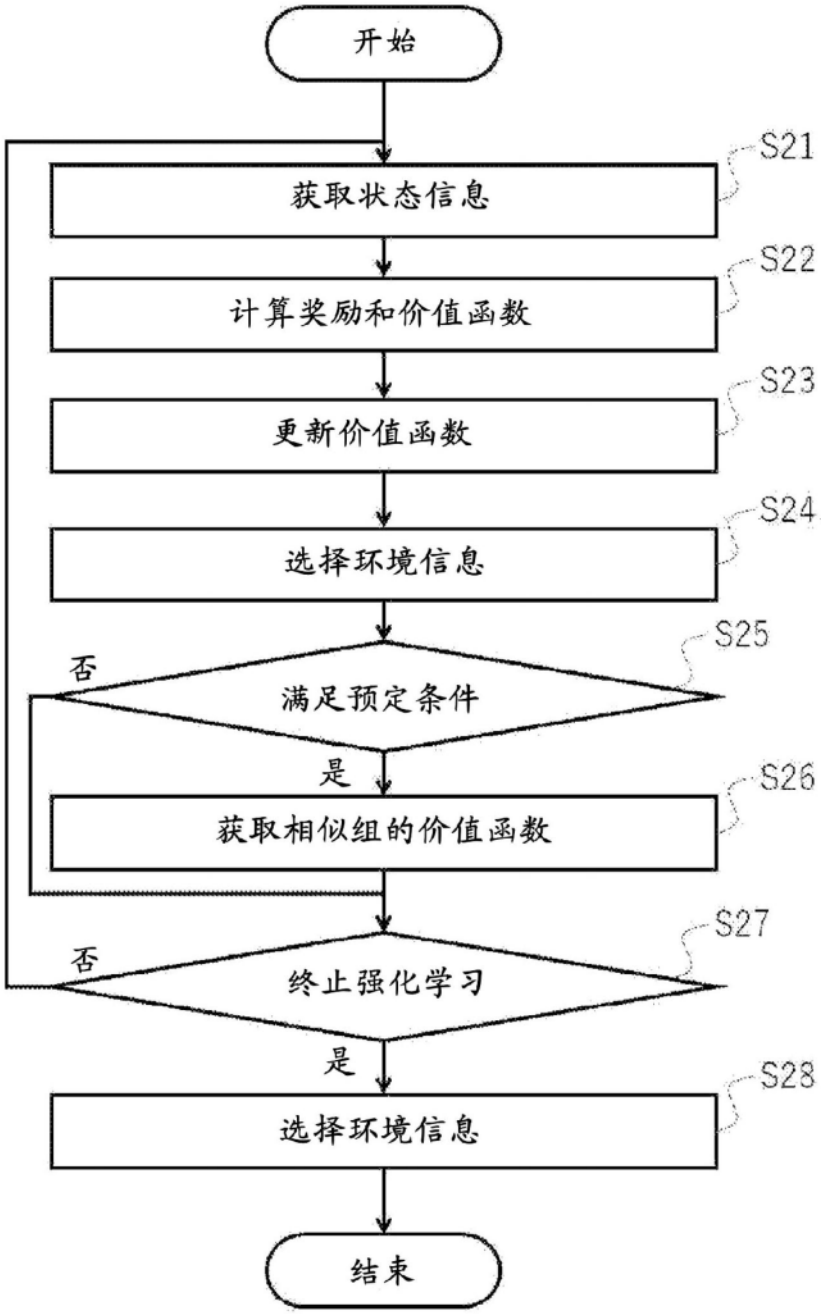

开始
S21
获取状态信息
S22
计算奖励和价值函数
S23
更新价值函数
S24
选择环境信息
S25
否
满足预定条件
是
S26
获取相似组的价值函数
S27
否
终止强化学习
是
S28
选择环境信息
结束

图20

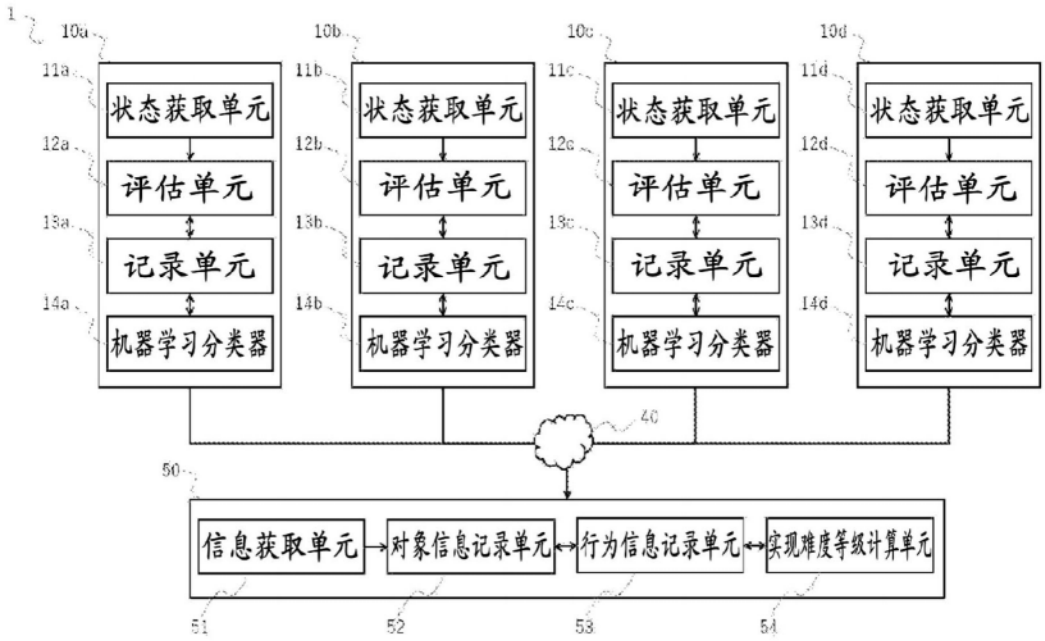

1
10a
10b
10c
10d
11a
11b
11c
11d
状态获取单元
12a
12b
12c
12d
评估单元
13a
13b
13c
13d
记录单元
14a
14b
14c
14d
机器学习分类器
40
50
信息获取单元
对象信息记录单元
行为信息记录单元
实现难度等级计算单元
51
52
53
54

图21

A

类别	子类别	子子类别
性别：男性或女性 年龄：20~59岁 AR：30% SAT：54个小时 NKV：2	性别：男性 年龄：20~59岁 AR：31% SAT：55个小时 NKV：1	性别：男性 年龄：20~39岁 AR：34% SAT：38个小时 NKV：1
		性别：男性 年龄：40~59岁 AR：28% SAT：72个小时 NKV：1
	性别：女性 年龄：20~59岁 AR：29% SAT：53个小时 NKV：3	性别：女性 年龄：20~39岁 AR：31% SAT：45个小时 NKV：2
		性别：女性 年龄：40~59岁 AR：27% SAT：61个小时 NKV：4

B

类别	子类别	子子类别
性别：男性或女性 年龄：20~79岁 AR：17% SAT：275.75个小时 NKV：1.5	性别：男性 年龄：20~79岁 AR：9.5% SAT：297.5个小时 NKV：1.5	性别：男性 年龄：20~49岁 AR：7% SAT：336个小时 NKV：2
		性别：男性 年龄：50~79岁 AR：12% SAT：259个小时 NKV：1
	性别：女性 年龄：20~79岁 AR：24.5% SAT：254个小时 NKV：1.5	性别：女性 年龄：20~49岁 AR：10% SAT：321个小时 NKV：2
		性别：女性 年龄：50~79岁 AR：39% SAT：187个小时 NKV：1

图22

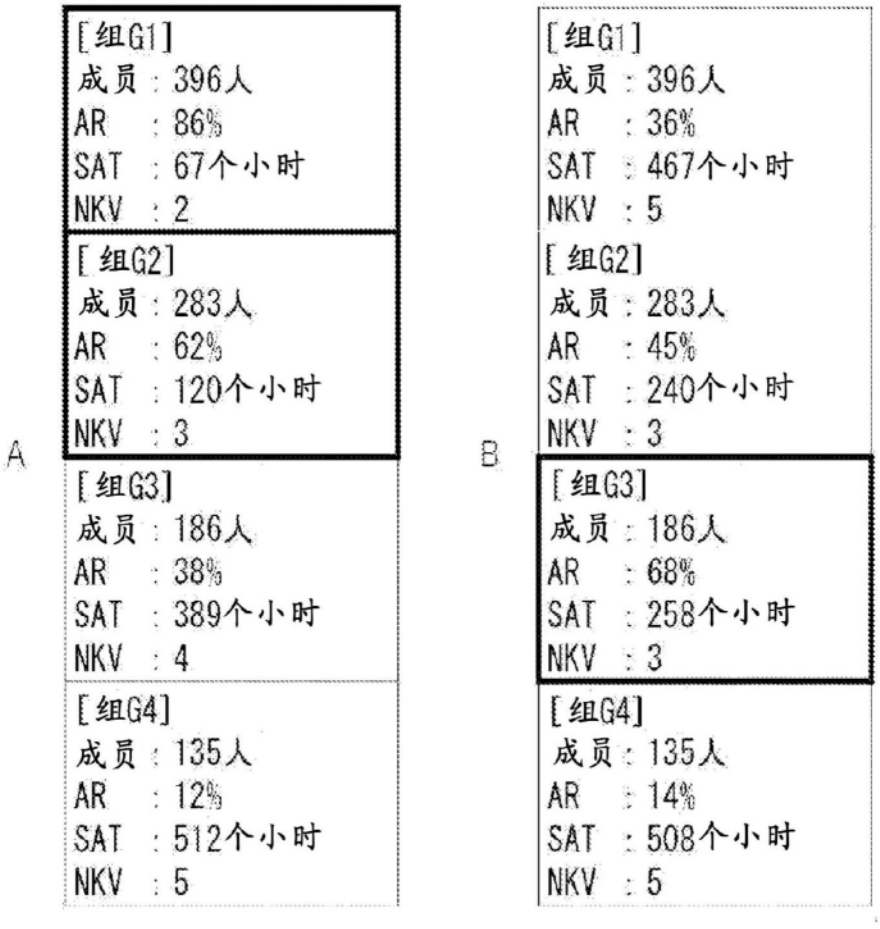

A
[组G1]
成员：396人
AR ：86%
SAT ：67个小时
NKV ：2
[组G2]
成员：283人
AR ：62%
SAT ：120个小时
NKV ：3
[组G3]
成员：186人
AR ：38%
SAT ：389个小时
NKV ：4
[组G4]
成员：135人
AR ：12%
SAT ：512个小时
NKV ：5
B
[组G1]
成员：396人
AR ：36%
SAT ：467个小时
NKV ：5
[组G2]
成员：283人
AR ：45%
SAT ：240个小时
NKV ：3
[组G3]
成员：186人
AR ：68%
SAT ：258个小时
NKV ：3
[组G4]
成员：135人
AR ：14%
SAT ：508个小时
NKV ：5

图23

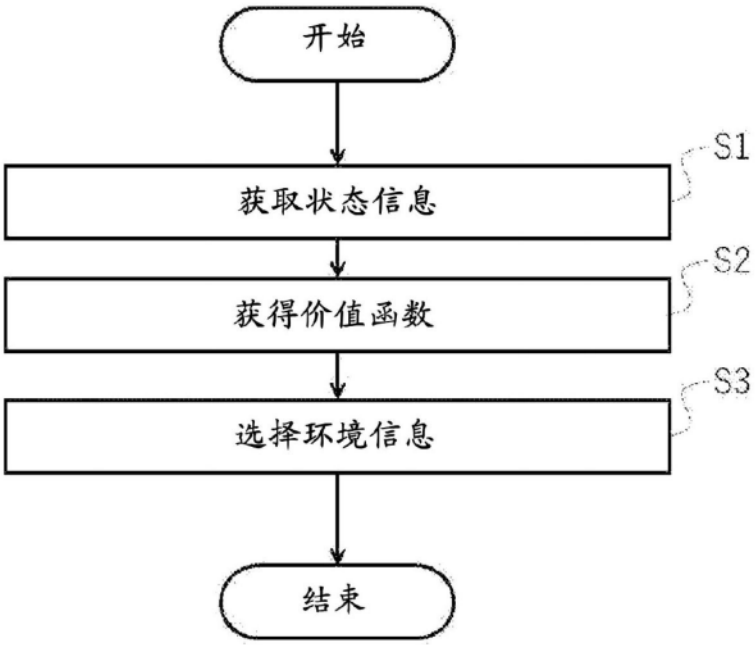

开始
S1
获取状态信息
S2
获得价值函数
S3
选择环境信息
结束

图24