**COMP9418**

**Assignment2 report**

Student Name：

Student zID:

对于本次作业，我们使用的是Hidden Markov Model. 为了简化模型，我们将房间作为每个单独的空间进行建模。

1. 数据预处理

由于我们只对房间开灯或者不开灯感兴趣，并不在意房间的人数。所以我们可以对训练的数据做了以下的简化：

1. 将r1-r10中的人数改为’on’, ‘off’，’on’对应人数大于0的情况，’off’对应人数为0的情况。
2. 将camera1-camera4中的人数改为’found’和’not found’，’found’对应人数大于0的情况， ‘not found’对应人数小于0的情况。
3. 将door\_sensor1-door\_sensor11中进出门的人数改为’not move’,‘even’和‘odd’。‘not move’对应门没有动的情况，‘even’对应人数为偶数的情况，‘odd’对应人数为奇数的情况。

下表为各个因素的outcomes表：

|  |  |
| --- | --- |
| room | (‘on’, ‘off’) |
| door\_sensor | (‘not move’, ‘even’, ‘odd’) |
| camera | (‘found’, ‘not found’) |

我们对测试数据sensor\_data中的camera,door\_sensor做了同上表一样的转化。

1. 数据建模

Hidden Markov Model需要三种概率分布，initial distribution P(X1), transition probability P(Xt | Xt-1)和emission probability P(Et | Xt)。

对于emission probability, 我们将每个房间中的sensor作为观察指向现在时间点的房间本身。

对于 ，我们将房间的上一个时间点指向现在的时间点。

以r1举例：

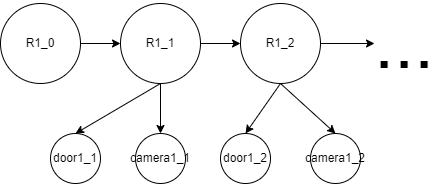


Figure 1: HMM model of r1

1. 公式计算

对于初始的P(X0)，我们将除了r3以外所有房间的概率设置为开灯为0.5, 关灯为0.5，将r3的概率设置为开灯为1，关灯为1。并将概率值存入字典distribution\_dict中。这是通过观察数据集data1.csv和data2.csv，在最初的时间里，r3有人的概率非常大。并且结合实际，所有人在8点进入大厦上班都需要进过r3。

对于P(Xt)的计算为:

P(Xt = ‘on’) = P(Xt-1 = ‘on’) \* P(Xt = ‘on’ | Xt-1 = ‘on’) \* P(et | Xt = ‘on’) + P(Xt-1 = ‘off’) \* P(Xt = ‘on’ | Xt-1 = ‘off’) \* P(et | Xt = ‘on’)

P(Xt = ‘off’) = P(Xt-1 = ‘on’) \* P(Xt = ‘on’ | Xt-1 = ‘off’) \* P(et | Xt = ‘off’) + P(Xt-1 = ‘off’) \* P(Xt = ‘off’ | Xt-1 = ‘off’) \* P(et | Xt = ‘off’)

将P(Xt = ‘on’)和P(Xt = ‘off’)进行归一化处理后，存入字典distribution\_dict中，作为下一轮的P(Xt-1)使用。

对P(Xt = ‘on’)和P(Xt = ‘off’)进行对比，如果P(Xt = ‘on’)大于P(Xt = ‘off’)，记录房间灯光情况为开灯，反之为关灯。

1. 时间复杂度
2. 结论