

**人工智能实验报告**

题 目 不确定性推理

专 业 人工智能（视听觉信息处理）

学　　 号 7203610121

姓 名 刘天瑞

同 组 人 员 无

1. **背景简介/问题描述**

**实验内容：**

参照课程第五部分讲授的贝叶斯网络完成。给定事件和事件之间的关系，并且给出每个事件的CPT图，编写程序自动读取文档实现基于贝叶斯网络的推理, 利用枚举或者变元消除法求出目标概率的精确解，利用采样方法求出目标概率的精确解。

**说明：**

在这里用到的贝叶斯算法是建立在有向无环图和CPT表的技术上实现的。

首先，给定的输入文件格式为：

N

rv0 rv1 ... rvN-1

0 0 1 ... 0

1 0 0 ... 1

...

0 1 1 ... 0

mat0

mat1

...

matN-1

在这里:

• N 是贝叶斯网络中随机事件的数目

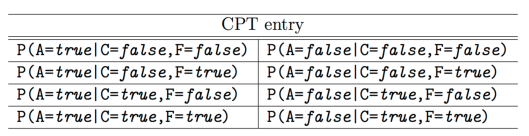
• rv 是随机事件的名字（字符串形式表示）

• mat是一个二维数组，分别表示从他的父亲到其本身的可能性概率。第一个元素表示发生的概率，第二个元素表示不发生的概率，显然两个元素相加为1。

在上述中mat即为CPT表（Conditional Probability Table），其被设计为如下格式：

对于每个节点，如果他有N个父节点，则其CPT表中有2^N列，我们记为标号0 - （2N-1），其行序号的定义方法如下，利用二进制分别表示对应的父亲为是否发生，1为发生，0位不发生，将得到的二进制数转化为十进制代表其对应的行号。举例如下：

A有两个父节点C，F，则CPT表如下表所示：



其次，编写程序对应的查询格式为：P(rvQ |rvE1=val, rvE2=val, ...)

rvQ表示查询的条件的名字，即在rvE1=val, rvE2=val,..发生的条件下，rvQ发生的概率。

RvEx表示条件的名称，而后面的val为true/false，分别表示发生和不发生。

最后，输出格式为两个数据分别表示P(QueryVar=true|...) 和P(QueryVar=false|...))。例如：

0.872 0.128

1. **算法介绍**

2.1 所用算法及的解题思路的一般介绍

2.1.1 所用方法一：采样分类法

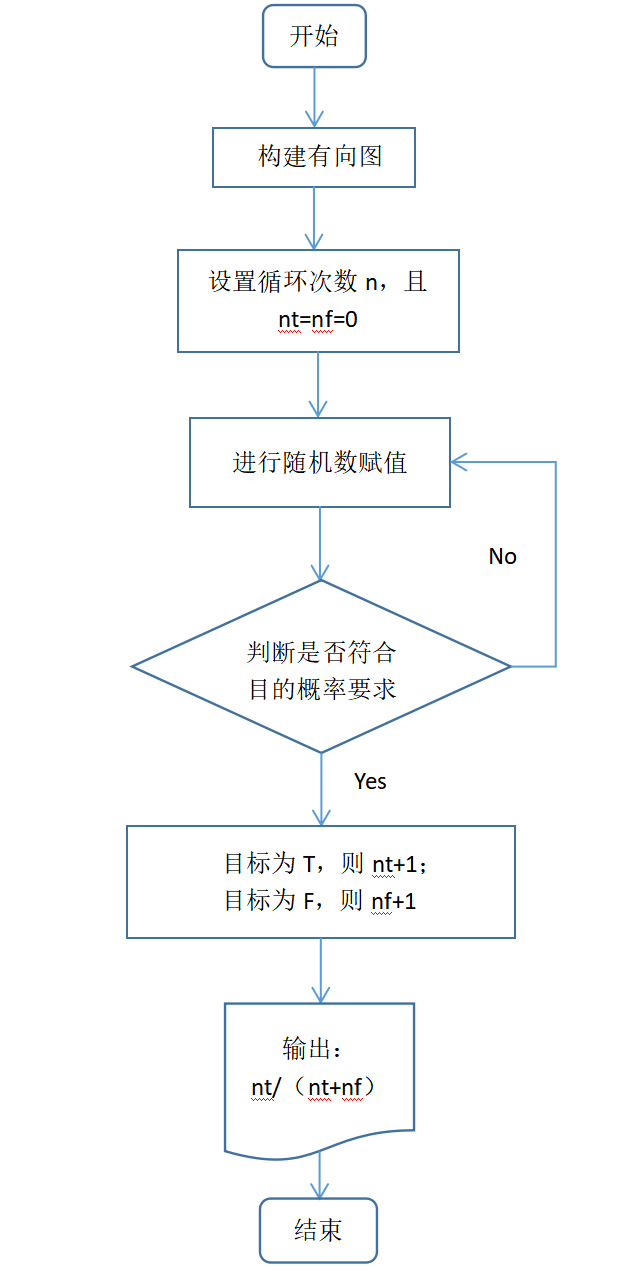
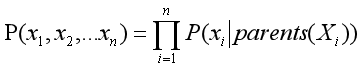
尝试使用近似法对概率进行求解。具体操作为构建完Bayes网络图后自上而下地进行根据节点概率来随机计算赋值的操作，然后统计其概率最终得到目标概率，如下图所示，为采样分类方法的具体流程图：

图1. 采样分类法的流程图1

2.1.2 所用方法二：枚举分类法

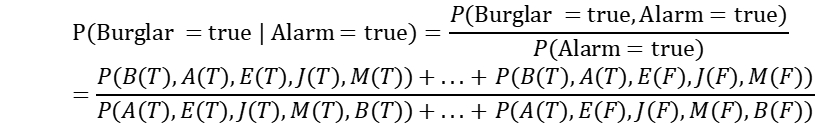
Bayes网络是一种概率图模型，是一种模拟人类推理过程中因果关系的一种不确定性推理处理模型，它的网络拓朴结构是一个有向无环图。 根据以上的实验文本内容来计算全概率，可以使用以下公式：

 式（1）

计算全概率的方法有很多，但是根据实验文本内容来讲可以只选择使用父节点条件概率来对全概率进行计算，再将所有情况计算出来后写于表中保存。而如何使用这些概率进行计算，这说明还需要用到经典Bayes公式：

 式（2）

如此代入便可以将上述需要计算的概率巧妙地转化为使用全概率来计算，例如假设需要计算的第一个概率：



这些概率都已通过父节点的条件概率计算得出，故均可得到，如下图所示，为枚举分类方法的具体流程图：

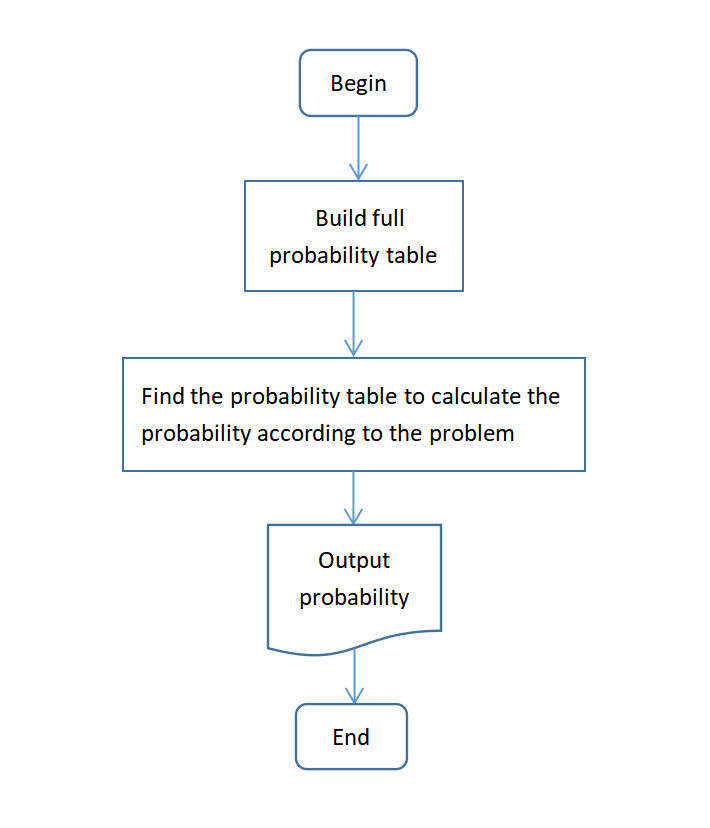


图2. 枚举分类法的流程图2

1. **算法实现**

3.1 实验结果

3.1.1 所用方法之一：采样分类方法：

重点主要在于计算所有全概率时计算在对应CPT表中的具体位置以及处理Bayes网络文本并返回概率表时，具体实现代码如下图所示：



图3. 计算所有全概率



图4. 计算在对应CPT表中的位置

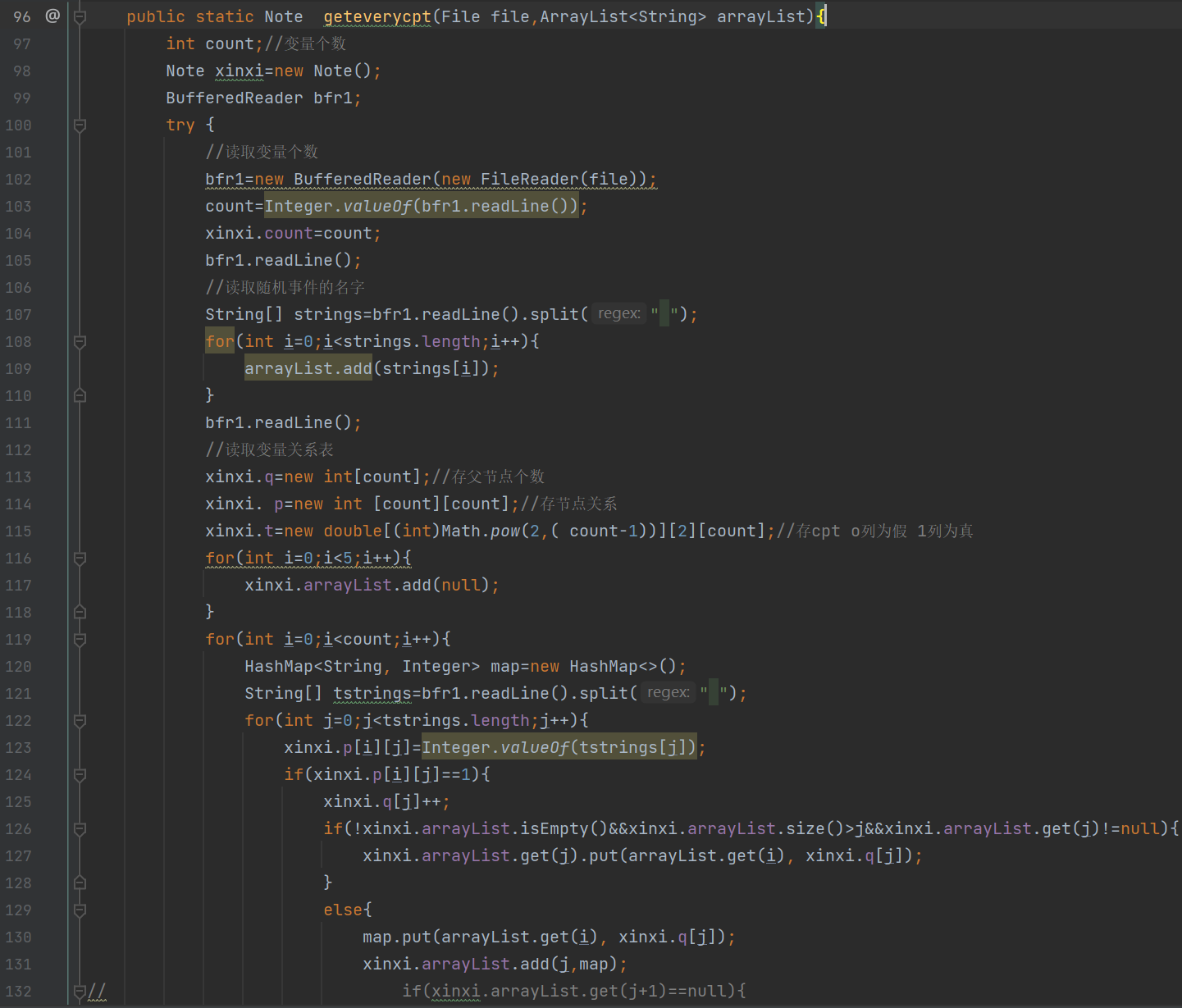
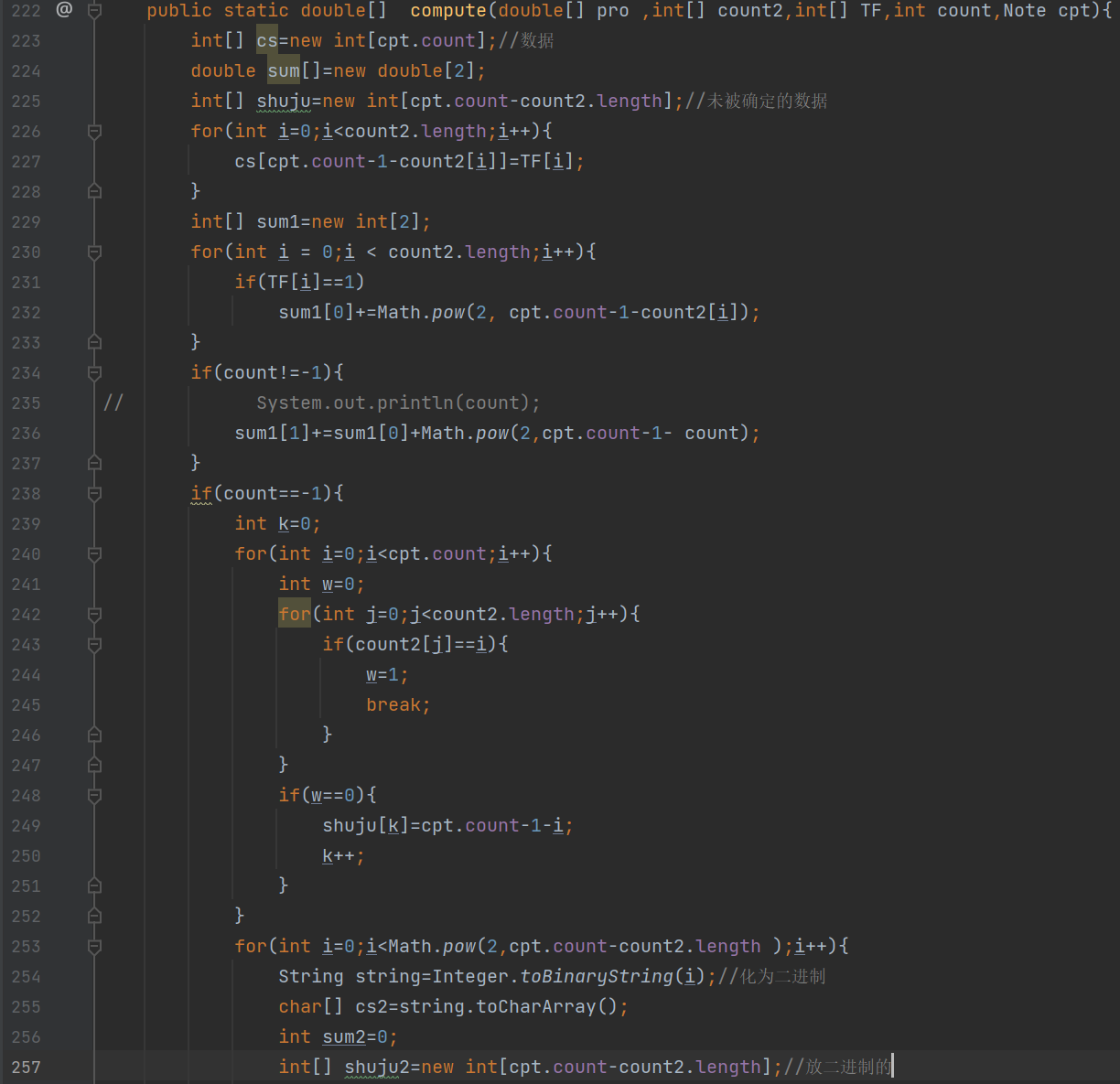




图5. 处理Bayes网络文本并返回概率表

3.1.2所用方法之二：枚举分类法：

主要实现部分在于构建全概率表及其使用计算，具体实现部分代码如下图所示：





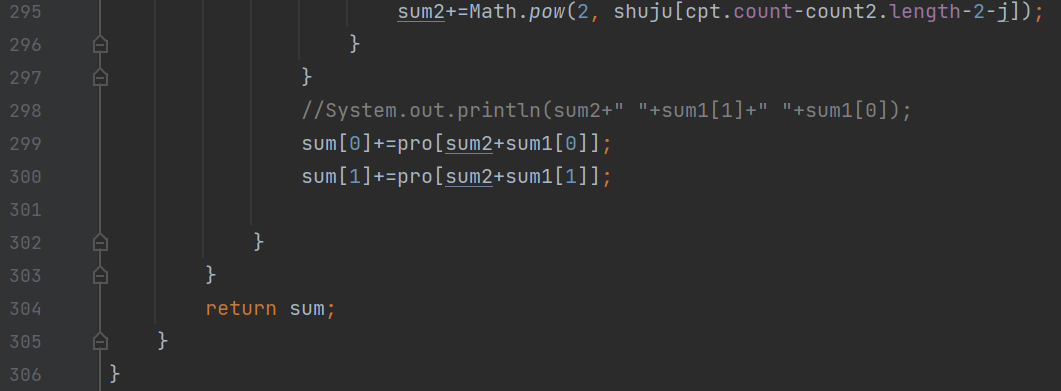


图6. 计算概率

1. **讨论及结论**

4.1 结论

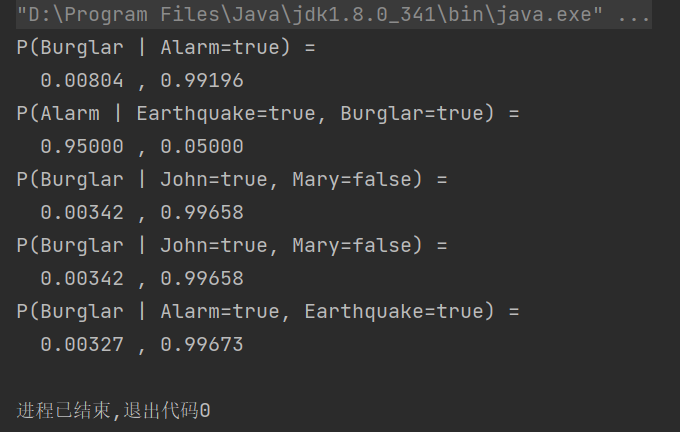


图7. Burglarnetwork实验结果

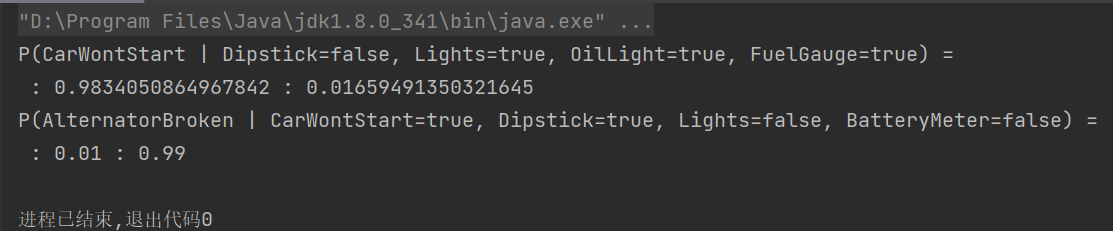


图8. Carnetwork实验结果

4.2 讨论

具体算法中采用的枚举分类法需要使用较为复杂的数据结构提供支持庞大繁琐的局量计算，所耗费的空间复杂度较大；而采样分类法如果要获取比较精确的概率结果则需要进行足够多次的循环与返回概率表，但是这样同时也导致了其运行时间较长，耗费的时间复杂度极大。

综上所述，以上这两种方法同时体现出了以空间换时间或者以时间换空间的算法优化思想。

**参考文献**

[1] [贝叶斯网络，看完这篇我终于理解了(附代码)！ - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/73415944) <https://zhuanlan.zhihu.com/p/73415944>

[2] [机器学习——贝叶斯算法 - 知乎 (zhihu.com)](https://zhuanlan.zhihu.com/p/158771382) <https://zhuanlan.zhihu.com/p/158771382>

注：报告中图的下方要有图题（如图1. XXXX），表格需要用三线表，表头需列在表的上方（如表1. XXXX）。图表居中排列。