

# 文献阅读作业：隐马尔可夫模型

## 文档简介

《语音信号数字处理》课程的文献阅读作业采用课下自主阅读、完成作业+课堂讲解、课堂提问的形式进行：由同学们自行查找并阅读隐马尔可夫模型（Hidden Markov Model, HMM）的相关文献资料，回答本文档后面的问题，在网络学堂上提交作业，并在课堂上进行抽查提问。锻炼同学们查找论文、阅读论文、归纳总结等能力。

课下自主阅读、作业需要在第四周周一之前完成，需通过查找资料理解隐马尔可夫模型基本原理，回答本文档后面的问题，并通过网络学堂提交（截止时间 10 月 4 日 23:59）。该作业旨在让同学们对 HMM 相关算法有更深入的理解，有助于第四周课堂讲授知识的理解。

课上讲解、课堂提问将于第四周周二（10 月 5 日）上午正常上课时间进行，课上将结合 HMM 讲义内容以及本文档所提之问题进行详细的讲解，并在相关问题处进行课堂提问（抽查或全体回答）。

## 中英文术语对照表

Markov Process: 马尔可夫过程

First Order Markov Process: 一阶马尔可夫过程

(First Order) Markov System: (一阶)马尔可夫系统

Hidden Markov Model (HMM): 隐马尔可夫模型

Gaussian Mixture Model (GMM): 高斯混合模型

Automatic Speech Recognition (ASR): (自动)语音识别

Text-To-Speech (TTS): (文语转换)语音合成

## 完成以下任务

### 任务一：查找阅读文献资料（30 分）

(1) 请查找并阅读至少 3 篇参考文献资料，其中：至少 1 篇为经过同行评审的正式发表的英文论文（会议或期刊均可）；不得多于 1 篇为知乎、博客、百科等非正式发表的科普性阅读资料或讲义。（20 分）

(2) 以标准参考文献的格式列出所阅读的参考文献列表。（5 分）

(3) 在回答相应的问题时给出明确的引用。(5 分)

## 任务二：回答以下问题（70 分）

(1) 什么是马尔可夫过程？一阶马尔可夫过程的特点是什么？（可结合图）对一阶马尔可夫过程及一阶马尔可夫系统进行明确的形式化定义。(5 分)

(2) 什么是隐马尔可夫模型？（可结合图）对隐马尔可夫模型进行明确的形式化定义。(5 分)

(3) 给出隐马尔可夫模型的 3 个典型问题的形式化描述。(5 分)

(4) 给定隐马尔可夫模型 $\lambda$ ，如何计算该模型生成某个观察序列  $O$  的概率？请给出数学公式及对应的推导。(10 分)

(5) 给定隐马尔可夫模型 $\lambda$ 、以及对应的观察序列  $O$ ，如何得到该模型产生该观察序列的最优隐含状态的序列  $Q$ ？请给出数学公式及对应的推导。(10 分)

(6) 给你一堆观察数据  $O_1, O_2, \dots, O_N$ ，并假设这些观察数据符合隐马尔可夫过程的假设，如何估计出一个最优的隐马尔可夫模型 $\lambda$ ，该模型产生这些观察数据的概率最优。(10 分)

(7) 什么是高斯混合模型？高斯混合模型如何与隐马尔可夫模型结合？(10 分)

(8) 请结合相关文献，介绍一个隐马尔可夫模型在语音及自然语言处理领域的最新应用实例，包括但不限于语音识别、语音合成等。注意需要解释清楚模型是如何应用到实际任务中的，以及是否存在方法上的改进。(15 分)

## 附录

### HMM 经典文献

- L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, vol. 77, no. 2, pp. 257–286, 1989.

### 语音领域相关会议、期刊名称

- IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)
- IEEE International Conference on Acoustics, Speech & Signal Processing (ICASSP)
- Annual Conference of International Speech Communication Association (INTERSPEECH)

### 关于作业参考文献的规范要求(参考 ICASSP 论文模板要求)

- REFERENCES: List and number all bibliographical references at the end of the paper. The

references can be numbered in alphabetic order or in order of appearance in the document. When referring to them in the text, type the corresponding reference number in square brackets as shown at the end of this sentence [2].

- The end of the document should include a list of references containing information similar to the following examples (Pay attention to all format details, including but not limited to authors' name list, lower/upper-cases, punctuations, etc.):

[1] D. E. Ingalls, "Image Processing for Experts," IEEE Transactions on Acoustics, Speech, and Signal Processing (TASSP), vol. TASSP-36, pp. 1932-1948, 1988.

[2] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, vol. 77, no. 2, pp. 257-286, 1989.

[3] Y. Lei, S. Yang, and L. Xie, "Fine-Grained Emotion Strength Transfer, Control and Prediction for Emotional Speech Synthesis," IEEE Spoken Language Technology Workshop (SLT), pp. 423-430, 2021.