

我们可以围绕一个结合新兴领域（如AI安全、智能交通、聊天机器人、图灵测试验证等）的主题，设计一个使用**有穷状态自动机（FSA）**的系统。我为你草拟了一个完整的大作业报告结构，主题为：

题目：基于有穷状态自动机的网络钓鱼邮件识别系统设计

以下是根据你上传的模板，填好的各部分内容草案：

1 引言

随着信息技术的迅速发展，电子邮件已成为重要的通讯工具。然而，网络钓鱼邮件层出不穷，给用户带来财产和信息安全的威胁。现有大多数识别方法依赖于机器学习模型，虽然精度高但可解释性差，难以适用于一些资源受限场景。本文尝试利用形式语言与自动机中的有穷状态自动机（Finite State Automaton, FSA）方法，对钓鱼邮件中的典型行为模式进行建模，构建一个可解释性强、逻辑清晰的轻量级识别系统，具有一定的创新性和实用价值。

2 问题描述

网络钓鱼邮件通常具有如下特点：

- 包含伪造的网址链接；
- 使用诱导性关键词（如“免费”、“紧急”、“立即行动”等）；
- 发件人地址可疑；
- 格式较混乱，常夹带不一致信息。

本设计目标是：使用确定有穷状态自动机（DFA），构建一个简化的钓鱼邮件识别模型，输入为邮件内容，输出为“正常”或“钓鱼邮件”判断结果。

3 解决方案

3.1 自动机设计思路

我们将对邮件文本按词或关键短语进行处理，用自动机的状态转换模拟文本中关键词及结构的模式识别。若路径最终落入“钓鱼判定状态”，则输出“钓鱼”。

3.2 五元组定义

- Q: 状态集合
{S0 (初始状态)、S1 (发现诱导词)、S2 (发现可疑URL)、S3 (发现伪发件人)、SA (钓鱼判定状态)、SN (安全状态)}
- Σ : 输入符号集
包括常见钓鱼邮件关键词，例如：
"免费"、"点击"、"紧急"、"http://"、"login"、"银行"、"更新信息"、"奖品"等

- δ : 状态转移函数
如:
 $\delta(S0, "免费") = S1$
 $\delta(S1, "http://") = S2$
 $\delta(S2, "login") = SA$
 $\delta(S0, "http://") = S2$
 $\delta(S2, "奖品") = SA$
..... (详见下方状态图)
- q_0 : 初始状态: $S0$
- F : 接受状态集: $\{SA \text{ (钓鱼邮件)}\}$

3.3 状态转移图 (δ 函数图)

(可绘制一个状态图，用圈表示状态，箭头表示转移，标注输入符号。)

3.4 DFA运行示例

输入: “免费赢大奖，请点击 <http://xxx.com> 登录获取奖励”
状态序列: $S0 \rightarrow S1 \text{ (免费)} \rightarrow S2 \text{ (点击)} \rightarrow SA \text{ (http)} \rightarrow \text{输出“钓鱼”}$

4 课程收获与感悟

通过本次作业，我们深刻体会到形式语言与自动机在现实中的应用价值，尤其是在网络安全等新兴领域的潜力。相比黑盒机器学习方法，基于自动机的设计更可解释、逻辑更清晰，适合对模型可控性有较高要求的系统。同时也锻炼了我们的建模与理论抽象能力，加深了对 DFA 结构的理解。

5 工作量描述

姓名	工作内容	自评分 (10分)
张三 (组长)	自动机模型设计、状态图绘制、文档整合	10
李四	文本分析与关键词提取、案例测试与模拟	9
王五	五元组定义、转移函数编写与验证	9

是否需要我帮你将这份内容整理成 Word 文档或 Jupyter Notebook 格式？如果你想使用不同的主题，例如 **语音识别指令控制系统**、**智能客服对话状态建模**、**自动交通灯控制逻辑**等，也可以告诉我，我可以为你换一个方向重新写。