

常见疾病诊断专家系统

史一哲 17300290027

引言

随着计算机和人工智能技术的不断发展，Expert System 的发展一直是最活跃的研究分支之一。它应用人工智能技术模拟人类专家求解问题的思维过程，从人类专家那里获得专业知识，并能通过一定的规则和学习策略进行知识的学习和更新，通过学习该系统的水平可以达到甚至超过人类专家的水平。Expert System 已经在各个领域都有了广泛的应用，该项目着眼于目前比较热门的医疗领域，结合医学专家的诊断思维过程，对常见的疾病诊断构建了专家系统，以对常见的疾病症状进行诊断并推荐相应的治疗方案。

目录

- [数据库的采集](#)
- [医疗诊断专家系统](#)
 - [医疗诊断专家系统的结构](#)
 - [规则的构建和表示](#)
 - [算法描述与分析](#)
- [用户界面](#)
- [测试用例](#)
- [意义与不足](#)

附录

[参考文献](#)

数据库的采集

对于一个疾病诊断系统而言，不仅要有规则数据库，更重要的是知道各种疾病的症状以及治疗方案。本专家系统数据库的数据主要来源于北京协和医院、复旦大学附属中山医院、浙大一院等知名医院官网；科普中国网、中国公众健康网等网站以及《默克家庭诊疗手册》等书籍(排名不分先后)，在此特别需要鸣谢医学院的学长学姐的热心帮助，帮助我从数据库中更正了很多用语规范以及剔除修改了一些不太准确的记录，使得整个系统的可信度更高。

下图展示了疾病数据库的结构和数据的部分截图，共记录了2000种疾病及其症状。

```
CREATE TABLE `disease` (
  `disease_id` int(11) NOT NULL AUTO_INCREMENT COMMENT '疾病编号',
  `disease_name` varchar(255) COLLATE utf8_unicode_ci NOT NULL COMMENT '疾病名称',
  `disease_trans` varchar(255) COLLATE utf8_unicode_ci NOT NULL COMMENT '疾病名称音译',
  `disease_alias` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病别名',
  `disease_intro` varchar(2000) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病简介',
  `disease_incidence_site` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病发病部位',
  `disease_contagious` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病的传染性',
  `disease_multiple_people` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病多发人群',
  `disease_symptom_early` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病早期症状',
  `disease_symptom_late` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病晚期症状',
  `disease_symptom_related` varchar(255) COLLATE utf8_unicode_ci DEFAULT NULL COMMENT '疾病相关症状',
```

198	胆汁返流性胃炎	dzflxwy	碱性返流性胃炎	胆汁返流性胃炎，也称胃	无传染性	成年人		
199	脂肪瘤	zfl		脂肪瘤(lipoma)是起源其他	无传染性	多见男性中青年		
200	肾病综合征	sbzhz	肾病综合征,肾综,肾	肾病综合征(nephrotic 肾	无传染性	所有人群		
201	分泌性中耳炎	fmxzey	胶耳,中耳炎,渗出性中	分泌性中耳炎(secretor 耳	无传染性	儿童。所有人群		
202	肛瘘	gl2	肛漏	肛管直肠瘘主要侵犯肛	无传染性	成年人		
203	沙眼	sy	颗粒性结膜炎,砂眼,衣	沙眼(trachoma)是由于	无传染性	所有人群, 少年人。男	患眼异物感、流泪; 受累的脸结膜发	
204	窦性心动过速	dxsdgs	心动过速	窦性心动过速(sinu	无传染性	中老年人		
205	胃炎	wy2	胃肠感染,肠胃炎	胃炎是指由各种因素引	无传染性	饮食不节人群, 男性多		
206	中暑	zs	伤暑	中暑是由于高温环境中	无传染性	无特殊发病群体, 大量出汗、口渴、明显面色苍白、恶		
207	强迫症	qpz		强迫性障碍(obsessive 心理	无传染性	有不良个性特征的群		
208	先天性心脏病	xtboxb	先心病,心脏病,小儿先	天性心血管病是先天心	无传染性	先天性婴幼儿		
209	支气管炎	zqgy	慢性支气管炎,慢支,支	气管炎是指气管、支气	无传染性	所有人群, 中老年人		
210	登革热	dgr	波儿加热,古典登革热,登	革热(dengue fever 全身	有传染性	所有人群, 小孩多数	有畏寒或寒战症状, 属为脑膜炎, 出	
211	老年痴呆症	aechmb	老年痴呆,阿尔茨海默病	阿尔茨海默病(Alzheim 颅脑	无传染性	老年人		
212	鸡眼	jy	胼胝, 肉刺	鸡眼和胼胝是皮肤由于	无传染性	所有人群		
213	喉癌	ha	喉府	喉癌(carcinoma of lai	无传染性	男性较女性多见, 约为	咽喉部异物感和疼痛, 在癌破溃后, 可	

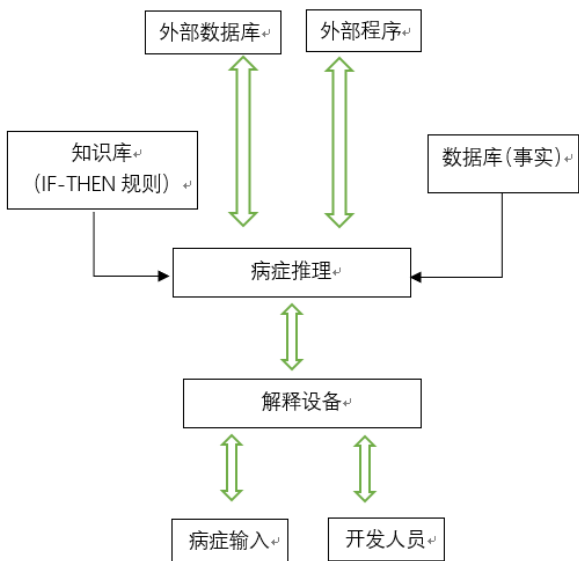
然后对数据进行清洗，清洗出来每种病症对应的症状，后面就可以通过清洗后的数据进行疾病诊断，然后对应清洗前的数据给出相应疾病的诊治措施。

```
INSERT INTO `t_symptom_di` VALUES ('2182', '剧痒', '365', '丘疹性荨麻疹');
INSERT INTO `t_symptom_di` VALUES ('2183', '斑丘疹', '365', '丘疹性荨麻疹');
INSERT INTO `t_symptom_di` VALUES ('2184', '丘疹', '365', '丘疹性荨麻疹');
INSERT INTO `t_symptom_di` VALUES ('2185', '咽喉干燥及灼热感', '366', '急性咽炎');
INSERT INTO `t_symptom_di` VALUES ('2186', '头痛', '366', '急性咽炎');
INSERT INTO `t_symptom_di` VALUES ('2187', '关节酸痛', '366', '急性咽炎');
INSERT INTO `t_symptom_di` VALUES ('2188', '咳嗽', '367', '结核性胸膜炎');
INSERT INTO `t_symptom_di` VALUES ('2189', '胸痛', '367', '结核性胸膜炎');
INSERT INTO `t_symptom_di` VALUES ('2190', '发烧', '367', '结核性胸膜炎');
INSERT INTO `t_symptom_di` VALUES ('2191', '气促', '367', '结核性胸膜炎');
INSERT INTO `t_symptom_di` VALUES ('2192', '低烧', '367', '结核性胸膜炎');
```

医疗诊断专家系统

医疗诊断专家系统的结构

医疗诊断专家系统主要由知识库、数据库、推理引擎、解释设备和用户界面等组成。系统结构如下：



采用数据推动的前向链接推理技术，患者通过提供自己的症状情况，专家系统对输入的疾病信息进行病症分解，抽取病症特征，进一步将问题转换成规则的描述形式（形如：RULE: IF …, THEN …），针对待决策的病症的特征，采用深度优先策略搜索知识库，同时在推理引擎中进行病症的推理和匹配。如果是不在范围内的病症，那么就不做处理，视为新病症，专门存储在新病症数据库中以待更新。

规则的构建和表示

由于疾病数据库过于庞杂，我只展示其中一部分规则：

```
{
  IF: ['集中式丘疹', '非病毒感染', '0']
  THEN: '集中式丘疹, 药物或过敏性皮疹'
  DESCRIPTION: '药物或者过敏源导致, 俗称荨麻疹'
}
```

```

{
    IF: ['集中式不规则疹','非病毒感染','0']
    THEN: '集中式不规则疹,药物或过敏性皮疹'
    DESCRIPTION: '药物或者过敏源导致,俗称荨麻疹'
}
{
    IF: ['四周型不规则疹','非病毒感染','0']
    THEN: '四周型不规则疹,药物或过敏性皮疹'
    DESCRIPTION: '药物或者过敏源导致,俗称荨麻疹'
}
{
    IF: ['四周型丘疹','非病毒感染','0']
    THEN: '四周型丘疹,药物或过敏性皮疹'
    DESCRIPTION: '药物或者过敏源导致,俗称荨麻疹'
}
{
    IF: ['集中式不规则疹','低热','1~']
    THEN: '集中式不规则疹,其他感染性皮肤病'
    DESCRIPTION: '由其他细菌或病毒感染所致'
}
{
    IF: ['集中式丘疹','低热','1~']
    THEN: '集中式丘疹,其他感染性皮肤病'
    DESCRIPTION: '由其他细菌或病毒感染所致'
}
{
    IF: ['幼儿','集中式丘疹','低热','<1']
    THEN: '幼儿风疹'
    DESCRIPTION: '由风疹病毒引起的急性传染性疾病'
}
.....

```

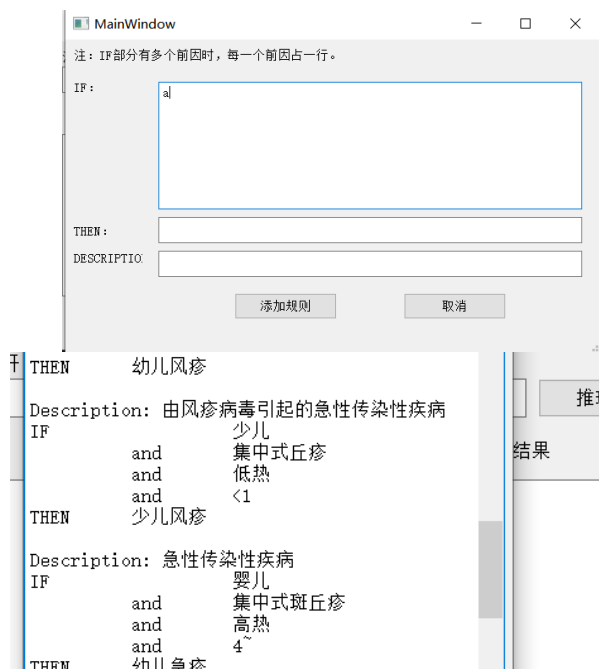
规则是将病症作为条件，将对应的可能病因作为结果，同时附加描述。

算法描述与分析

专家系统的核心算法是在症状的推理过程中，案例匹配采用最近邻法对病症案例的相似度进行计算，计算结果从高到低排序，相似度取值最高者作为疾病最佳匹配结果。

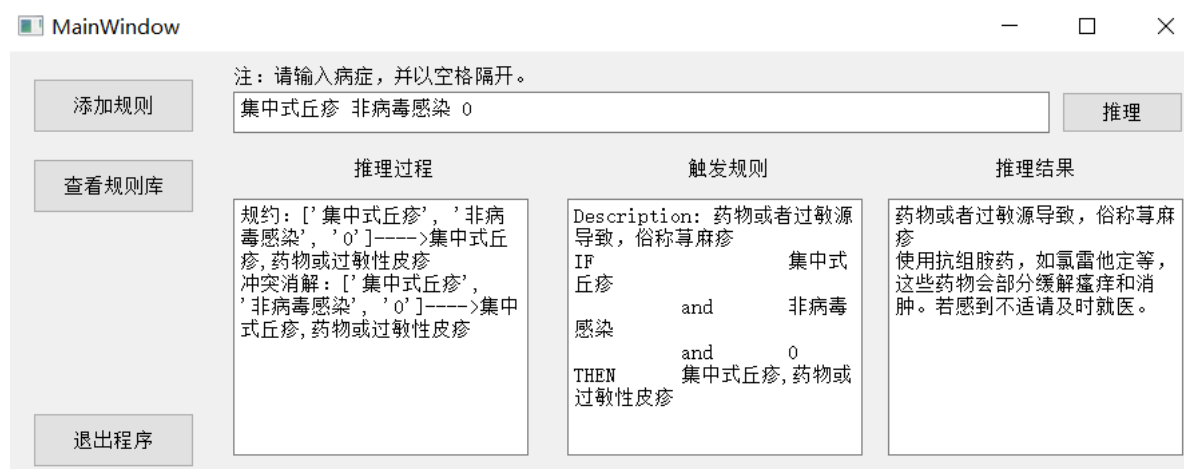
用户界面





使用python的模块 Pyqt5 实现了简易的可视化界面，左边添加规则可以更新知识库，新增 IF ... THEN ... 条件，查看规则库可以显示出来现有的规则。

测试用例



如果不能检索到相应病症，会进行提示：



意义与不足

本项目构建的常见疾病诊断的专家系统在已有医学知识库的支持下，以最近邻算法为基础进行病症相似度的匹配计算，已给出较为准确的诊断结果并推荐相应的治疗方案。该系统首先可以大量节约医疗资源，使得病患可以有目标地选择科室诊治，同时在紧急情况下可以及时处理，并且该专家系统引入智能交互界面，提供了友好的人机交流，对于基础医疗知识的普及有着特别重要的意义，并能为医院解决患者就医难、等待时间长等问题提供一定的帮助。

尽管测试案例全部通过，但是由于知识库规模的有限性以及规则的局限性，所以不能保证诊断结果完全准确。