

# 限定领域智能导学系统问题生成及对话管理技术<sup>①</sup>

商雄伟, 张志祥

(海军工程大学, 武汉 430033)

**摘要:** 智能导学是当前信息化教学研究的一个热点领域, 将导学技术应用到难于识记并需要严格遵守的规章制度类知识领域, 将极大地提高领域人员对相关知识的快速理解与掌握能力. 采用面向对象的知识表示方法, 提高了对通用主题和问题模板表示的灵活性. 设计了基于模板的领域知识表示方法和导学问题生成方法, 可实现一个主题知识生成多种问句. 采用约简的有限状态转移方法, 避免了对话管理中状态数量庞大的问题. 最后, 本文以船舶损管领域为例, 将此方法运用到船舶损管训练语音导学原型系统中. 系统运行良好, 充分表明这种设计方法是切实可行的.

**关键词:** 导学系统; 对话管理; 自然语言生成; 导学问题生成; 损管训练

## Interrogative Sentence Generation and Dialogue Management for Intelligent Tutoring System in Limited Domain

SHANG Xiong-Wei, ZHANG Zhi-Xiang

(Naval University of Engineering, Wuhan 430033, China)

**Abstract:** Intelligent tutoring is a hot field in the study of information-based teaching. Applying this technology to knowledge of regulations and operational procedures that are expected to be remembered accurately and obeyed strictly, will improve the ability of grasping and using related knowledge. Using the method of object oriented, improve greatly the flexibility of general theme representation and question template. Template-based methods of domain knowledge representation and tutoring questions generation are all presented. It can generate multiple tutoring questions for different attributes. The reduced state transition method can avoid efficiently the problem of state explosion. Finally, taking shipboard damage control as one application background, a prototype system was developed and worked well using this method, illustrating the feasibility and practicality of the design method.

**Key words:** tutoring system; dialogue management; natural language generation; interrogative sentence generation; damage control

智能导学系统 (ITS) 是一种基于计算机的教学系统, 力图模仿自然人的教学方法和对话方式, 与学生进行及时的、按需的交互<sup>[1]</sup>. ITS 可以回答学生的问题、对学生的动作进行反馈并提供个性化的指导. 现实生活中, 大部分的操作规程、处置方案等文案, 比如:《煤矿安全规定》、《建设工程安全生产管理条例》、《舰艇条例》等, 这种规章、法规类文案通常有以下特点: (1) 文字表述科学、严谨、简洁, 一般不允许随意使用替代词; (2) 语句组成形式有一定规律性; (3) 重要内容

需要领域人员准确识记并严格遵守. 由于文案通常比较枯燥, 准确理解识记此类知识存在一定难度, 需要用户反复学习理解, 耗费大量的时间与精力, 设计开发规章制度类领域导学系统有着重要意义.

近年来, ITS 在特定领域教学中应用越来越广泛, 美国斯坦福大学 Pon-Barry 等人开发了智能对话损管训练系统 SCoT, 将 ITS 应用到船舶损管教学领域<sup>[2]</sup>; A. C. Graesser 等人设计开发 AutoTutor<sup>[3]</sup>, 东北师范大学赵蔚等开发出自适应教育超媒体系统, 将 ITS 应用到

<sup>①</sup> 收稿时间:2015-03-09;收到修改稿时间:2015-05-12

计算机教学领域。在面向特定领域的智能导学系统中有两个关键问题：一是导学问题生成；二是对话流程管理<sup>[4]</sup>。人工智能中的自然语言生成(Natural Language Generating, NLG)是自然语言处理的两大领域之一。对话过程中的自然语言生成实际上是在确定“说什么”后选择“如何说”的问题。近年来,在许多应用系统中都运用了基于 NLG 的文本自动生成技术,虽然在文档规划、微观规划及表层实现上有一定的差别,但是所用技术并无太大差别。开发人员一般会以生成文本所要表达能力的强弱为标准,选择不同的技术来实现。目前常用技术有以下几种:模板生成技术(Template)、管道技术(Pipeline approaches)、规划技术(Planning approaches)和统计技术(Statistical approaches)<sup>[5]</sup>。对话管理在对话系统中处于核心地位,其设计优秀与否关系到整个对话系统性能。对话管理的任务是控制对话流程,帮助用户高效自然地完成对话。对话管理根据其所提供的控制类型可以分为 3 类:系统主导,用户主导和混合主导<sup>[6]</sup>。在对话过程中,用户的回答或提问可能是含糊不清或者是不完整的,对话管理必须引导用户说明自己的意图,并提供完成任务所需要的信息。为了完成和用户的交互行为,对话管理应该记录对话历史,并根据对话上下文正确理解用户输入,产生合适的对话反馈。

## 1 领域知识的表示

知识的表示就是对知识的一种描述,将其进行符号化、形式化的过程,即通过特定的形式把知识表示为计算机可识别的数据的数据结构<sup>[7]</sup>。知识表示是构建知识库的关键,知识表示方法是否合理不仅关系到知识库中知识的有效存储,而且也直接影响着系统的知识推理效率和对新知识的获取能力<sup>[8]</sup>。知识表示方法主要有谓词逻辑表示法、产生式表示法、框架表示法、语义网络表示法、面向对象知识表示法等。

现代汉语中一个完整简单的句子成分主要有六种,即主语、谓语、宾语、定语、状语和补语。我们通常使用的简单句由六种成分中的一种或若干种组成。结合规章制度类知识特点,规定一个简单句为一个主题,一个复杂句由若干主题组成。一个主题由若干属性组成,在不同的规章类知识领域,可以划分为不同属性,如条件属性(condition),主语属性(subject),时间属性(time)等等,具体属性划分需根据规章制度类知识的内

容和组织结构而定。每一属性可以进一步划分为五种类型:OR, AND, SEQ, UNESS 和 UEXIS。如果一个属性的类型为 OR,说明该属性对应的答案有若干可选项;如果一个属性的类型为 SEQ,说明该属性对应的答案有若干个有序项;AND 类型与 SEQ 类型类似,只不过不要求答案项的有序性;如果一个属性的类型为 UNESS,说明该属性为非考核性知识,在后续的考核问题语句生成中,不针对该属性生成专门的疑问句。如果在某个主题中,模板中某一属性不存在,该属性类型记为 UEXIS。对于每个主题,可以针对除 UNESS 类型外各属性生成疑问句,向用户提问。

结合上文规章制度类领域知识特点,采用面向对象的知识表示方法,知识表示类图如图 1 所示。

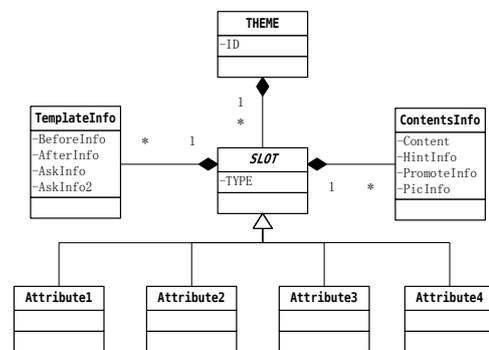


图 1 知识表示类图

按照上图知识表示方法,主题表示如下:

THEME :=(Attribute1, Attribute2, Attribute3,.....)

其中每个属性值包含三类信息。一是属性类型信息 TYPE,即上文中提到的五种类型中的一种;二是内容提示类信息 PromptInfo,主要包括暗示性信息(HintInfo)、提示性信息(PromptInfo)和多媒体提示信息(PicInfo)等属性,主要对用户进行导学;三是问句生成模板辅助信息 TemplateInfo,主要包括属性前件信息(BeforeInfo)、属性后件信息(AfterInfo)、属性问题信息(AskInfo1 和 AskInfo2),用于生成导学问题。以损管领域主题“主机舱起油火(condition),船舶消防队(subject)可以用卤代烷灭火器、泡沫灭火器或者二氧化碳灭火器(tool)灭火(action)”为例。上述主题中“tool”的“BeforeInfo”为“使用”,“AfterInfo”为“NULL”,“AskInfo1”为“使用什么工具”,“AskInfo2”为“还应使用什么工具”。在针对“tool”属性生成问句时,如果针对主题中“tool”以外的其它属性提问,使用“BeforeInfo +

tool + AfterInfo”,即可生成“...使用 tool...”;当针对“tool”属性提问时,如果用户未回答出任何答案则使用“AskInfo1”。当用户已回答出部分答案,则使用“AskInfo 2”。具体使用方法下文会详细阐述。由于采用了面向对象的知识表示方法,通过更改相关属性值信息,即快速移植到不同的规章制度类领域。

相对于传统 ITS 导学材料库中大量的<Question,Answer>一对一匹配对知识<sup>[9]</sup>,通过采用上述模板的知识表示方法,系统可以针对一个主题中的不同属性生成不同的问句,即一个主题可以生成多种导学问题,节省了领域知识库开发设计所需要的时间,提高了开发效率。

## 2 导学问题生成

基于模板的自然语言生成技术需要手工创建模板,生成的文本不够灵活,但其思路简单、工作效率高、较容易实现,考虑到规章制度等限定领域应用实际,我们仍采用基于模板的自然语言生成技术<sup>[10]</sup>。根据上文主题<Attribute1,Attribute2,Attribute3,.....>表示形式,针对不同属性设置不同模板,系统可以根据主题若干属性自动生成  $n(n \geq 1)$ 种考核问句,其中,当某一属性类型为 UEXIS 时,则在问句中不需要生成相应的属性部分,当针对某一属性提问时,使用该属性的 AskInfo1/AskInfo2 值和其他属性的 BeforeInfo、Content、AfterInfo 值即可组成导学问题。

结合前面的知识表示方法,以 Attribute2 属性为例,问题模板设置如下:

Template\_Att2-1: Attribute1.BeforeInfo + Attribute1.Content + Attribute1.AfterInfo + Attribute2.AskInfo1 + .....

Template\_Att2-2: “除了<A>”, +Attribute1.BeforeInfo + Attribute1.Content + Attribute1.AfterInfo + Attribute2.AskInfo2 + .....

在对话系统中,系统问题的生成还要考虑对话的上下文,根据具体情况匹配相应的模板,如果用户首次作答或未给出任何有效答案,使用 Template\_Att2-1,如果用户答案不完整,使用 Template\_Att2-2,并根据用户回答情况,适时调整模板中的属性<A>。

如果属性 Attributei 类型为 SEQ,其参考答案为长度为  $n$  的有序表  $\{S_1, S_2, \dots, S_n\}$ ,如果用户已经正确回答了前  $k-1$  项,即用户的回答  $\{A_1, A_2, \dots, A_m\}$  中前  $k-1$  项与参考答案相同,第  $k$  项出现错误,即  $k = \min\{i | S_i \neq A_i\}$ ,为了使用户完整地解决问题,需要生成疑问句对用户提问,

针对 Attributei 的导学问题如下:“除了<A>,还要.....”。其中  $A = \{S_1, S_2, \dots, S_{k-1}\}$ 。

如果属性类型为 AND,其参考答案为长度为  $n$  的无序表  $\{S_1, S_2, \dots, S_n\}$ ,如果用户已经正确回答若干项,为了使用户完整地解决问题,针对 Attributei 的疑问句如下:“除了<A>,还要.....”。其中  $A = \{A_i | A_i \in \{S_1, S_2, \dots, S_n\}\}$ 。

## 3 语音合成

语音合成(Text To Speech, 简称 TTS),是将文字信息转换为自然语音的一项技术,在人机语音交流、文字信息处理领域有着广泛的应用。

微软 Speech SDK 是微软视窗环境下的大型语音开发工具包,含有语音应用设计接口(Speech API)、连续语音识别引擎(MSCR)以及语音听写引擎(TTS)等,提供了强大的语音识别和语音合成功能,它采用 COM 标准开发,底层协议都以 COM 组件的形式完全独立于应用程序层,为应用程序设计人员屏蔽掉复杂的语音技术,语音识别由语音识别引擎管理,语音合成由语音合成引擎管理负责,开发人员只需调用相关的语音应用程序接口来实现语音识别与合成功能。

## 4 答案评判

当用户输入答案  $\{A_1, A_2, \dots, A_n\}$  后,系统通过评判答案  $A_i$  的正确性,了解目前用户知识掌握情况,便于在对话管理模块生成基于上下文的对话,为此,定义答案评判函数  $F$ ,用于判断用户答案  $A_i$  的正确性,当  $A_i$  为正确答案时,  $F(A_i)$  取 1; 反之,  $F(A_i)$  取 0。标准答案主要有 SEQ、AND、OR 三种类型,不同类型答案评判标准不同,对于 SEQ 类型答案,当用户前  $k-1$  项答案正确,第  $k$  项答案错误,第  $k+1 \sim n$  答案正确性未知,规定前  $k-1$  项为正确答案,第  $k \sim n$  项为错误答案,对于 AND 和 OR 类型答案,当  $A_i \in \{S_1, \dots, S_n\}$ ,为正确答案,否则为错误答案,在一个主题的  $n(n \geq 1)$  轮对话中,由于用户一直在完善答案,所以需要每一轮对话中都更新  $F$  函数值。

表 1 评判函数更新表

(更新前)					(更新后)				
S	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>	S	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>4</sub>
A	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>		A	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	
F	1	0	0	0	F	1	1	1	0
问句	除了 A <sub>1</sub> 之外,还要使用什么器材?				问句	除了 A <sub>2</sub> 、A <sub>3</sub> 之外,还要使用什么器材?			

下面以 SEQ 类型某属性为例,说明该评判函数.如表 1 中所示.标准答案为  $\{S_1, S_2, S_3, S_4\}$ , 用户第一轮对话中答案为  $\{A_1, A_2, A_3\}$ . 当  $A_1$  正确,  $A_2$  错误时,  $F$  函数取值为  $\{1, 0, 0, 0\}$ . 系统生成问句“除了  $A_1$  之外, 还要执行什么动作?”. 在第二轮对话中, 用户补充答案  $\{A_2, A_3, A_4\}$ . 当  $A_2, A_3$  回答正确时, 更新  $F$  函数取值为  $\{1, 1, 1, 0\}$ . 系统生成问句“除了  $A_2, A_3$  之外, 还要执行什么动作?”. 以此类推.

## 5 对话管理

导学过程中对话管理过程通常有以下三个步骤, 对话管理流程如图 2 所示.

STEP1: 生成问句

STEP2: 用户输入答案信息

STEP3: 系统评判, 对于用户不完整/不正确答案给出提示信息, 引导用户朝着正确的方向作答, 并返回 STEP1<sup>[11]</sup>

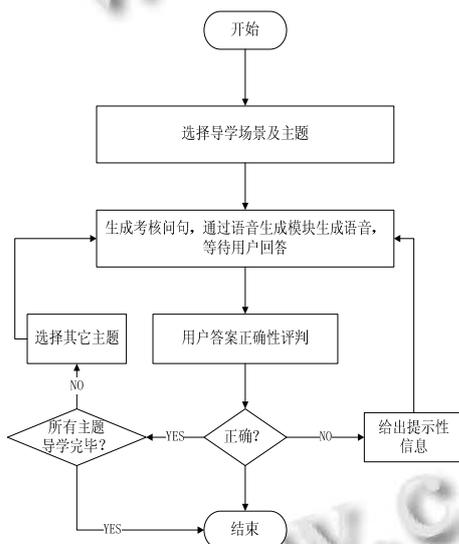


图 2 面向问答的对话管理流程

对话管理过程的重点是根据用户当前输入信息, 给出合适的反馈信息. 对话管理的方法主要有基于状态图的结构、填充槽结构和基于任务的结构三种. 采用基于有限状态图方法, 处理速度快, 实现简单. 比较适合于相对简单的对话过程. 但当对话过程变得复杂而庞大时, 状态转移网络中的状态节点会成倍增加, 变得难以控制. 而用户的语句偏离系统计划, 使得系统无法处理的可能性也大大增加. 例如, 对一个 SEQ 类型标准答案有  $n$  个  $(S_1, S_2, \dots, S_n)$  的主题来说, 设每个项

有 2 种可能的用户输入. 对于标准的有限状态模型, 该主题有  $2^n$  个状态, 状态数量庞大.

为了避免状态爆炸, 本文通过计算初始错误位置算法, 求出导致用户答案错误的首要因素, 对状态进行约简合并. 对于一个标准答案  $(S_1, S_2, \dots, S_n)$  的主题, 我们定义状态  $\text{AnsT}_i$  表示用户第  $1 \sim i-1$  个答案正确, 第  $i$  个答案错误, 第  $i+1 \sim n$  个答案正确性未知. 那么系统状态将约简为  $n+1$  种. 该算法如下:

```

K=1, N=1;
while(K>0)
{
    获取语音输入信息, 并进行语音识别, 分析结果得到答案
     $A_1, A_2, \dots, A_m$ 
     $K = \text{ErrorIndex}(\text{type}, S_i, S_{i+1}, \dots, S_n, A_1, A_2, \dots, A_m)$ 
    if(K>0)
        语音输出第 K 项对应的提示信息
        生成第 N++ 轮问句
    End if
}

```

该算法的关键是  $\text{ErrorIndex}$  函数. 该函数根据题目类型,  $S_i, S_{i+1}, \dots, S_n$  以及用户答案  $A_1, A_2, \dots, A_m$  找出第一个错误答案编号并返回; 如果答案正确, 返回 0. 对于 SEQ 和 AND 类型题目,  $\text{ErrorIndex}$  有如下处理方法:

```

if(type==AND)
     $K = \min\{i | S_i \notin \{A_1, \dots, A_n\}\}$ 
else if(type==SEQ)
     $K = \min\{i | S_i \neq A_i\}$ , 其中  $i \geq \min\{j | F(A_j) = 0\}$ 

```

## 6 实验与结论

基于上述方法, 我们以船舶损管训练领域为例, 在 Visual Studio 2010 平台上开发出原型系统, 并进行了实验. 我们从答案评判准确度、问句生成准确度、提示信息准确度三个方面对系统进行测试. 在对话过程中, 如果系统能够和真实老师一样判断答案正误, 按照预定义好的知识库给出提示信息并且生成下一轮问句, 我们则认为系统响应正确, 否则系统响应错误. 在实验室环境下, 采集了 4 个人, 共进行了 4 轮对话, 根据题目设置, 我们给每个人分配了不同的回答方式. 从表 2 中我们可以清楚的看到, 基于本文的设计方法, 测试集总的正确率均达到 94% 以上, 这样的结果还是比较令人满意的. 分析测试数据, 导致系统响应错误原因主要有以下几点: 语音数据不清晰; 噪音影响干扰; 对用户输入的信息分词发生错误; 领域知识库知识表示不完整.

表2 正确率统计表

编号	答案评判	问句生成	提示信息
	正确/错误	正确/错误	正确/错误
1	10/0	10/1	6/0
2	15/1	15/1	10/0
3	13/0	13/0	10/0
4	13/0	13/1	8/1
正确率	98.08%	94.44%	97.14%

下面是一个比较典型的对话片段。

**系统 S1:** 主机舱起油火。船舶消防队应该用什么灭火器进行灭火?

**用户 U1:** 卤代烷灭火器

**系统 S2:** 回答正确

**系统 S3:** 本题错误 0 次

**系统 S4:** 手持式灭火器没能扑灭。您首先应怎么办?

**用户 U2:** 应该封舱灭火

**系统 S5:** 回答正确

**系统 S6:** 本题错误 0 次

**系统 S7:** 封舱灭火有哪些步骤?

**用户 U3:** 我不知道

**系统 S8:** 影响船舶封舱灭火效果的决定性因素是舱内温度和氧气浓度。

首要措施应该尽可能降低船舶舱室氧气浓度。封舱灭火有哪些步骤?

**用户 U4:** 关通风、人员疏散

**系统 S9:** 第一步正确。除了降低氧气浓度外,还应切断火源。除了关通风外,封舱灭火还要做什么?

**用户 U5:** 关油阀。关舱门

**系统 S10:** 回答正确。

**系统 S11:** 本题错误 1 次

从上述对话可以看出,在每轮对话中,系统首先判断用户答案的正确性,根据用户当前回答情况,给出提示性信息,并继续生成下一轮问题,逐步引导用户给出正确的答案。在用户给出正确回答后,统计错误次数。比如在“系统 S9”中有三部分内容:答案评判(“第一步正确”)、提示(“应切断火源”)和问句生成(“除了关通风外,封舱灭火还要做什么?”)。从对话结果来看,系统基本可以正常进行语音问答,准确地对用户答案进行评判,给出用户提示信息,根据用户答案自动生成导学问题。系统基本达到了设计要求。

## 7 结语

导学问题生成和对话流程管理是 ITS 的重要研究内容,本文设计了基于模板的领域知识表示方法和导学问题生成方法,实现了针对一个主题不同属性生成多个问句。采用面向对象的知识表示方法和基于状态约简的导学对话管理策略,方便在不同领域个性化设置问题生成模板,使导学对话管理过程简单实用,对

于规章制度类限定领域的导学系统有着一定借鉴意义。由于问题生成模块、对话管理模块与领域知识模块的相互独立性,确保了本系统具有一定的可扩展性。但是基于模板的导学问题生成方法受模板影响较大,在模板设置不合理时,所生成的文本质量不高、难以满足使用需求。如何生成更高质量的自然语言文本,也是我们下一步的研究重点。

## 参考文献

- 1 Latham A. Personalising Learning with Dynamic Prediction and Adaptation to Learning Styles in a Conversational Intelligent Tutoring System[Ph.D. Thesis]. Manchester:the Manchester Metropolitan University,2011.
- 2 Pon-Barry H, Clark B, Schultz, et al. Contextualizing reflective dialogue in a spoken conversational tutor. Educational Technology & Society, 2005, 8(4): 42-51.
- 3 Graesser AC, Olney A, et al. AutoTutor: an intelligent tutoring system with mixed-initiative dialogue. IEEE Trans. on Education, 2005, 48(4): 612-618.
- 4 Lemon O. Learning what to say and how to say it: joint optimization of spoken dialogue management and natural language generation. Computer Speech and Language, 2011, 25(11): 210-221.
- 5 Srinivasan CJ. Learning User Modelling Strategies for Adaptive Referring Expression Generation in Spoken Dialogue Systems[Ph.D. Thesis], University of Edinburgh, 2011.
- 6 王菁华,钟义信,王枫等.口语对话管理综述.计算机应用研究,2005,10(4):5-8.
- 7 彭丹华.基于知识推理的武器装备运用教学仿真系统研究[硕士学位论文].长沙:国防科学技术大学,2012.
- 8 刘建伟,燕路峰.知识表示方法比较.计算机系统应用,2010, 20(3):242-246.
- 9 Norzaidah Md Noh, Ahmad A, Halim SA, et al. Intelligent tutoring system using rule-based and case-based: a comparison. The 3rd International Conference on e-Learning ICEL. Bandung. Elsevier Ltd. 2011. 454-463.
- 10 Oh AH. Stochastic natural language generation for spoken dialog systems. Computer Speech and Language, 2002, (16): 387-407.
- 11 Graesser AC. AutoTutor: A simulation of a human tutor. Journal of Cognitive Systems Research, 1999, 1(17): 35-51.