

智能问答系统实践

第五课扩展：人机对话



姜文斌

北京师范大学人工智能学院

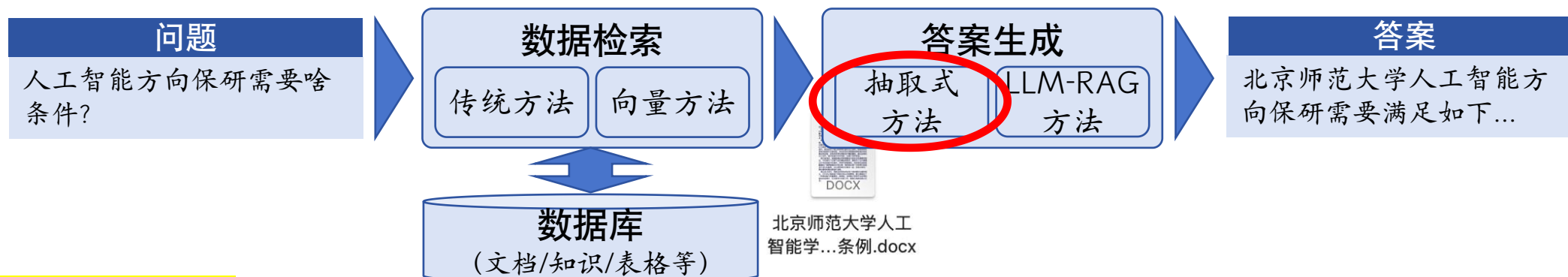
2025.03.27

我的位置



智能问答系统

针对用户提出的自然语言问题，从数据库中检索相关信息，并依据相关信息作出回答



智能问答线上处理流程

智能问答线下处理模块

问答数据库构建

基于传统方法的数据建库

基于向量方法的数据建库

数据检索模块构建

传统语义匹配模型构建

向量语义匹配模型构建

答案生成模块构建

抽取式答案生成模型构建

RAG式答案生成模型构建

效果评估模块构建

文档检索效果评估

问答整体效果评估

目录



- 人机对话
- 对话系统流程
- 对话系统发展
- 总结

人机交互



- 人机交互：人类使用机器时，需要给机器输入**指令**，并接收机器的**反馈**



按键，拨片，摇杆，旋钮，语音，体感...

指令

指示灯，声音，图像，各操作件状态...

反馈



人机交互



传统机器

按键, 拨片, 摇杆...

机器设备

声音, 指示灯, 状态...



计算机

键盘, 鼠标, 触摸板/屏...

计算机

声音, 图像, 视频...



智能机器

语言, 动作...

机器人

语言, 动作...



人机对话

■ 人机对话：以人类语言对话的方式，实现信息获取或完成特定任务

■ 任务型对话：最为经典的对话形态，应用也最为广泛

■ 问答 / 推荐：都可以视为一种广义上的人机对话形态

	任务型 Task	聊天 Chat	知识问答 Knowledge	推荐 Recommendation
目的	完成任务或动作	闲聊	知识获取	信息推荐
领域	特定域（垂类）	开放域	开放域	特定域
以话轮数评价	越少越好	话轮越多越好	越少越好	越少越好
应用	虚拟个人助理	娱乐、情感陪护	客服、教育	个性化推荐
典型系统	Siri、Cortana、Allo、度秘、灵犀	小冰	Watson、Wolfram Alpha	Quartz/今日头条

智能硬件



无人机
智能家居

汽车



车载语音
智能汽车

个人助理



订餐
出行

咨询类机器人



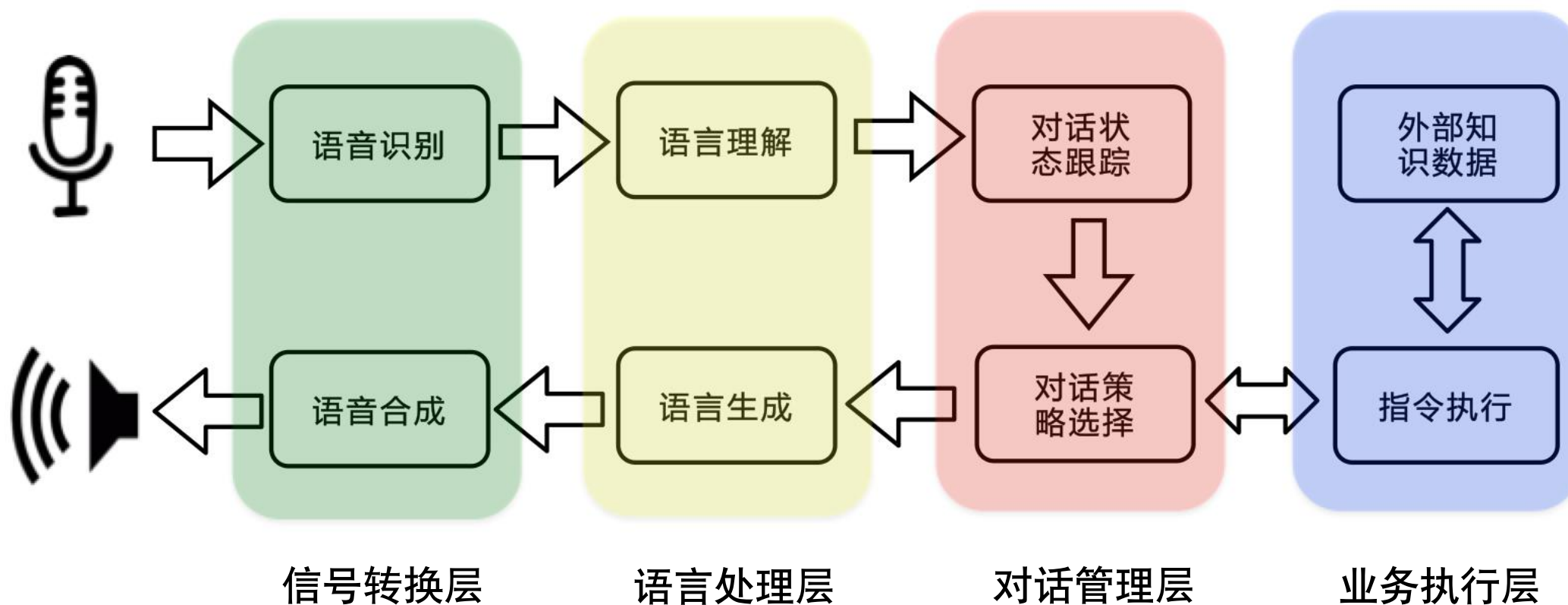
导诊机器人
电商客服

目录



- 人机对话
- 对话系统流程
- 对话系统发展
- 总结

对话流程



对话示例



U: 清华大学有多远

B: 起点是哪里

- 理解用户询问清华大学的距离，但未提供起点位置，需要向用户询问

对话理解

U: 北京师范大学

B: 大约7公里

- 理解用户给出了起点位置，执行查询两地距离的指令，并反馈给用户

对话管理

U: 怎样才能过去

B: 公交还是驾车

- 理解用户询问如何从起点到终点，但未提供出行方式，需要向用户询问

指令执行

U: 公交

B: 地铁12号线人民大学换4号线

- 理解用户给出了出行方式，执行查询导航路径的指令，并反馈给用户

语言生成

对话理解



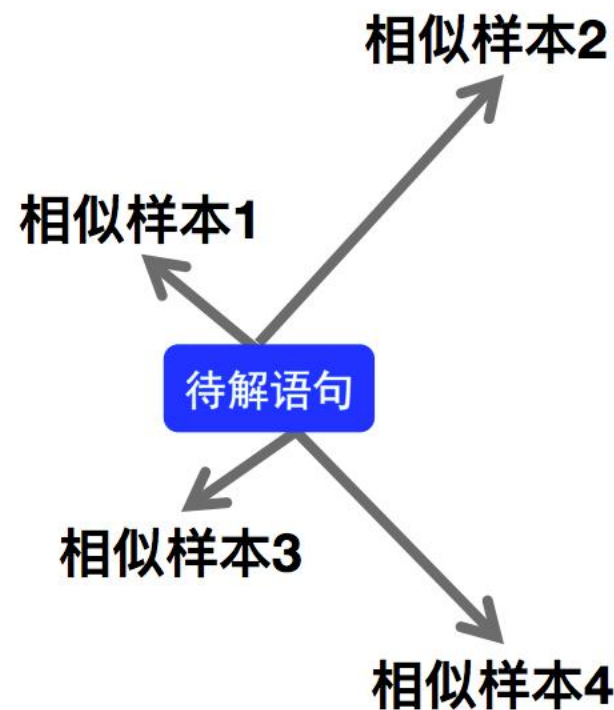
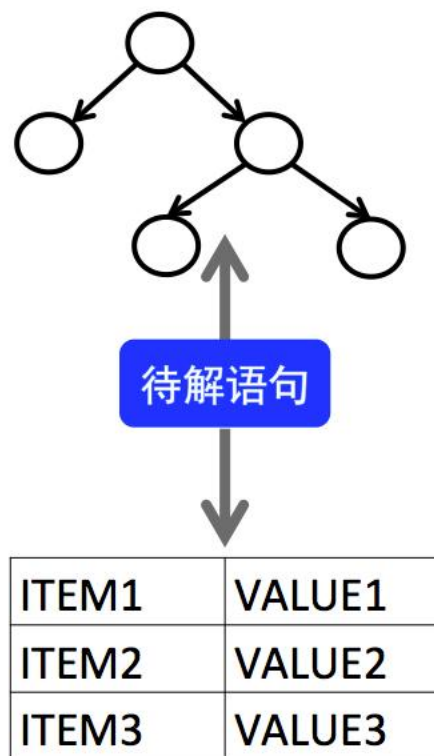
■ 理解用户请求所包含的语义信息，用于信息查询或指令执行

■ 任务难点

- 自然语言歧义性
- 表达方法多样性
- 叙述风格口语化

■ 对话理解的经典模式

- 基于语义解析的对话理解
- 基于语义匹配的对话理解



对话理解



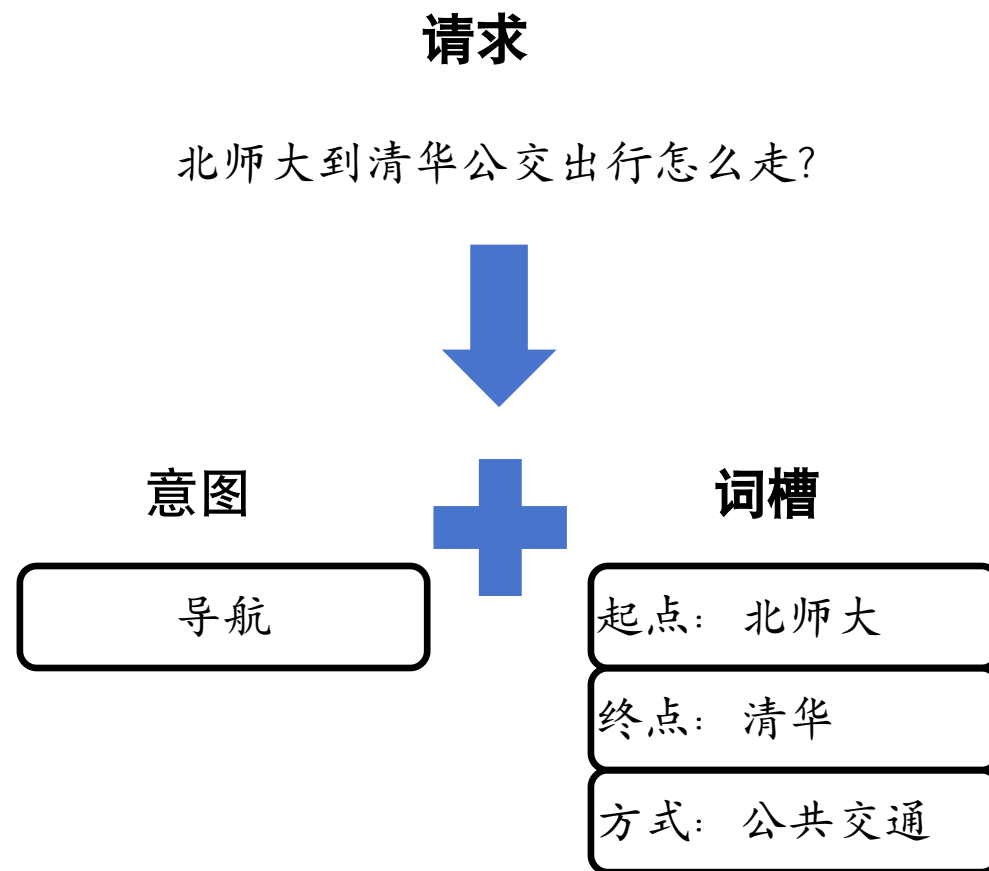
■ 将用户请求解析为所包含语义信息的结构化表达

■ 示例：意图-词槽模式

- 意图：描述用户的核心诉求
- 词槽：描述意图的关键信息

■ 语义解析的常用方法

- 基于知识规则的方法
- 基于机器学习的方法
- 基于融合策略的方法



对话管理

■ 基于对话状态实施对话策略，从而实现多轮对话逻辑

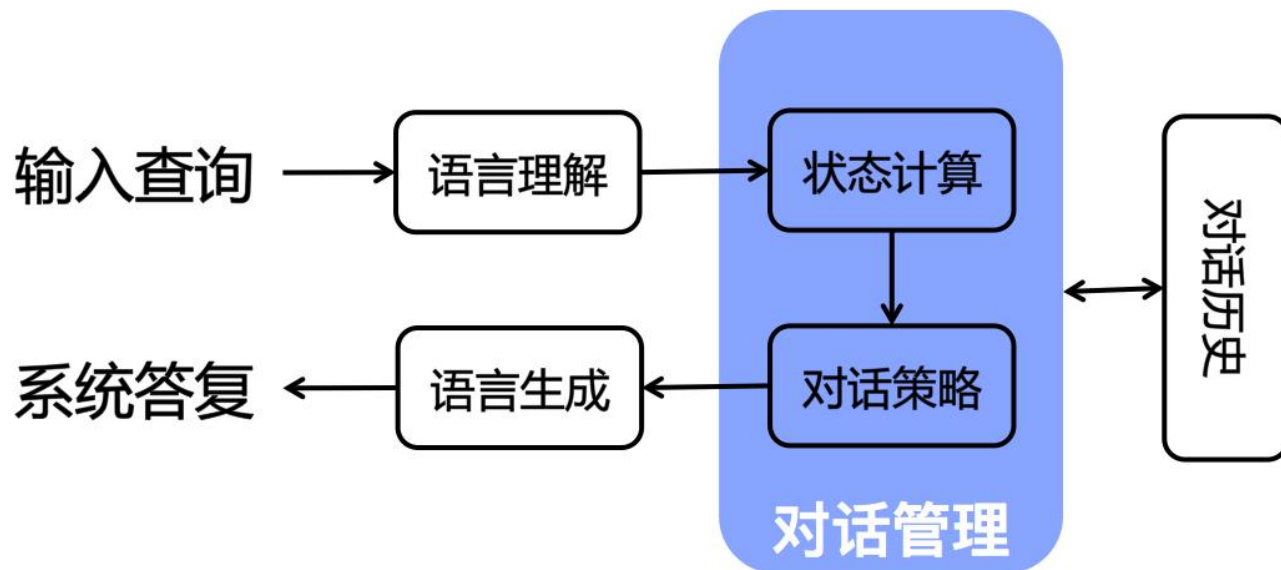
- 通过多轮对话逻辑支撑业务逻辑，处理歧义纠正错误

■ 任务难点

- 状态计算的不确定性
- 不确定环境下的决策

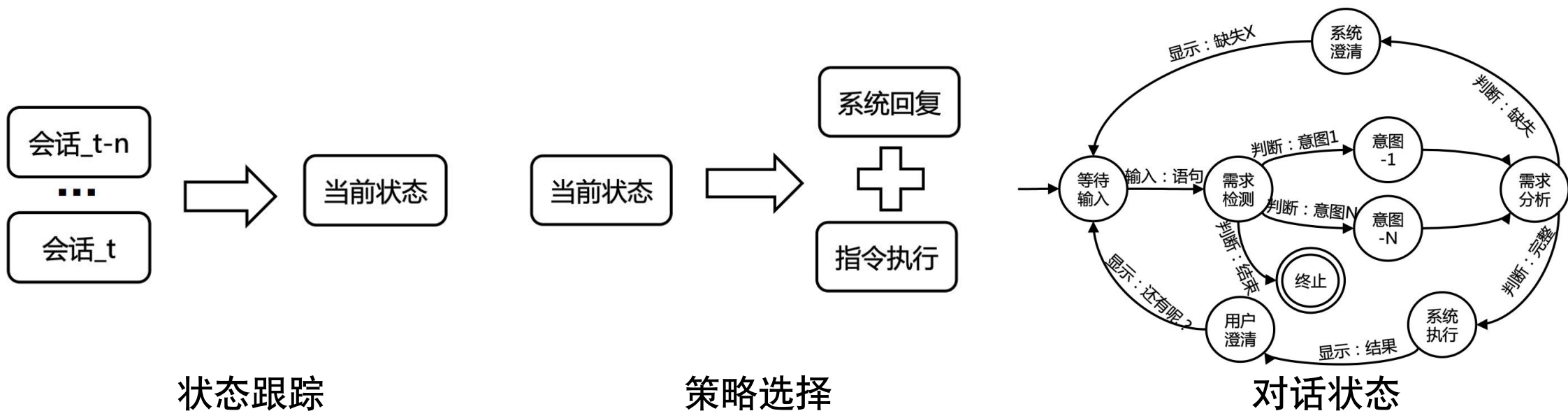
■ 对话管理的核心任务

- 对话状态跟踪
- 对话策略选择



对话管理

- 对话状态：反映当前对话进展的语义状态，用作系统回复的依据
- 状态跟踪：根据对话历史计算当前对话状态，管理并更新对话历史
- 策略选择：根据当前对话状态，选择接下来最恰当的操作



目录



- 人机对话
- 对话系统流程
- 对话系统发展
- 总结

发展历程



人工模板规则

■ 优点

- 快速启动
- 便于理解与优化

■ 缺点

- 需要专业知识
- 难以跨领域
- 鲁棒性难提升

传统机器学习

■ 优点

- 无需手工规则
- 鲁棒性较好

■ 缺点

- 需要设计特征模板
- 需要较大语料库
- 较难理解与优化

深度机器学习

■ 优点

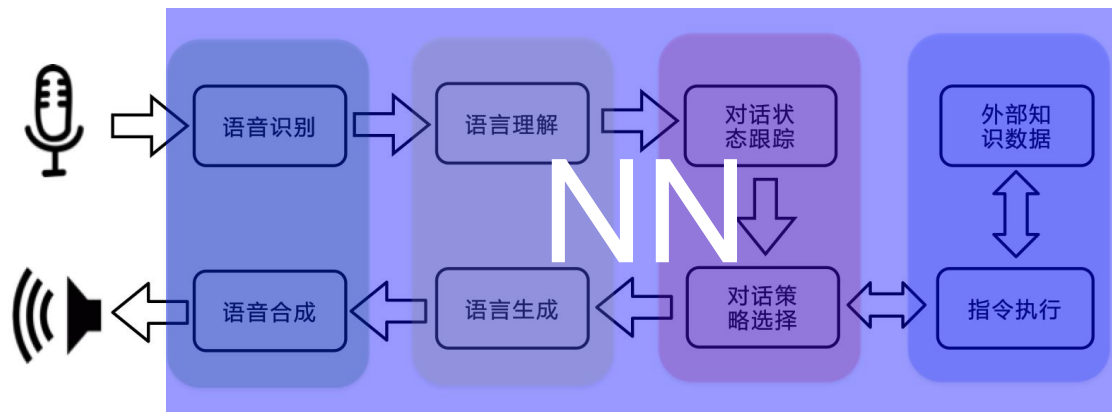
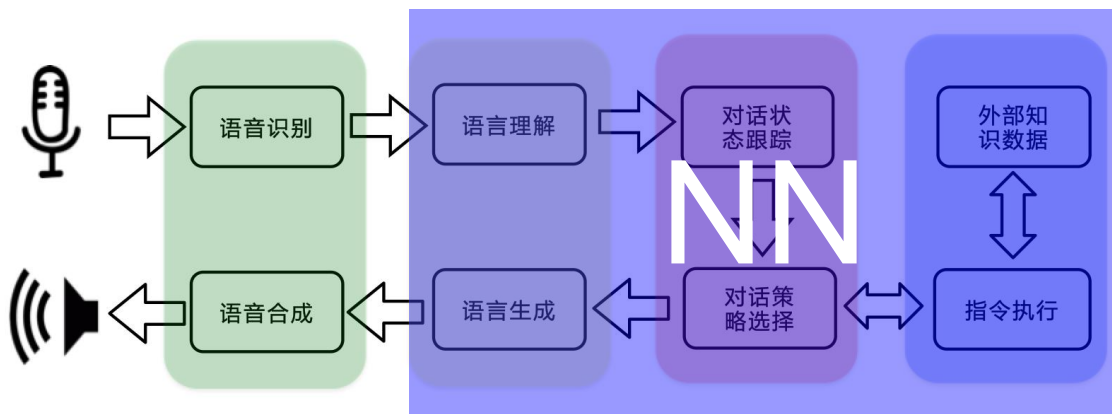
- 不需设计特征模板
- 端到端简单高效

■ 缺点

- 需要更大语料库
- 更难理解与优化

端到端建模

- 端到端的对话系统，即将特定的处理链条用统一的神经网络建模
 - 将语音处理之外的处理链条用统一神经网络建模
 - 将语音处理纳入统一模型能够提供更实时更共情的对话

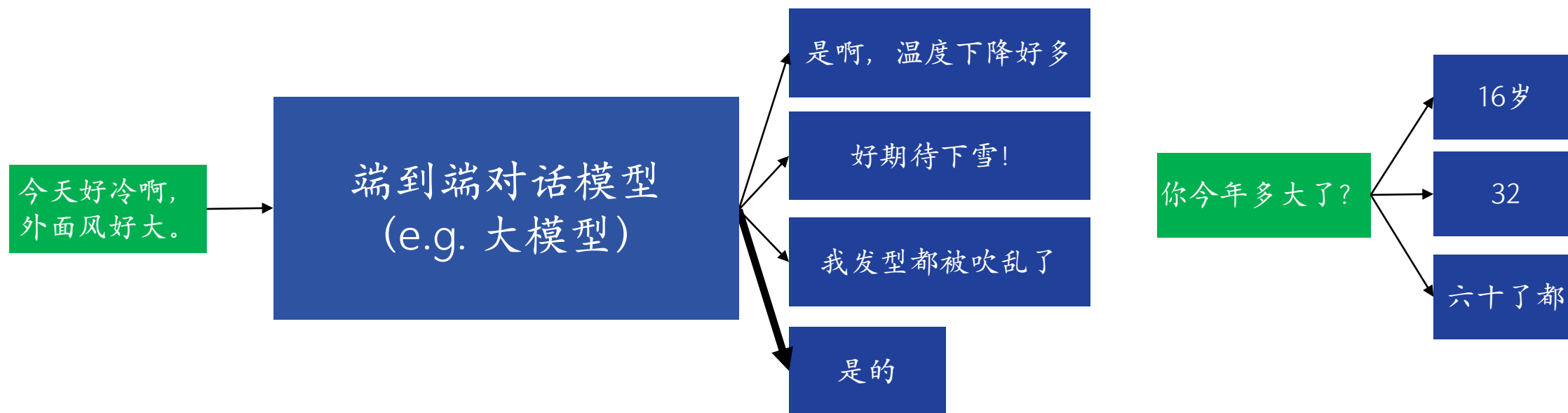


端到端建模



■ 端到端的对话系统：以编码器-解码器方式，类似于机器翻译的方式运行

■ 问题：同一句上文可能存在多样化的下文，标准端到端模型倾向于产生安全回复



模型一对一的设定导致被较高概率的安全回复主导

背景知识缺乏导致回复不可控

目录



- 人机对话
- 对话系统流程
- 对话系统发展
- 总结

对话系统



■ 机器满足人类以语言的方式获取信息或服务的需求

- 任务型对话：自然语言，获取信息（天气、问答等）或服务（遥控、订餐等）
- 闲聊型对话：自然语言，获取服务（消遣、情感等）
- 问答：自然语言，获取信息
- 推荐：历史信息和行为，获取信息（主动推荐）
- 搜索：关键词，获取信息

■ 各类型人机对话任务的共性

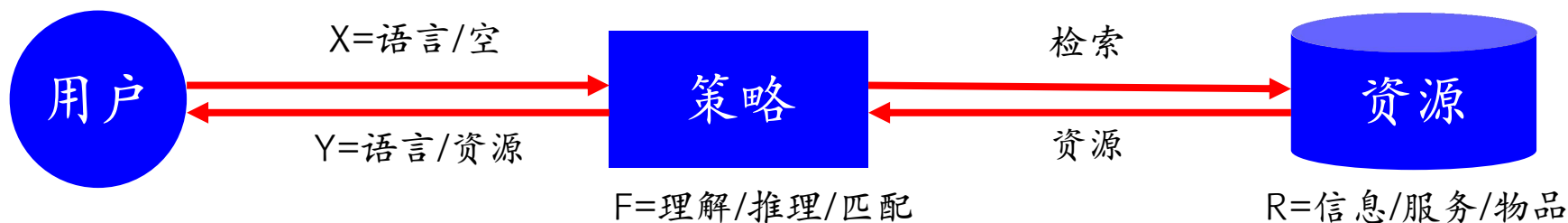
- 语言理解：理解人类以语言或词语形式发出的表达（词法、句法、语义等）
- 用户建模：刻画用户的兴趣偏好和行为习惯等（浏览、询问、购物历史等）
- 资源建模：理解用以满足用户需求的信息资源（图谱/表格/网页/文本/图像/视频/商品）
- 需求满足：建立从用户需求到信息资源的映射关系（问答、搜索、推荐、对话等模型）

抽象模型



■ 人机对话模型：在用户和资源之间建立映射

- 用户：人
- 资源：信息、服务、物品



■ 各具体交互形态的定义

- 闲聊/对话: $X = \text{语言}$, $F = \text{理解} + \text{推理}$, $R = \text{空/信息/服务}$
- 问答/搜索: $X = \text{语言}$, $F = \text{理解} + \text{推理/匹配}$, $R = \text{信息/物品}$
- 推荐: $X = \text{空}$, $F = \text{推理} + \text{匹配}$, $R = \text{信息/服务/物品}$

谢谢大家！

