

人机对话系统体验提升之道

刘升平 博士

资深技术专家，云知声AI Labs

目 录

- 人机对话系统介绍
- 体验提升之道
 - 全双工交互
 - 从语义到语用
 - 从闲聊到知性
- 人机对话系统的产业化

人机对话系统（会话交互，CUI）有望替代APP，成为IoT时代主要人机交互方式



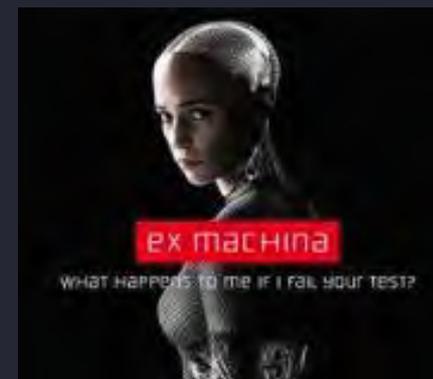
Apple SIRI



Amazon Echo



Jarvis@Iron Man

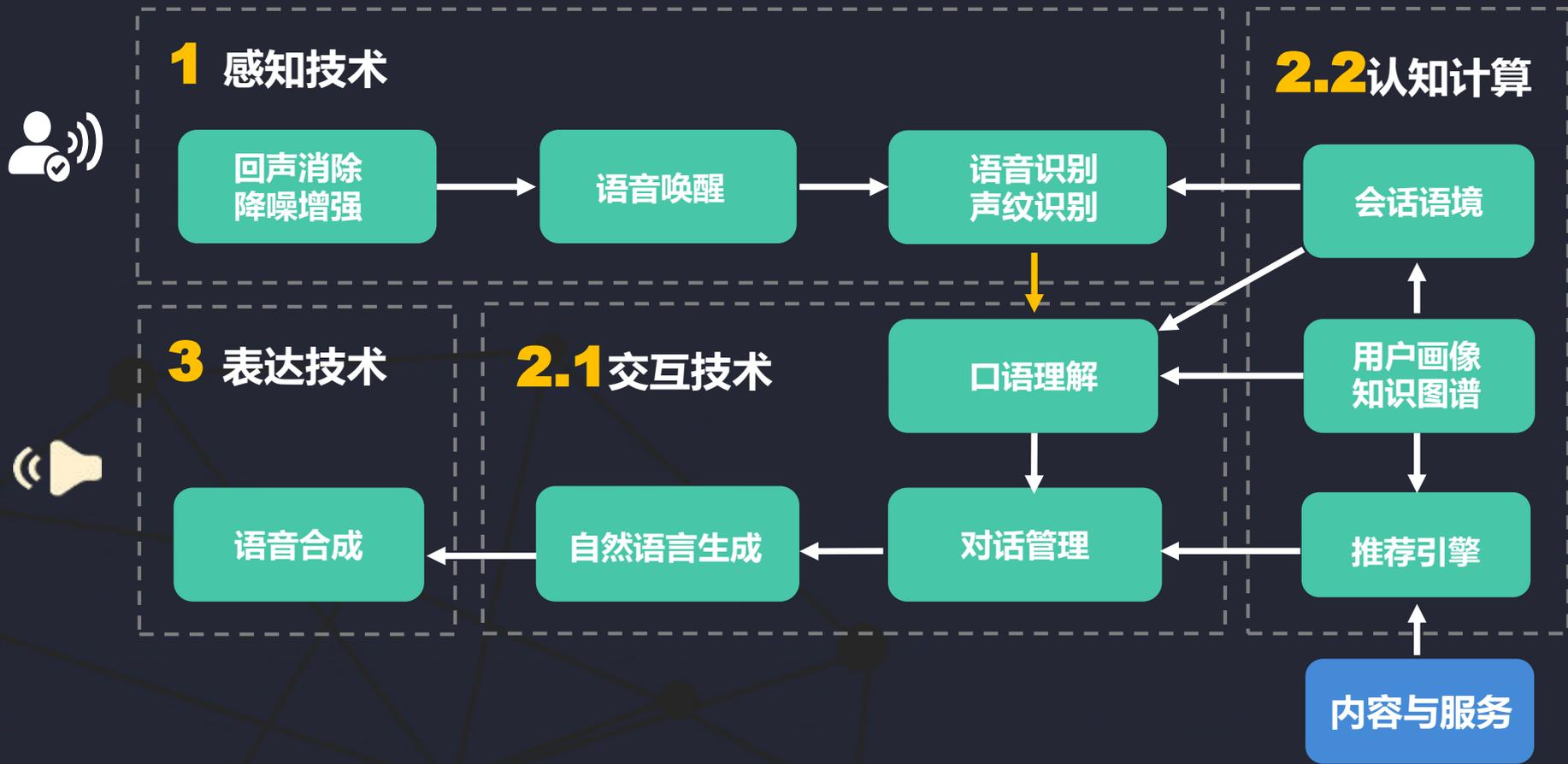


Eva@Ex Machina

人机对话系统



人机对话系统架构：复杂版



人机对话系统中的机器角色



体验提升之道

- 全双工交互
 - 基于语音和语言紧耦合
- 从语义到语用
- 从闲聊到知性

全双工交互

基于语音和语义的紧耦合

语音唤醒

■ 为什么需要语音唤醒？

- 进入语音交互场景，多设备区分：“知道是叫我”
- 无关语音过滤：“不叫我就不理你”

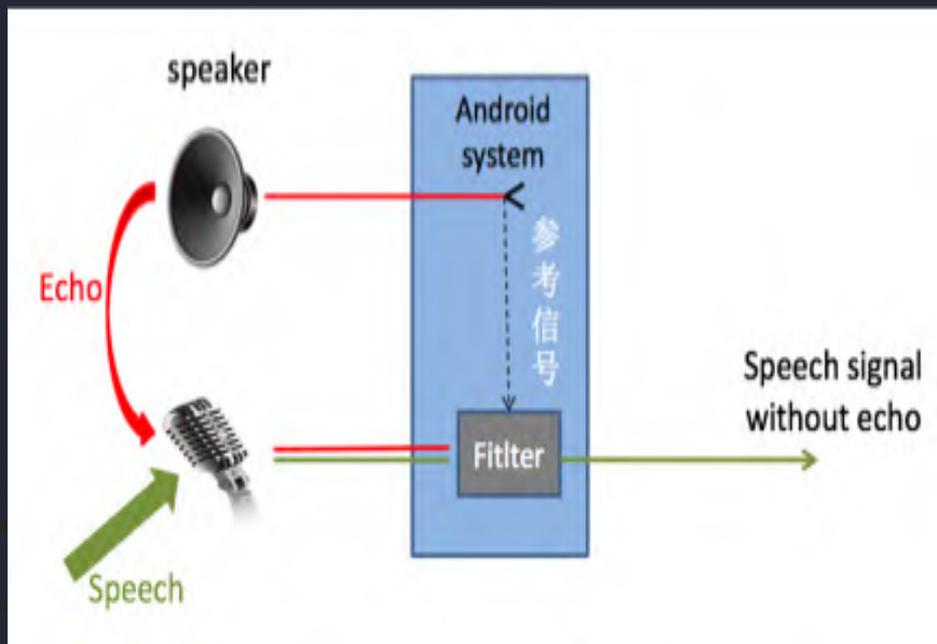
■ 语音唤醒的技术特点

- 持续监听：离线，超低功耗
- 冷唤醒：在未开机只通电情况下，能够唤醒
- 基于关键词检出，并不是做语音识别
- 唤醒词是否可定义：需录音/开发时/运行时

■ 评测方法

- 唤醒率：如 95%
- 误唤醒率：如：1次/24h

语音打断



Automatic Echo Cancellation:

- **TTS Barge-In** (提示音打断)
- **Music Barge-In** (音乐打断)

应用条件：

- ①单mic支持双通道录音;
- ②硬件设计上需要将一路信号输入给硬件作为参考信号;
- ③只能作用于设备自身的喇叭播放的声音

利用唤醒词打断

利用语音指令打断

智能断句 和 无效语音拒识

导航到。。。东单。。。公园。。。嗯啊。。。

■特征选取

- 声学信息：停顿时长，声音能量值，信噪比
- 声纹信息：说话人跳变检测
- 语言信息：语义完整性检测，上下文主题相关性检测

■分类模型

- 基于SVM

从语义到语用

基于语用的口语理解，生成和交互

口语理解

- 把用户的语句转换为机器可执行的指令

自然语言

分词及实体识别

业务逻辑

机器可处理的查询

北京明天下雨吗?

北京/city 明天/time 下雨 吗

任务：天气查询
时间：2016-08-27
地点：北京市
主题：雨



从语义到语用



语用学

• 语用学

- 是语言学各分支中一个以语言意义为研究对象的新兴学科领域
- 它研究在特定情景中的特定话语，研究如何通过语境来理解和使用语言。

• 计算语用学

- 理解：如何从语境中解码相关属性
- 生成：如何基于语境产生语言表达

语用计算：面向人与机器的对话式交互服务

理解 生成 交互



语用计算

语用计算是人机对话交互的核心

物理语境

时间/地点/场所

物理世界 (Sensors)

设备显示 (屏幕内容)

情绪和情感

非言语行为

言语语境

上下文

主题及焦点

设备反馈

知识语境

人类常识

领域知识

用户画像

Robot 画像

设备信息库

语境的生命周期：

请求级

会话级

长期

口语理解：会话意义理解

字面意义 + 语境 → 会话意义

字面意义理解

- 用逻辑语言USR(Underspecified Semantic Representation)来表示
- 把仅依赖语言学知识不能理解的部分留空，而不是枚举所有可能的语义

会话意义理解

- 结合语境信息，做指代和省略消解
- 结合语境信息，推理出会话含义
- 结合语境信息，对候选字面意义重新排序

User：雁栖湖在哪儿

Sys：北京市怀柔区怀北镇北部雁水路3号

User：那儿天气怎样？

Sys：天气很好，晴，20-28度，微风。

User：附近有什么农家院

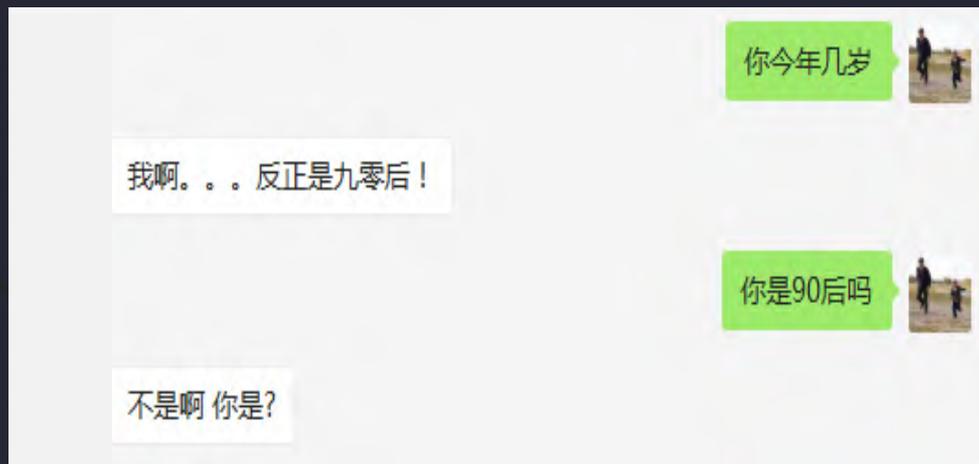
Sys：雁栖湖附近的农家院有....

自然语言生成

人机对话交互中，用户唯一能体验到的是
机器的应答！

个性化
(User Profile)

一致性
(Robot Profile)



基于语用计算的人机交互框架

以语境为中心
跨垂直领域
跨交互形式



User : 你好 (聊天)

Sys : 下午好

User : 明天青龙峡天气怎样? (问答)

Sys : 天气很好, 晴, 20-28度, 微风。

User : 怎么这么冷啊 (聊天)

Sys : 山区温度低, 记得多带些衣服

User : 订一家那附近的农家院 (任务)

Sys: 要什么价位的?

User : . . .

从聊天到知性

基于多源知识融合的聊天，问答和推荐

知性会话：基于知识的个性化专业化会话服务



知性会话交互示例

- Q: 贝加尔湖畔是谁唱的?
- A: 贝加尔湖畔的演唱者是李健(KBQA)
- Q: 他的音乐风格是什么?
- A: 李健的风格，有民谣的简洁，但比民谣华丽得多；有流行音乐的流畅，但又比流行更雅致，他纯净、复古的嗓音更温暖，更能受到文艺青年们的青睐。(DocChat)
- Q: 有哪些代表作?
- A: 李健代表作有传奇、风吹麦浪、贝加尔湖畔、向往(KBQA)
- Q: 哪个学校毕业的?
- A: 清华大学(电子工程系) (MetaQA)



人机对话系统 的产业化

智能人机交互的商业价值——云端芯



AI Service
提供云端智能服务

应用领域：医疗与教育

AI 芯
解决终端感知

应用领域：家居

AIUI
解决智能终端交互

应用领域：车载

语义云服务

- 语义解析
- 自动问答
- 聊天机器人
- 个性化推荐
- 多轮对话
- 机器人调教
- 领域支持：60+
 - 通讯/资讯
 - 常用工具
 - 娱乐休闲
 - 交通出行
 - 家居控制
- 语义解析准确率：95%

实时定制优化，快速Bug-Fix

语义解决方案

	主要领域	主要亮点	重点客户
电视	电视操控, 电视节目查询, 视频搜索, 影讯, 网站导航, 购物	<ul style="list-style-type: none"> • 视频新词更新, 语义解析准确 	乐视, TCL, 长虹, 风行者等
车机	<ul style="list-style-type: none"> • 通讯: 电话, 微信, 短信, 联系人, 黄页 • 交通: 导航, 路况, 限行, 违章, 油价, 地图操作 • 多媒体: 音乐, 有声读物, 广播, 天气, 股票, 新闻 	<ul style="list-style-type: none"> • 导航POI解析准确率业内领先 • 跨领域上下文理解 • 支持口语自由说 	同行者, 品网, 小蚁等50+ 厂商
音箱	音乐, 有声读物, 提醒, 备忘, 闹铃, 日历	<ul style="list-style-type: none"> • 对音乐和有声读物的深度优化 • 基于事件的提醒 • 关于音乐的问答和聊天 	融帆, 火灵鸟, 三诺, 海尔
家居	50+种家居设备单控, 中控, 菜谱, 食物管理	<ul style="list-style-type: none"> • 语用计算支持 • 个性化语言反馈 • 决策模型自动生成动作序列 	海尔, 美的, 格力, 长虹, TCL等
儿童教育	百科知识, 儿童聊天, 唐诗宋词, 翻译, 计算器	<ul style="list-style-type: none"> • 百科知识内容丰富 	巴巴腾, 小U机器人等

案例应用：智能家居

在智能家居领域，云知声结合技术与工程之优势成功达成与多方知名品牌的合作，包括美的，格力，华帝，海信的产品以落地，**并联手打造标杆产品**

主打产品：智能语音模组（IVM），面向家居智能化的云端芯一体化解决方案

合作客户及应用案例



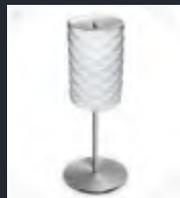
美的挂式空调



格力立式空调



乐视超级电视



小巢台灯



海尔立式空调



康力优蓝机器人



火百灵音箱



华帝抽油烟机

远讲识别

面向家居和车载领域的远讲识别和降噪技术取得突破，可以提供超过5米距离的远讲识别方案以及高性价比的双麦克风和四麦克风方案

冷唤醒

可以实现在智能硬件、应用休眠的状态下，通过个性化唤醒词唤醒

全语音自然交互

全程自然语言般的对话，可以听懂自然语言，并通过TTS的方式以自然语言的方式与人交互

案例应用：车联网

在智能车载领域，云知声已服务近百家方案商、品牌商，在车载后装市场上占有率超过60%，位列第一。目前搭载云知声整体软件方案的设备出货量已经接近200万台。另外，搭载云知声SDK的第三方车机出货量则已经超过600万台

主打产品：车载语音交互软件（AIUI for Unicar），方案包括拾音降噪、语音交互、云端计算和内容服务一揽子解决方案

合作客户及案例

场景化识别

软硬降噪方案

软件AEC方案



车载中控、HUD



智能后视镜

合作客户包括：

互联移动、飞歌科视、小蚁、水清光电、远峰电子等企业



导航

导航回家



电话

打电话给老王

28



娱乐

我想听小苹果



社交

发微信给小云



资讯

今天会下雨吗

智享未来



谢谢