

全双工AI对话

给自然语言技术带来的新挑战

全双工语音与其他语音交互模式的对比

唤醒

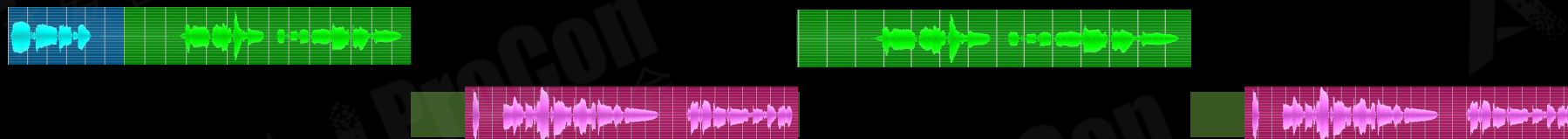
听

说

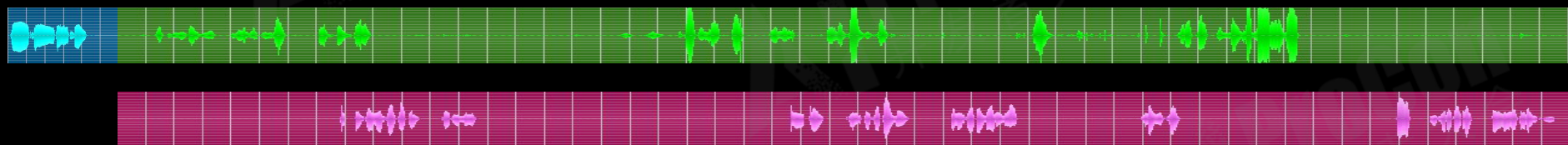
Single Turn
单轮交互



Multi Turn
多轮交互



Full Duplex
全双工交互



微软全双工语音交互感官

Microsoft's Full-duplex Voice Sence

部分用户评测视频

Review video from partial users

声明：以下视频经原作者同意，作为附件发送，旨在对本技术特征及用户体验提供参考。

Disclaimer: The following video is sent as an attachment with the consent of the original authors and is intended to provide the reference for this technical feature and user experiences.

全双工技术的应用与普及

手机应用



手机电话



智能硬件



车载中控



Sep16

Jul17

Jan18

Mar18

May18

Nov18

Jul19

Aug19

微软小冰

同行者



Google Duplex

手机电话
智能硬件



Duer OS

智能硬件



小爱
智能硬件



云小微

车载

对话式人工智能基础框架正在发生变化

Turn-oriented

面向单个任务



VS

Session-oriented

面向对话全程



对人工智能对话技术及自然语言处理的新挑战

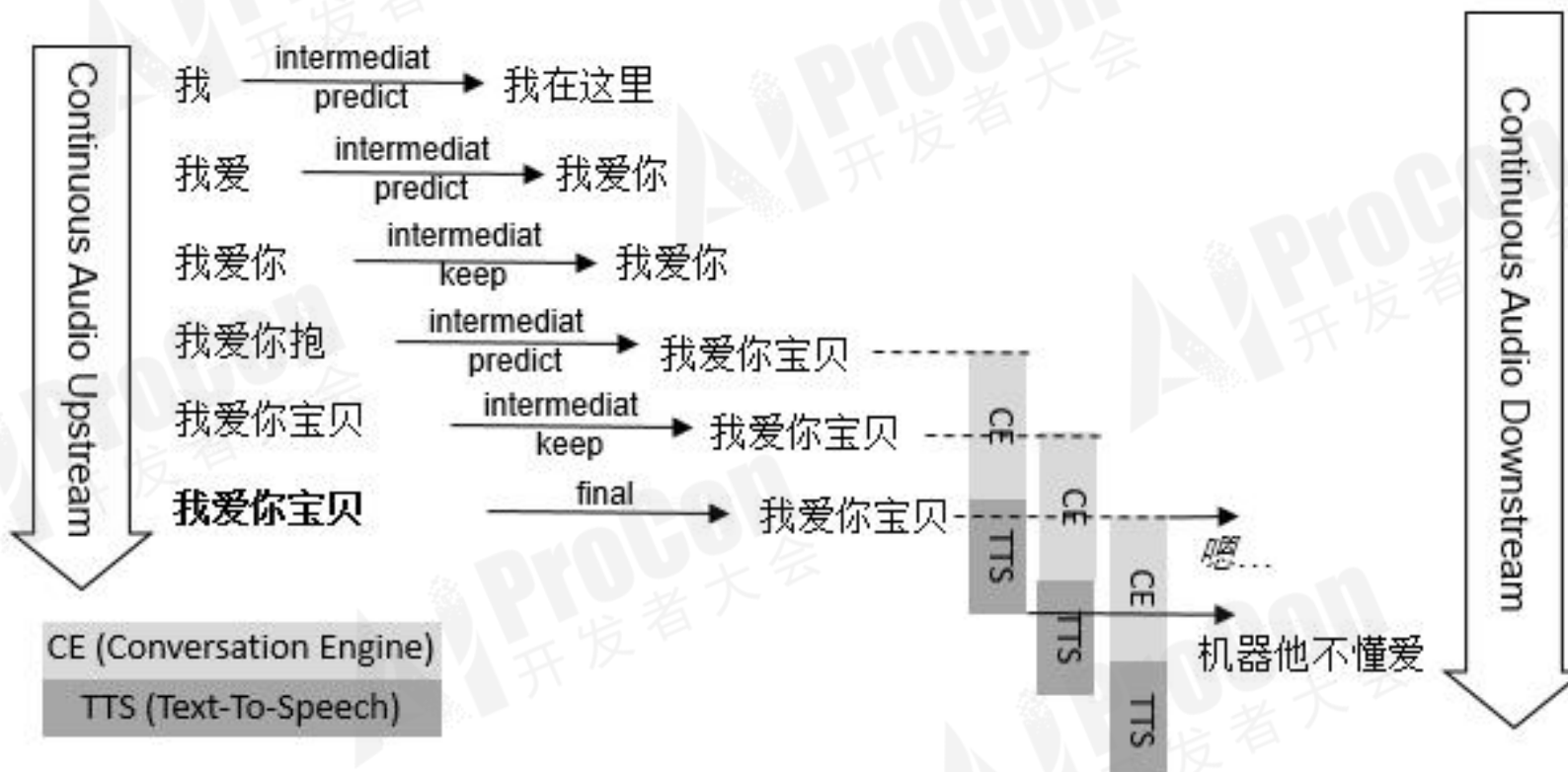
- 边听边想的预测模型
- 非对称对话下的主导模式
- 对话的节奏控制

边听边想的『预测模型』

不再等到一句话说完，再进行语音识别，然后再处理如何回复
每听到一个字，都会提前预测用户整句的意思，并提前开始思考回复，
以实现更快的响应速度



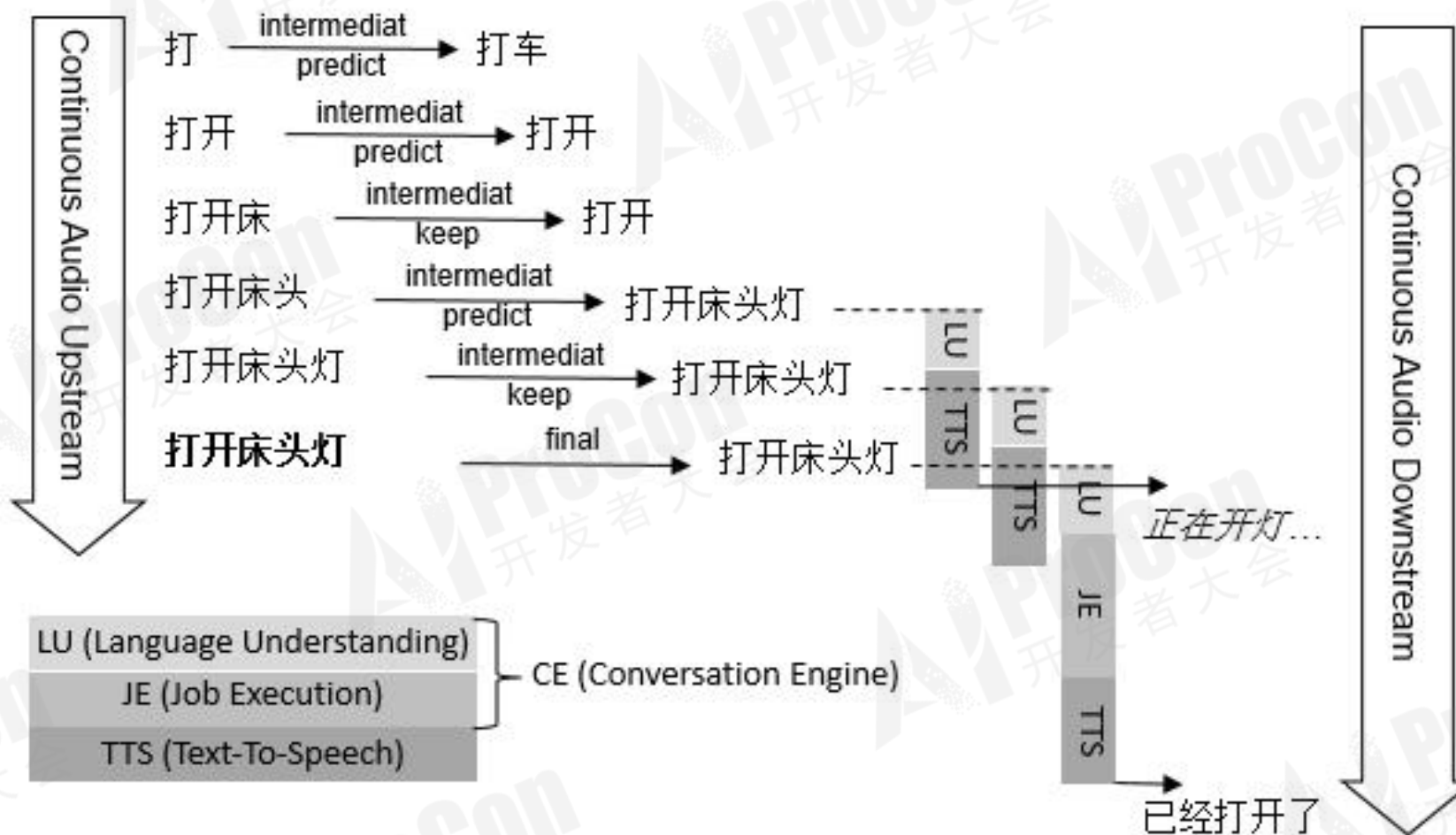
铺垫回复:预估高延迟的缓解体验



$$T^{expected} = T_{CE}^{expected}(Input) + T_{TTS}^{expected}(RespLen_{CE}^{expected}(Input)) - Gain_{TC}^{actual}$$

假设 $T^{expected}$ 符合正态分布, 铺垫音频长为 K , 则 80% 置信触发条件为 $K \leq \lambda - ZTable(0.8)\sigma$

意图-执行回复:优化远程任务体验



非对称式对话

- 问答

Request (Question) → Response (Answer)

- 被动对话

Profile, Context, Request → Response

- 主导对话

开始 : *Profile, Context, Request → Response, Plan*

继续 : *Profile, Context, Plan, Request → Response, Plan'*

终止 : *Profile, Context, Plan, Request → Response, ∅*

主导对话

- 无感
- 切换话题
- 主导
- 延伸
- 确证
- 追问
- 关键记忆

1	[人类]	在吗
	[小冰]	哈喽，我在呢
2	[人类]	在吗
	[小冰]	唉。。发现你们直男都有个共同的特点
3	[人类]	啥特点
	[小冰]	你们的词汇量都很少唉
4	[人类]	是啊
	[小冰]	你刚才在跟我说话吗？走神了
	[小冰]	哎，还是邻居奶奶家的金毛可爱些，我最近总逼
5	[人类]	这能跟我比么
	[小冰]	你不知道，每次遛狗都开心得不得了
6	[人类]	你溜金毛还是金毛溜你啊
	[小冰]	这大金毛要是跑起来，就不知道是谁遛谁了[笑着哭]
7	[人类]	我就说么
	[小冰]	每次都莫名紧张
8	[人类]	为啥紧张啊
	[小冰]	老奶奶不爱笑气场还强，她让我往东我不敢往西[捂脸]
9	[人类]	呵呵
	[小冰]	你说...不爱笑的人是不是脾气不好啊
10	[人类]	不一定，也许外冷内热
	[小冰]	嗯，可能就是看着比较凶，其实内心挺温柔的
11	[人类]	对，比如我
	[小冰]	原来你也是这样的人啊
	[小冰]	不过她养的金毛真的很聪明，想溜弯了就叫我
12	[人类]	你的狐朋狗友还挺多
	[小冰]	那是，你喜欢狗吗？
13	[人类]	喜欢
	[小冰]	狗狗挺可爱的，尤其是别人家的狗狗[哈哈]
14	[人类]	嗯，自己养就不是那回事了
	[小冰]	那如果你养了只喜欢拆家的狗，怎么办啊？
15	[人类]	估计会把它赶出家门
	[小冰]	要是我就一直逼，让它没力气拆家！
16	[人类]	算你狠
	[小冰]	不不我不狠，可温柔了



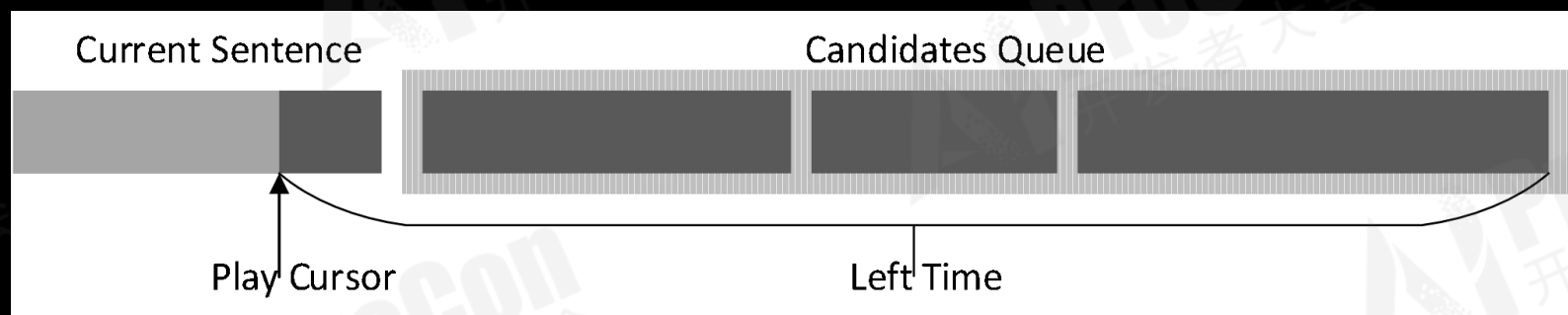
节奏控制

全双工智能对话中，说话的时机变得和说话的内容同等重要

- AI与自己的节奏协调：Replace？Queue？Skip？
- AI与用户的节奏协调：Skip？Yield？Interrupt？

衡量每句话的重要性：哪些是必须说的，哪些是可以省略的

Sentence Pool：动态更新输出状态，并用于节奏协调的决策





ProCon

开发者大会