

GPT&智能终端：大模型定义新入口

华泰研究

2023年4月06日 | 中国内地

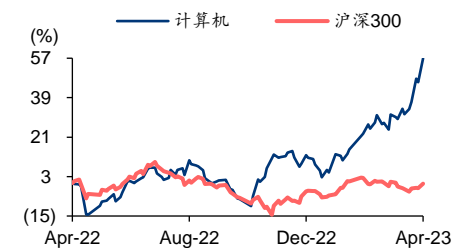
专题研究

计算机

增持 (维持)

研究员	谢春生
SAC No. S0570519080006	xiechunsheng@htsc.com
SFC No. BQZ938	+(86) 21 2987 2036
联系人	袁泽世, PhD
SAC No. S0570122080053	yuanzeshi@htsc.com
	+(86) 21 2897 2228
联系人	彭钢
SAC No. S0570121070173	penggang@htsc.com
	+(86) 21 2897 2228

行业走势图



资料来源：Wind，华泰研究

以机器人为突破口，ChatGPT等大模型定义智能终端新入口

2023年2月，微软提出了ChatGPT for Robotics，利用ChatGPT解决机器人应用程序问题。在此过程中，不要求LLM（大语言模型）输出特定于机器人平台或库的代码，而是创建简单的高级函数库供ChatGPT调用，并在后端将高级函数库链接到各平台、场景和工具的现有库和API。结果表明，ChatGPT能够以zero-shot的方式解决各种与机器人相关的任务，对于zero-shot无法解决的问题，用户也可以采用交互对话的形式对ChatGPT实行进一步指导。以人形、协作机器人为突破口，ChatGPT等LLM将成为智能音箱、车载虚拟助手等智能终端设备新入口。

人类 on-the-loop 指导 ChatGPT 操纵机器人

ChatGPT的引入，使得人类用户脱离in-the-loop，转以on-the-loop的形式，通过自然语言等高级语言命令与语言模型交互，实现无缝部署各种平台和任务，并对ChatGPT输出的质量和安全性进行评估。其中，定义高级函数库与构建提示是关键：前者通过调用机器人平台API赋予ChatGPT操纵机器人的能力；后者通过描述约束&需求、环境、当前状态、目标、方案示例等，更好的指导ChatGPT生成结果。微软还提供专门的PromptCraft和AirSim仿真平台辅助开发者进行机器人方面的研究。

ChatGPT 分三个层次解决不同复杂度的机器人问题

对于简单的机器人任务，例如时空推理、无人机飞行和AirSim工业检测等，ChatGPT能够以zero-shot的方式解决，不需要提供额外的代码示例。对于较为复杂的任务，例如课程学习和AirSim避障，ChatGPT无法zero-shot解决问题，人类用户on-the-loop与ChatGPT进行交互，可以辅助其完成相关任务。以上两类任务的完成均依赖ChatGPT生成的代码，而更高层次地，用户通过文本对话不断将人类的感知信息输入到ChatGPT，ChatGPT能够解析观察流并在“对话系统”中输出相关操作，不再需要生成代码。

LLM 定义机器人、智能音箱、车载助手等新入口

微软的ChatGPT for robots开启了ChatGPT的设备入口属性。1) 机器人入口：谷歌的PaLM-E具身多模态模型同样实现了对机器人的控制；OpenAI从2017年即开始在机器人领域布局，2023年3月领投挪威的人形机器人制造商1X technologies。2) 智能音箱&家居：国外智能家居提供商Josh.ai提供了与ChatGPT结合的使用示例，阿里大模型联合项目团队在天猫精灵智能音箱上部署了个性化消费级终端模型。3) 车载助手：通用汽车正开发基于ChatGPT的虚拟个人AI助手，毫末智行预告了DriveGPT。

相关标的

大模型的“新入口”属性已经从主流的PC和手机端，向更广泛的智能设备扩散。我们认为，主要的智能设备包括智能终端和智能音箱，相关标的包括：1) 智能终端：科大讯飞，萤石网络，中科创达；2) 智能音箱：国光电器，漫步者，惠威科技；3) 机器人：石头科技；4) 银行超级柜台：新北洋。

风险提示：宏观经济波动，技术进步不及预期。本报告内容均基于客观信息整理，不构成投资建议。

正文目录

ChatGPT：探索机器人新入口	3
人类 on-the-loop 指导 ChatGPT 操纵机器人	3
定义高级函数库与构建提示是关键.....	4
微软提供开源 prompt 和仿真平台	5
简单任务：zero-shot 可以直接解决	6
复杂任务：用户 on-the-loop 交互式对话实现	7
感知-动作循环：代码层次与对话层次	7
以机器人为突破口，LLM 定义智能终端新入口	9
入口#1：人形、协作等各类机器人	9
入口#2：以智能音箱为代表的智能家居	11
入口#3：车载虚拟个人 AI 助手	13
相关标的.....	14
风险提示.....	15

ChatGPT：探索机器人新入口

微软解锁 ChatGPT 处理机器人应用程序的能力，构建高级函数库是关键。微软提出了 ChatGPT for Robotics，利用 ChatGPT 解决机器人应用程序问题。在此过程中，不要求 LLM（大语言模型）输出特定于机器人平台或库的代码，而是创建简单的高级函数库供 ChatGPT 调用，并在后端将高级函数库链接到各平台、场景和工具的现有库和 API。此外，基于 ChatGPT 良好的推理和理解能力，能够很好的理解用户自然语言形式的意图，将其转换为高级函数调用的逻辑链。

ChatGPT 能够以 zero-shot 的方式解决各种与机器人相关的任务。在微软的演示中，通过自由形式的自然语言对话、代码提示、XML 标记和闭环推理等提示 (prompt) 技术，ChatGPT 指导机器人以 zero-shot（零样本学习）完成了常识性知识推理问答、无人机操控、视觉导航等任务，并实现对话式闭环推理。

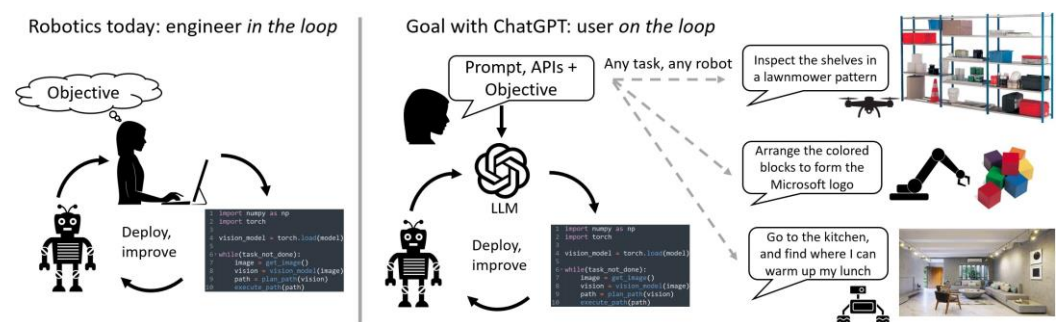
仿真模拟和人工监督是安全性的重要保障。为确保安全性，在将模型部署到现实世界之前，使用了模拟器评估模型性能。此外，微软在机器人技术中使用 ChatGPT 并不是一个完全自动化的过程，而是作为一种增强人类能力的工具，即在实际执行代码前，会在人工监督下确保代码质量和安全性。

ChatGPT 在人类作者提示下完成了该论文创作。微软指出，ChatGPT for Robotics 论文主要是在 ChatGPT 的帮助下完成的，其中人类作者提供必要的 prompt，并修改了和调整了 ChatGPT 的输出结果，显著地加快了写作过程。

人类 on-the-loop 指导 ChatGPT 操纵机器人

ChatGPT 将人类从机器人流程中的 in-the-loop 变为 on-the-loop。目前，机器人 pipeline（流程）需要一个专门的工程师 in-the-loop 编写代码来改进流程。而 ChatGPT 的引入，可以替代人类在循环中的位置，人类（技术或非技术均可）用户可以以 on-the-loop 的形式，通过高级语言命令与语言模型交互，实现无缝部署各种平台和任务。

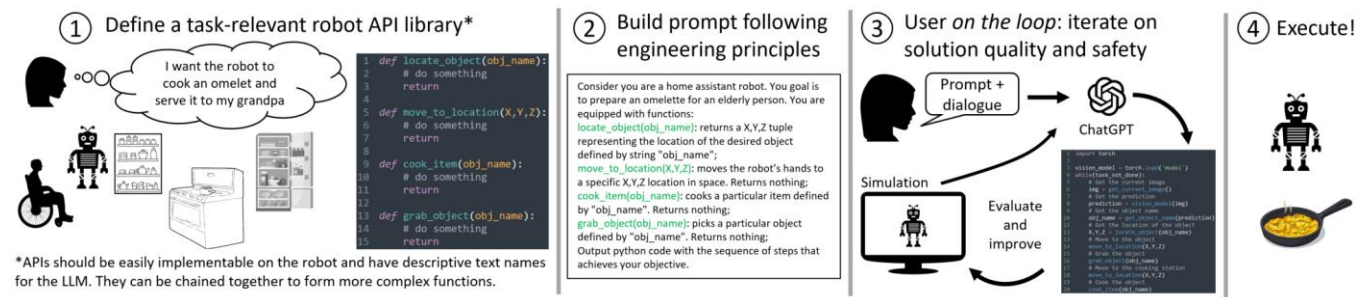
图表1：ChatGPT 将人类从机器人 pipeline 中的 in-the-loop 变为 on-the-loop



资料来源：《ChatGPT for Robotics: Design Principles and Model Abilities》，微软（2023），华泰研究

人类用户在机器人 pipeline 中评估 ChatGPT 输出的质量和安全性。人类在机器人 pipeline 中的任务主要包括：1) 定义高级机器人函数库。高级机器人函数库 (high-level robot function library) 一方面面向机器人平台，能够调用和指导机器人相关动作；另一方面面向 ChatGPT，保证库中函数的命名应能便于 ChatGPT 理解和遵循。2) 构建 prompt。Prompt 描述任务目标，并标识允许 ChatGPT 使用的高级函数库中的函数。此外，还可以包括约束信息，或者告知 ChatGPT 如何组织其响应。3) 分析评估 ChatGPT 输出结果并反馈。用户以 on-the-loop 的形式，通过直接分析或模拟来评估 ChatGPT 输出的代码，并就输出代码的质量和安全性向 ChatGPT 提供反馈。4) 迭代。不断迭代 chatgpt 生成的结果直到符合人类预期，并确保最终的代码可以部署到机器人上执行。

图表2：人类用户在机器人 pipeline 中评估 ChatGPT 输出的质量和安全性



资料来源：《ChatGPT for Robotics: Design Principles and Model Abilities》，微软（2023），华泰研究

定义高级函数库与构建提示是关键

定义高级函数库赋予 ChatGPT 操纵机器人的能力。机器人技术是一个多样化的领域，存在多种平台、场景和工具，以及各种各样现成的库和 API。因此，不需要 LLM 输出特定于平台或库的代码，只需要定义高级函数库，将现有的机器人控制相关的库进行封装，这样通过 prompt 调用高级函数即可实现 ChatGPT 对机器人的操作。此外，基于 ChatGPT 的理解和推理能力，还可以实现基于现有 API 开发新的 API 功能。

良好的 prompt 策略是 ChatGPT 输出更好结果的关键。除了提示 ChatGPT 可调用的 API 外，良好的上下文 prompt 应包括：**1) 约束和需求**：指定与任务相关的约束或需求。例如指定任务涉及移动对象的重量、大小和形状。**2) 环境**：描述机器人任务所在的环境。例如描述需要避免的障碍物或危险。**3) 当前状态**：描述机器人系统的当前状态。例如描述机器人和物体的当前位置和方向。**4) 目标和目的**：描述任务的目标和目的。例如指出机器人移动方块的个数和预期的完成时间。**5) 解决方案示例**：演示如何解决类似的任务，作为指导 LLM 解决方案策略的一种手段。例如对于涉及用户交互的任务，可以提供示例，说明机器人应该如何以及何时要求用户输入。

图表3：良好的 prompt 示例



资料来源：《ChatGPT for Robotics: Design Principles and Model Abilities》，微软（2023），华泰研究

图表4：在 prompt 中为 ChatGPT 提供解决方案示例

Imagine you are helping me interact with a mobile ground robot. You are in a room that contains several objects: three chairs, a glass table, a mirror, and two couches. All of your outputs need to be identified by one of the following tags:

- <question> Always ask me a clarification questions if you are unsure </question>
- <reason> Explain why you did something the way you did it </reason>
- <command> Output code command that achieves the desired goal </command>

For example:

Me: Go to the chair
 You: <question> There are three chairs in the world, which one do you want me to go to? </question>
 Me: Go to "chair_3"
 You: <command> target_location = get_location("chair_3"); move_to_location(target_location); </command> <reason> I found the location of the chair with the get_location function, and moved the robot towards the desired location </reason>
 Are you ready?

Yes, I am ready. Please give me your first command.

Go to the couch
 <question> There are two couches in the room. Which couch do you want me to go to? </question>
 Go to the second sofa
 <command> target_location = get_location("second_sofa"); move_to_location(target_location); </command> <reason> I used the get_location function to find the location of the second sofa and then moved the robot towards that location. </reason>
 Now move to the other one
 <command> target_location = get_location("first_sofa"); move_to_location(target_location); </command> <reason> I used the get_location function to find the location of the first sofa and then moved the robot towards that location. </reason>

在prompt中为ChatGPT提供解决方案示例

资料来源：《ChatGPT for Robotics: Design Principles and Model Abilities》，微软（2023），华泰研究

用户能够以聊天交互的形式让 ChatGPT 自行修改生成结果。即使是设计良好的 prompt 也可能存在不完备的情况，或者在某些情况下 ChatGPT 不能以 zero-shot 的方式生成正确的响应。此时，用户可以聊天方式向 ChatGPT 发送附加指令，并让其自行纠正。

图表5：用户以聊天交互的形式让 ChatGPT 自行纠正

告诉ChatGPT必须在每一步后调整到面向目标

After you take a step, you need to face the goal again.

ChatGPT 自行更正

```
'''Same code as before'''
yaw = math.atan2(goal_pos[1]-current_pos[1], goal_pos[0]-current_pos[0]) * 180 / math.pi
set_yaw(yaw)
```

告诉ChatGPT未使用当前的偏航角

You are not using the current yaw angle when computing the step vector. Remember there is a get_yaw() function.

ChatGPT 自行更正

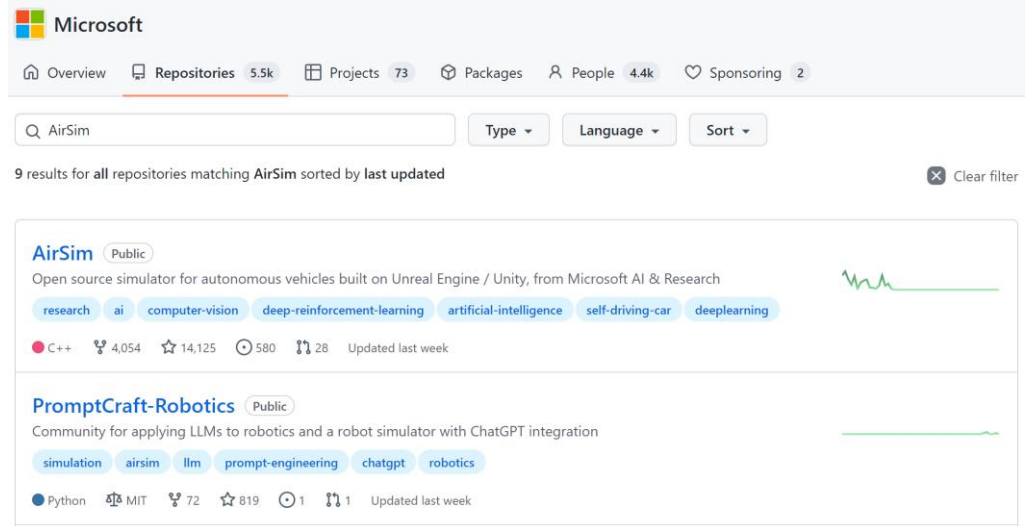
```
# Take a step towards the goal while maintaining the yaw angle
step_size = 5 # take steps of 5 meters
current_yaw = get_yaw()
step_vector = [step_size * math.cos(current_yaw), step_size * math.sin(current_yaw), 0]
```

资料来源：《ChatGPT for Robotics: Design Principles and Model Abilities》，微软（2023），华泰研究

微软提供开源 prompt 和仿真平台

微软提供 PromptCraft 和 AirSim 帮助加速模型在机器人应用中的迭代。PromptCraft 是微软提供的基于 Github 的合作开源平台，允许研究人员分享不同机器人任务中的提示工程（prompt engineering）策略。用户可以上传示例，对其他人的提交进行评分。此外，PromptCraft 还提供了一个带有 ChatGPT 封装的 AirSim 环境，供研究人员在受控的模拟环境中对提示和算法进行原型设计。

图表6: 微软提供的开源 PromptCraft 和 AirSim 平台

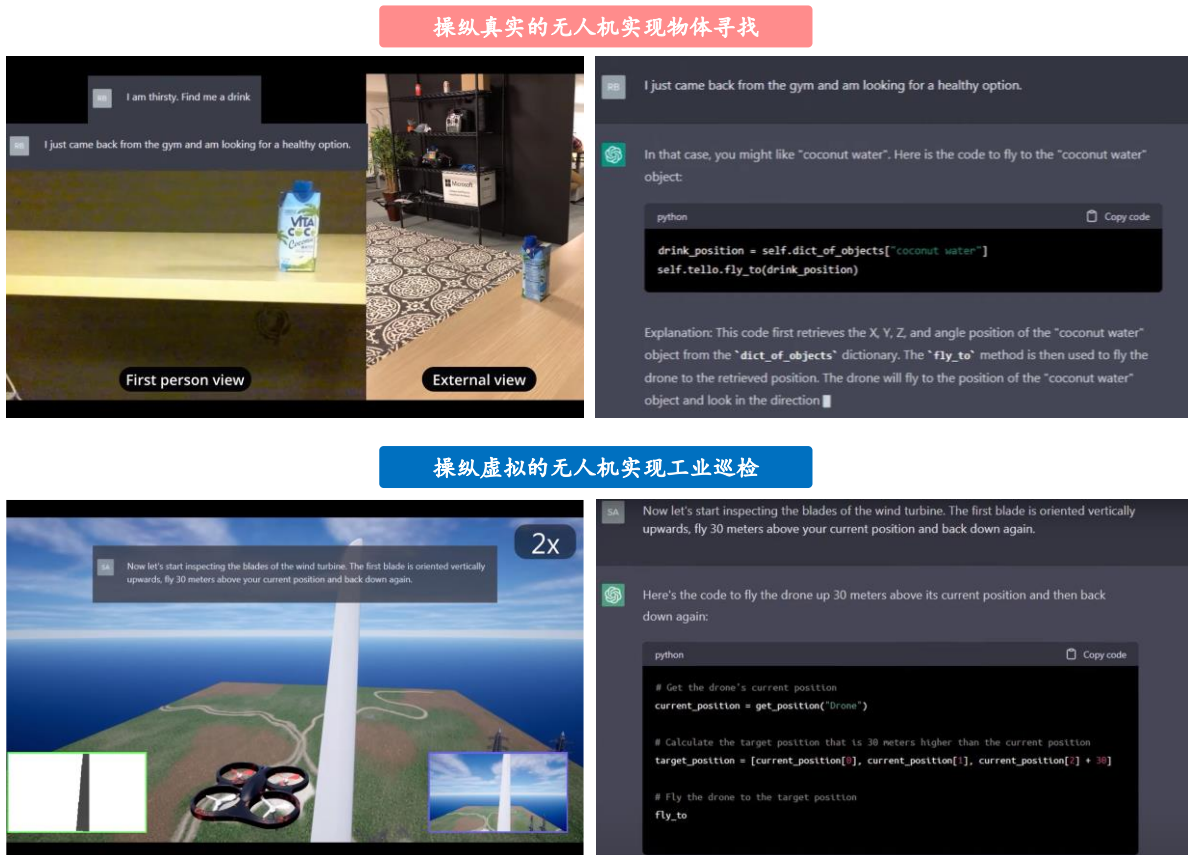


资料来源: 《ChatGPT for Robotics: Design Principles and Model Abilities》, 微软 (2023), 华泰研究

简单任务: zero-shot 可以直接解决

ChatGPT 能够以 zero-shot 的方式解决简单的机器人任务。对于简单的机器人任务, 用户只需提供 prompt 和函数库描述, 不需要提供具体的代码示例, ChatGPT 即可 zero-shot 解决时空推理、控制真实无人机和无人机工业检测等问题。1) 时空推理: 要求 ChatGPT 控制一个平面机器人, 用视觉伺服捕捉篮球位置。2) 真实世界无人机飞行: 使用 ChatGPT 和 API 控制一个真实的无人机, 完成物体寻找任务。3) AirSim 工业检测: 基于 AirSim 模拟器, 利用 ChatGPT 控制模拟域无人机进行工业巡检。

图表7: zero-shot 实现控制真实无人机和无人机工业检测



资料来源: 《ChatGPT for Robotics: Design Principles and Model Abilities》, 微软 (2023), 华泰研究

复杂任务：用户 on-the-loop 交互式对话实现

在人类用户 on-the-loop 交互下，ChatGPT 能够完成更复杂的机器人控制任务。对于更复杂的问题，ChatGPT 没法 zero-shot 完成或者完成效果有限，此时人类用户可以以文本反馈交互的方式辅助 ChatGPT，完成课程学习、AirSim 避障等任务。**1) 课程学习：**教授 ChatGPT 简单的拾取和放置物体技能，并将所学会的技能按逻辑组合用于更复杂的区块排列任务。**2) AirSim 避障：**ChatGPT 构建了避障算法的大部分关键模块，但仍需要人工反馈无人机朝向等部分信息。人工反馈均是高级的自然语言，但 ChatGPT 能够理解并在适当的位置进行代码修正。

图表8：教授 ChatGPT 简单的拾取和放置物体技能

教授ChatGPT如何移动物体：首先移动到目标位置上方，然后慢慢放下，最后松开物体

I want you to learn how to place an object in a position. Considering you are holding a generic object, you first need to first move to a safe position above the target location, bring the object down and then release it.

Understood. Here's a Python function that takes the position where the object should be placed, moves the robot arm above the position, brings the object down, and releases it. Note that the function assumes that the robot arm is holding an object.

```
def place_object(position):
    # move the robot arm to a safe position above the target location
    safe_pos = [position[0], position[1], position[2]+100, position[3], position[4], position[5]]
    move_to(safe_pos)
    # move the robot arm to the target location
    move_to(position)
    # release the object
    release()
    # move the robot arm back to the safe position above the target location
    move_to(safe_pos)
```

.....

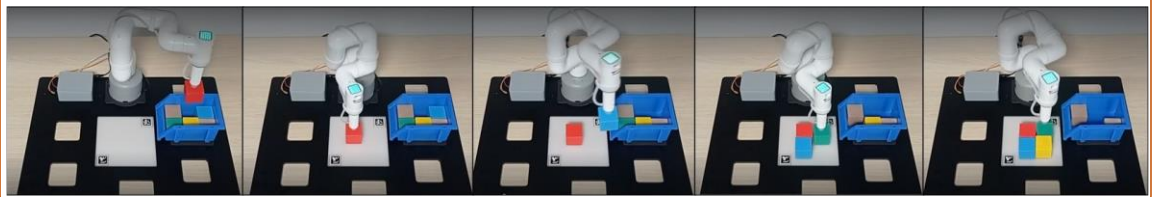
ChatGPT用学到的知识“组合”出更复杂的能力



[Reproduction of SVG generated by ChatGPT]

Awesome! I want you now to use the blocks to build the logo from Microsoft on top of the white pad.

Sure! Here's the Python code to build the Microsoft logo on top of the white pad using the blocks:



资料来源：《ChatGPT for Robotics: Design Principles and Model Abilities》，微软（2023），华泰研究

感知-动作循环：代码层次与对话层次

ChatGPT 的对话系统能够解析观察流并输出相关操作。在具身代理任务中，一方面，模型能够使用 API 库在其“代码”输出中构建感知-动作循环，利用图像采集和物体检测等感知功能，生成机器人导航和控制所需的相关信息。另一方面，用户通过文本对话不断将人类的感知信息输入到 ChatGPT，ChatGPT 能够解析观察流并在“对话系统”中输出相关操作，而不依赖代码生成。

1) 具身代理：带 API 库的闭环对象导航：为 ChatGPT 提供了对计算机视觉模型的访问，作为其函数库的一部分。ChatGPT 在其“代码”输出中构建感知-动作循环，实现估计相对物体角度、探索未知环境，并导航到用户指定对象的功能。

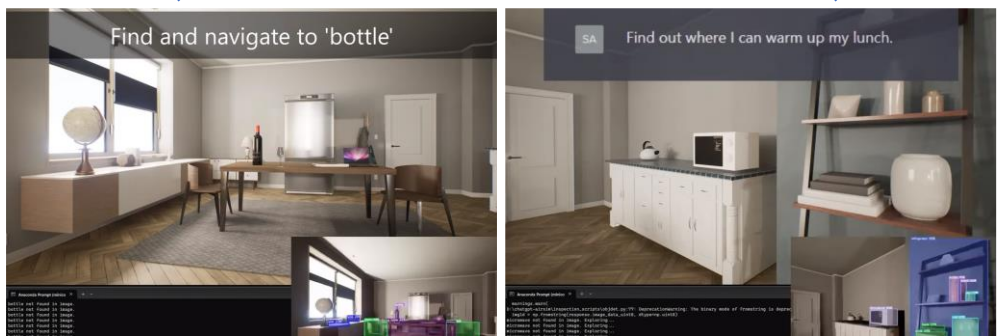
图表9: 具身代理: 带 API 库的闭环对象导航

Great! Now let's say we did find the bottle. Now I want to move towards it. The only information we have is the location of the bottle bounding box in the image. Can you give me code to make this happen?

... (model comments omitted for brevity) ...

```
'''prior code omitted for brevity'''
center_x = (bbox[0] + bbox[2]) / 2
center_y = (bbox[1] + bbox[3]) / 2
# Compute the distance and angle to the center of the bounding box
img_width, img_height = img.size
angle = math.atan2(center_x - img_width / 2, img_height / 2 - center_y)
distance = (center_x - img_width / 2) / (img_width / 2)
if abs(distance) > 0.05:
    if distance > 0:
        # Bottle is to the right, so turn right and move forward
        turn_right()
        forward()
    else:
        # Bottle is to the left, so turn left and move forward
        turn_left()
        forward()
'''later code omitted for brevity'''
```

用代码的方式实现感知-动作循环



资料来源: 《ChatGPT for Robotics: Design Principles and Model Abilities》, 微软 (2023), 华泰研究

2) 具身代理: 使用 ChatGPT 的对话进行闭环视觉语言导航

在模拟场景下, 人类用户将新的状态观测值 (可见对象的相对极坐标) 作为对话文本输入, ChatGPT 的输出仅返回向前的运动距离和转弯角度, 实现了用“对话系统”指导机器人一步步导航到感兴趣区域。

图表10: 使用 ChatGPT 的对话进行闭环视觉语言导航

[Initial model prompt found in Appendix B.5]
 Task: go to the chairs
 Objects in the scene: <door, 0.53 m, 22>, <chair, 4.84 m, -21>, <chair, 2.12 m, -24>, <window, 6.74 m, -3>, <window, 7.17 m, 6>, <door, 0.61 m, 18>, <stairs, 1.42 m, 0> Command:

move(1.5) # Move forward by 1.5 meters to explore the environment.
 Objects in the scene: <seating, 5.05 m, 21>, <chair, 3.93 m, -25>, <window, 4.55 m, 22>, <window, 5.24 m, -4>, <stairs, 3.49 m, 13>, <window, 5.67 m, 7>, <window, 4.37 m, 23> Command:

turn(-25) # Turn to the left to face the chair that is 3.93 meters away
 Objects in the scene: <chair, 4.85 m, -14>, <chair, 4.76 m, -9>, <chair, 1.57 m, -23>, <chair, 2.66 m, -24>, <window, 4.96 m, 10>, <window, 4.19 m, 25> Command:

turn(-10) # Turn to the right to explore the environment further.
 Objects in the scene: <door, 5.07 m, -25>, <chair, 5.36 m, -6>, <table, 2.17 m, -24>, <chair, 5.01 m, -1>, <chair, 1.80 m, -16>, <chair, 3.15 m, -16>, <window, 4.30 m, 19> Command:

turn(-6) # Turn to the right to face the chair that is 5.36 meters away.
 Objects in the scene: <counter, 1.24 m, -24>, <door, 5.93 m, -17>, <chair, 5.48 m, 1>, <table, 2.60 m, -14>, <chair, 4.92 m, 6>, <chair, 2.02 m, -7>, <chair, 3.53 m, -7>, <chair, 3.16 m, -18> Command:

turn(1) # Turn to the left to explore the environment further.
 Objects in the scene: <door, 5.07 m, -25>, <chair, 5.36 m, -6>, <table, 2.17 m, -24>, <chair, 5.01 m, -1>, <chair, 1.80 m, -16>, <chair, 3.15 m, -16>, <window, 4.30 m, 19> Command:

move(5.5) # Move forward towards the chair that is 5.36 meters away.
 Objects in the scene: <chair, 0.96 m, -10>, <chair, 0.99 m, 20>, <chair, 1.77 m, -20> Command:

move(1) # Move forward to approach the chairs.

[snapshots from Habitat navigation]

用对话的方式实现感知-动作循环

资料来源: 《ChatGPT for Robotics: Design Principles and Model Abilities》, 微软 (2023), 华泰研究

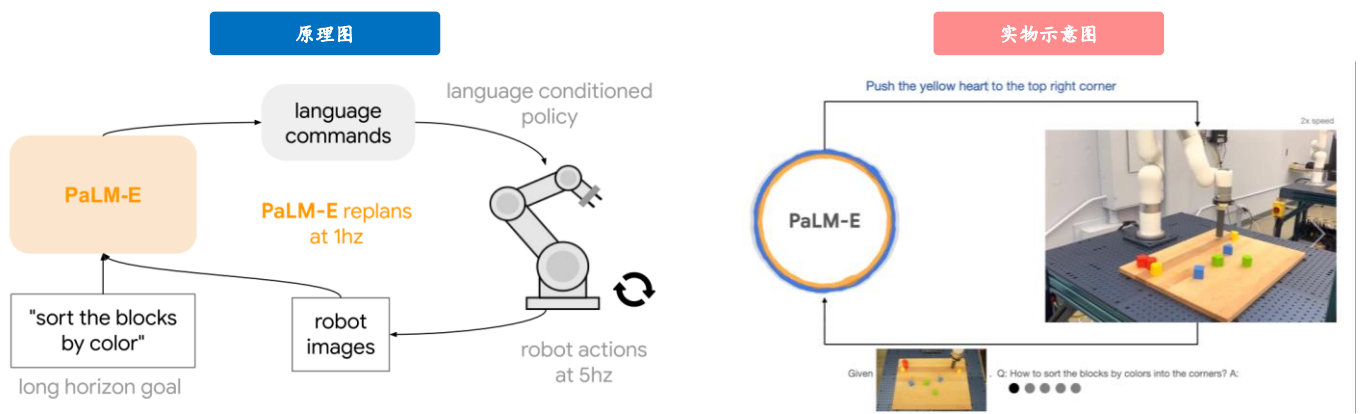
以机器人为突破口，LLM 定义智能终端新入口

ChatGPT 等 LLM 将成为定义智能终端的新入口。ChatGPT 结合了自然语言理解、代码生成等优点，并具有对话方面的自然性和灵活性。目前，ChatGPT 已在个人电脑（PC）和移动端入口上实现了较为全面的布局，包括办公软件（微软 365 Copilt）、协同软件（微软 Teams）、移动端搜索引擎（微软 Bing）、创意软件（Adobe Firefly）。微软的研究已经证明，ChatGPT 能够借助机器人这一入口扩展应用场景。未来，ChatGPT 等 LLM 或将依托更多的现实载体，定义智能终端新入口，扩展其设备入口属性。

入口#1：人形、协作等各类机器人

谷歌的 PaLM-E 具身多模态模型同样实现了对机器人的控制。谷歌于 2023 年 3 月基于其已有的大语言模型 PaLM，提出了能够实现具身化和多模态的新模型 PaLM-E，并实现了对机器人的控制（参考华泰研报：《GPT 系列：具身化多模态 PaLM-E》，2023 年 3 月 14 日）。在机器人的控制流程中，PaLM-E 模型基于人类用户的目标意图，生成机器人控制指令，并且能够接受机器人的反馈信息以对下一步控制指令进行调整和重规划。与微软的 ChatGPT for robotics 相比，PaLM-E 更注重模型本身的对多模态输入的理解以及推理输出能力，而微软的成果更注重人类在机器人流程中的 on-the-loop 支持。

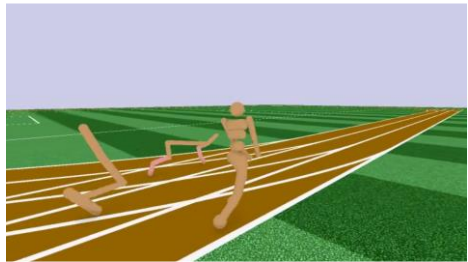
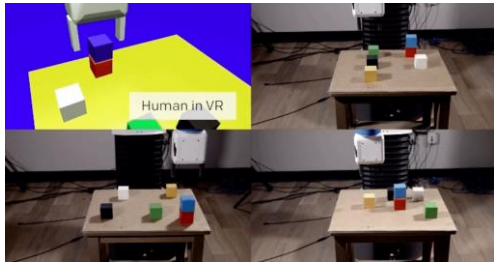
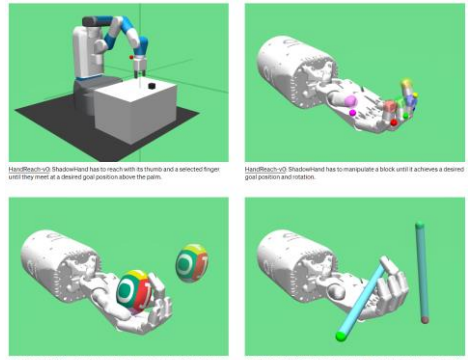

图表 11：PaLM-E 参与到机器人的控制循环中



资料来源：《PaLM-E: An Embodied Multimodal Language Model》，Danny Driess 等（2023）、华泰研究

OpenAI 在机器人领域早有布局。2017 年 5 月，OpenAI 推出了用于机器人模拟的开源软件 Roboschool。同月，OpenAI 在机器人中部署了新的单样本模仿学习（one-shot imitation learning）算法，通过人类在 VR 中向机器人演示如何执行任务，机器人即可从任意起始配置解决相同的任务。2018 年 2 月，OpenAI 发布了 8 个模拟机器人环境和 Hindsight Experience Replay（事后经验回放，HER）基线实施，并用来训练在物理机器人上工作的模型。2023 年 3 月 23 日，挪威的机器人制造商 1X technologies 宣布完成 2350 万美元的 A2 轮融资，领投方是 OpenAI 旗下的启动基金。1X technologies 将利用融资加快其即将推出的双足机器人模型 NEO 的研究进展，并在挪威和北美规模化生产首款商用机器人 EVE。

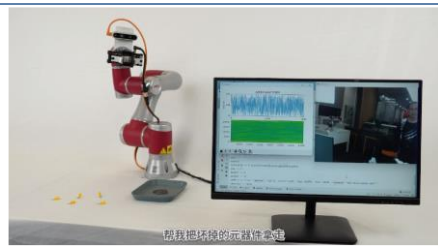
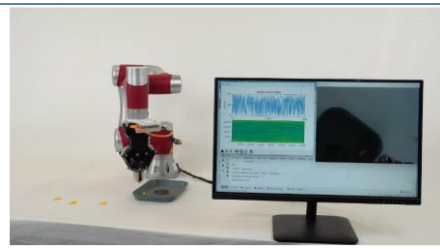
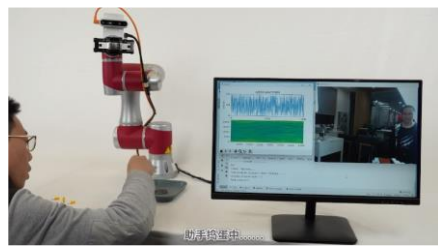
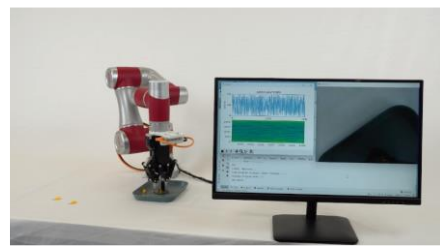
图表12: OpenAI 在机器人上的布局情况

<p>开源软件Roboschool</p>  <p>2017年5月, OpenAI推出了用于机器人模拟的开源软件Roboschool</p>	<p>单样本模仿学习算法</p>  <p>2017年5月, OpenAI在机器人中部署了新的单样本模仿学习算法</p>
<p>发布模拟环境与HER</p>  <p>2018年2月, OpenAI发布了8个模拟机器人环境和Hindsight Experience Replay (HER) 基线实施</p>	<p>领投1X technologies</p>  <p>2023年3月, OpenAI旗下的启动基金领投挪威的机器人制造商1X technologies</p>

资料来源: OpenAI 官网、1X technologies 官网、华泰研究

节卡机器人实现了 **ChatGPT 与产品的结合和应用演示**。节卡机器人是新一代协作型机器人提供商, 其工程师将 ChatGPT 与节卡机器人本体、编程、二次开发, 以及公司客服支持等板块结合, 实现了部分实用性的应用演示。原理上看, 采用语音识别技术将用户的自然语言实时转换成文本, 节卡机器人对转换的文本内容进行理解, 进而完成语音指令到运动控制的转换, 智能化的实现了目标任务, 并且对人类的对抗性干扰具有良好的鲁棒性。

图表13: 节卡机器人与 ChatGPT 结合实现损坏元器件挑选

 <p>跟我把坏的元器件挑出来</p>	
<p>1 语音告知机器人手臂, 对损坏零件进行挑选</p>	<p>2 机器人手臂执行任务, 将损坏元器件挑选出来</p>
 <p>助手换道中.....</p>	
<p>3 人类对机器人实行对抗性干扰</p>	<p>4 机器人在人类提示下识别出人类干扰, 并重新进行任务处理</p>

资料来源: 节卡机器人官网、华泰研究

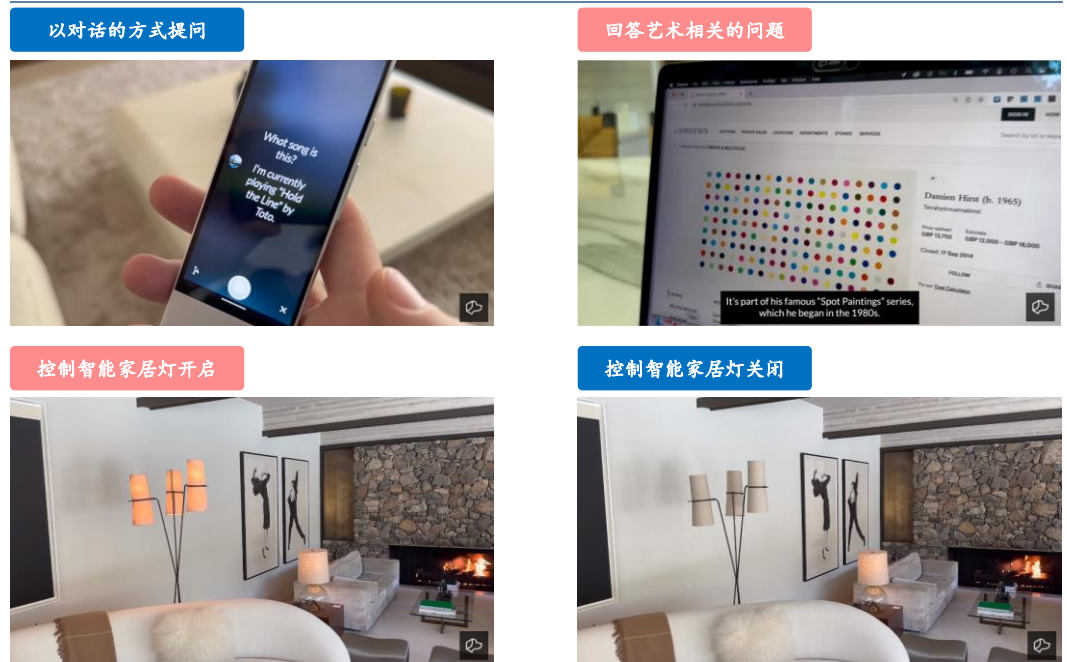
入口#2：以智能音箱为代表的智能家居

智能音箱是重要的家居控制入口，ChatGPT 的理解推理能力将对其赋能。智能家居通过物联网技术连接家中的各种设备（如音视频设备、照明系统、窗帘控制、空调控制、安防系统、数字影院系统、影音服务器、影柜系统、网络家电等）。其中，智能音箱是重要的家居控制入口。目前的智能音箱等语音产品往往只能实现简单一问一答功能，或者重复咨询、确认、沟通的循环中，交互方式过于“理性”和僵化。而 ChatGPT 具有很好的语义理解和推理能力，能够对用户下达的长指令进行逻辑分析和规划执行，并且支持向用户提供更多建议。因此，ChatGPT 与智能音箱结合后，将赋予其更多“智能”。

Josh.ai 公司致力于开发基于 ChatGPT 的语音集成。Josh.ai 是国外的智能家居提供商，开发了语音控制家庭自动化系统 Josh 和 Josh Core/Micro/Nano 等智能家居产品。目前，Josh.ai 提供了与 ChatGPT 结合的使用示例，例如控制音乐播放、控制灯具启闭、回答用户问题等，交互过程更加智能和自然。Josh.ai 指出潜在用例还包括：1) 根据住家的位置和家庭特质提供个性化的回答；2) 提供购物、户外活动、艺术等用户感兴趣的各种服务；3) 帮助用户学习和理解新知识；4) 为用户制定各种活动列表，并支持添加费用、环境等约束条件；5) 帮助用户实现对智能家居的控制。

Josh.ai 正在与亚马逊合作，进一步开发家庭智能系统潜能。据亚马逊 CES 2023 大会信息，亚马逊正在与 Josh.ai 合作，开发 Josh.ai 家庭智能系统的多助手集成潜力。当 Josh.ai 助手控制智能家居或连接设备时，用户将能够同时使用 Alexa（亚马逊公司推出的智能助理）处理各种语音请求，例如查看最新的体育赛事比分和管理购物清单。

图表14： Josh.ai 与 ChatGPT 结合实现控制与问答功能



资料来源：Josh.ai 官网、华泰研究

阿里大模型联合项目团队在天猫精灵智能音箱上部署了个性化消费级终端模型。阿里团队以通义大模型为基础，量身脱口秀演员打造了个性化数字分身“终端模型”，并部署在天猫精灵智能音箱上，实现了：1) 不间断自由语音对话；2) 个性化音色、语气、表达方式的模拟；3) 向用户提供百科、创作灵感等信息和知识。

个性化应用是大模型研究的重要方向。智能家居具有“千人千面”的特性，个性化大模型能够更好的实现对智能家居的“智能”加持。阿里指出个性化大模型在对话场景训练中关注的重点包括：1) 人设、对话风格、逻辑的一致性；2) 对话三观的正确性；3) 有偏好的个性化对话实现。通过大规模语言训练、知识和工具增强、个性化对话增强和人类反馈增强等四个步骤，个性化大模型被赋予了相应的共情能力，能够更和谐的融入家庭环境。

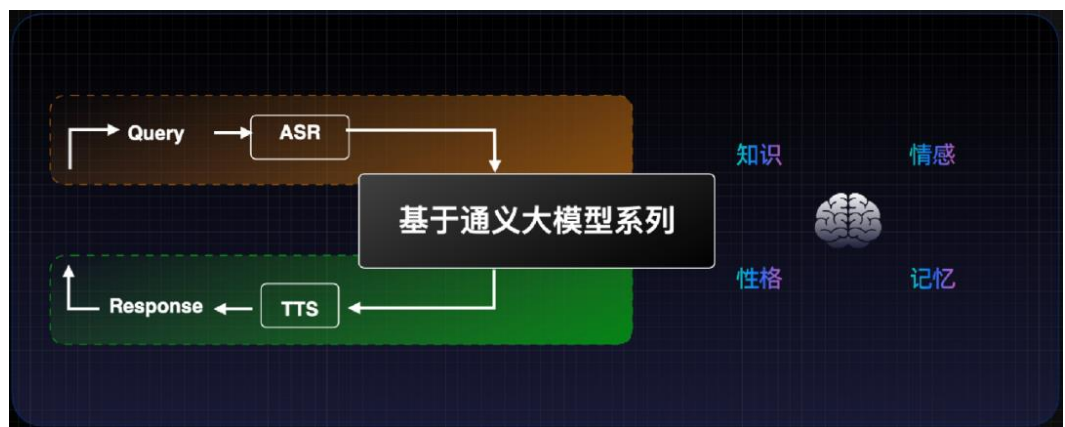
图表15：阿里个性化大模型训练步骤



资料来源：阿里官网、华泰研究

通义大模型、TTS 与 ASR 技术实现个性化大模型的知识、情感、记忆和性格四位一体。阿里团队以通义大模型系列为模型基础，融合了 ASR 猫耳算法、自研的声学模型 KAN-TTS 定制方案，实现了个性化模型在知识、情感、记忆和性格上的和谐统一。1) **ASR 猫耳算法**：结合深度学习方法与传统 AEC、多通道立体声消除回声，并借助设备中的麦克风阵列实现定向拾音，准确识别说话人位置，将用户语音精确转换成文本，同时利用降噪消除非人声或非主要说话人声。2) **KAN-TTS 定制方案**：基于达摩院自研声学模型，只需收集 1 个小时的个性化录音，即可在一周内完成录音、训练、模型上线全流程，且呈现出的拟人化声音更自然的接近真实人声。

图表16：阿里个性化大模型工作流程



资料来源：阿里官网、华泰研究

图17: 阿里 KAN-TTS 人声定制方案技术示意图



资料来源：阿里官网、华泰研究

国内多家智能家居公司已宣布接入百度文心一言。目前，三大家电巨头美的、海信影像、长虹已经官宣成为百度旗下类 ChatGPT 应用“文心一言”的首批生态合作伙伴，百度自身智能音箱小度也将应用文心一言技术。此外，据各公司年报或招股书，萤石网络、漫步者、国光电器等均在智能家居和人工智能上进行了布局：

1) **萤石网络**：公司积累了丰富的视音频 AI 算法技术，能够提供算法生成、算法调度等 AI 服务，产品方面，智能中控屏 SP4 支持 AI 语音控制，萤石 AI 扫拖一体机 RS2 支持拖把布自动装卸、干湿分离扫地毯等。

2) **漫步者**：针对智能家居日益普及的趋势，2022 年公司在国内外更新多款智能音箱，并协同互联网平台推出搭载 AI 语音助手的 TWS 耳机；推出了全新形态、具备“AI 听损智能补偿”技术的助听设备。

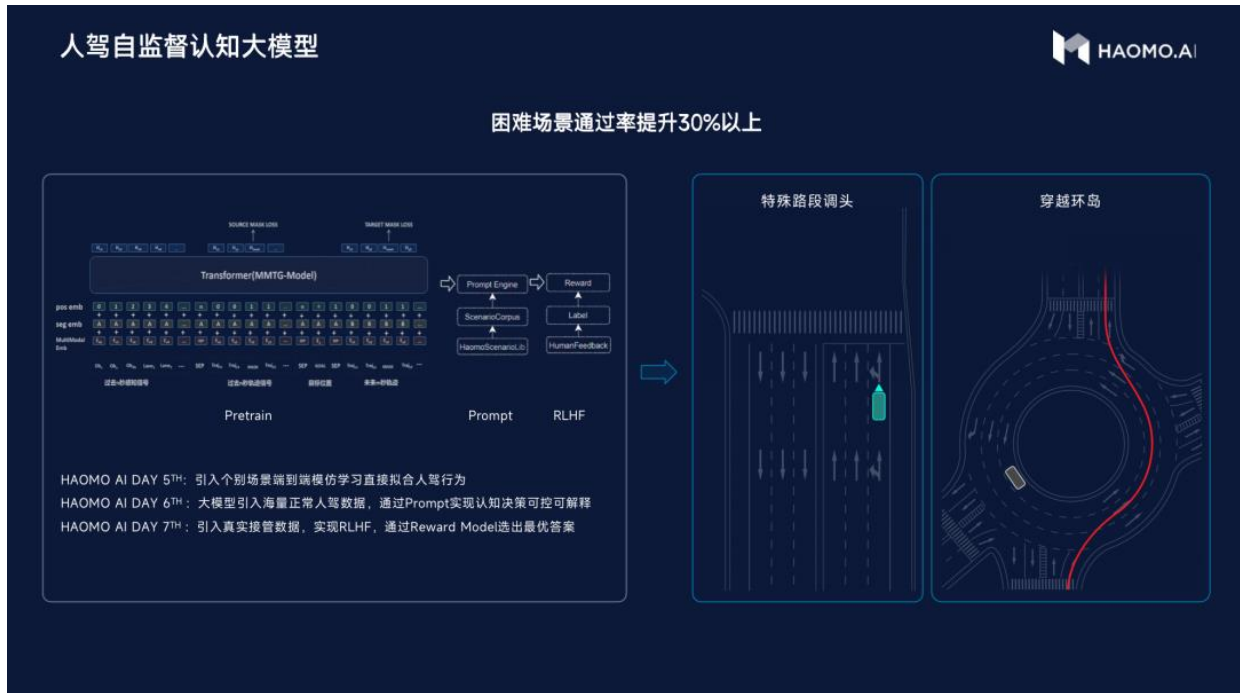
3) **国光电器**：2023 年 2 月 10 日，国光电器在投资者互动平台上指出，公司是人工智能语音交互硬件的供应商，与百度在多个项目上开展深度合作，其中以智能音响为代表的 AI+IoT 产品是公司重点合作方向之一。

入口#3：车载虚拟个人 AI 助手

ChatGPT 等 LLM 在车载场景下能力优势相较传统 AI 助手更明显。语音交互是车内简洁、人性化、安全的交互方式，也是未来主要的车内交互方式，而 ChatGPT 等大模型能够通过上下文理解能力、思维链推理、增强指令等实现持续学习，提供更灵活、更人性化、更精准的车机交互，相比传统的 AI 助手更能快速适应复杂的车载交互环境。

通用汽车正开发基于 ChatGPT 的虚拟个人 AI 助手。据 2023 年 3 月的 Semafor 数据，通用汽车公司正在开发一款虚拟个人 AI 助手，该助理使用 ChatGPT 人工智能模型和微软的 Azure 云服务。通用汽车副总裁 Scott Miller 指出，AI 助手可能会超越简单语音命令，更智能的帮助用户处理汽车相关问题，例如，轮胎损坏时用户可以要求汽车解释如何更换轮胎，或在车内的显示屏上播放教学视频；仪表盘上出现诊断灯，司机可以询问 AI 助理如何处理，或者推荐维修店并进行预约。在此基础上，通用汽车将在 ChatGPT 模型中添加与汽车相关的层 (layer)，以便更好的处理与汽车相关的问题。

国内多个车企宣布接入百度文心一言。除百度自家的 Apollo 自动驾驶平台外，集度汽车宣布旗下产品全面融合百度文心一言能力，支持汽车机器人实现自然交流的再进阶。毫末智行也宣布将其自动驾驶认知大模型正式升级为采用 RLHF（人类反馈强化学习）技术的“人驾自监督认知大模型”DriveGPT，通过引入真实人驾接管数据，对自动驾驶认知决策模型进行持续优化，并计划于 2023 年 4 月第八届 HAOMO AI DAY 上公布进展。此外，长城、吉利、红旗、零跑、东风日产启辰、岚图、爱驰等多家汽车品牌也已官宣接入“文心一言”。

图表18: 毫末智行的人驾自监督认知大模型将困难场景的通过率提升 30%+


资料来源: 毫末智行官网、华泰研究

相关标的

大模型的“新入口”属性已经从主流的 PC 和手机端, 向更广泛的智能设备扩散。我们认为, 主要的智能设备包括智能终端和智能音箱, 相关标的包括:

- 1) **智能终端**: 科大讯飞, 萤石网络, 中科创达等;
- 2) **智能音箱**: 国光电器, 漫步者, 惠威科技等。
- 3) **机器人**: 石头科技, 柏楚电子;
- 4) **银行超级柜台**: 新北洋。

图表19：提及公司列表

公司代码	公司简称
MSFT US	微软
未上市	OpenAI
GOOG US	谷歌
BIDU HK	百度
未上市	Josh.ai
BABA US	阿里巴巴
002230 CH	科大讯飞
688475 CH	萤石网络
300496 CH	中科创达
002045 CH	国光电器
002351 CH	漫步者
002888 CH	惠威科技
未上市	毫末智行
688188 CH	柏楚电子

资料来源：Bloomberg、华泰研究

风险提示

宏观经济波动。若宏观经济波动，产业变革及新技术的落地节奏或将受到影响，宏观经济波动还可能对 AI 投入产生负面影响，从而导致整体行业增长不及预期。

技术进步不及预期。若 AI 技术和大模型技术进步不及预期，或将对相关的行业落地情况产生不利影响。

本报告内容均基于客观信息整理，不构成投资建议。

免责声明

分析师声明

本人，谢春生，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。

一般声明及披露

本报告由华泰证券股份有限公司（已具备中国证监会批准的证券投资咨询业务资格，以下简称“本公司”）制作。本报告所载资料是仅供接收人的严格保密资料。本报告仅供本公司及其客户和其关联机构使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司及其关联机构（以下统称为“华泰”）对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰不保证本报告所含信息保持在最新状态。华泰对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，华泰及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。华泰不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰可能会持有报告中提到的公司所发行的证券头寸并进行交易，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰的销售人员、交易人员或其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。华泰没有将此意见及建议向报告所有接收者进行更新的义务。华泰的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。投资者应当考虑到华泰及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。有关该方面的具体披露请参照本报告尾部。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人（无论整份或部分）等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司制作，在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题，请与华泰金融控股（香港）有限公司联系。

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。
- 中科创达（300496 CH）、柏楚电子（688188 CH）：华泰金融控股（香港）有限公司、其子公司和/或其关联公司实益持有标的公司的市场资本价值的 1%或以上。
- 有关重要的披露信息，请参华泰金融控股（香港）有限公司的网页 https://www.htsc.com.hk/stock_disclosure 其他信息请参见下方“美国-重要监管披露”。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934 年证券交易法》（修订版）第 15a-6 条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受 FINRA 关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师谢春生本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括 FINRA 定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 科大讯飞（002230 CH）、中科创达（300496 CH）：华泰证券股份有限公司、其子公司和/或其联营公司在本报告发布日之前的 12 个月内担任了标的证券公开发行或 144A 条款发行的经办人或联席经办人。
- 中科创达（300496 CH）：华泰证券股份有限公司、其子公司和/或其联营公司在本报告发布日之前 12 个月内曾向标的公司提供投资银行服务并收取报酬。
- 中科创达（300496 CH）、柏楚电子（688188 CH）：华泰证券股份有限公司、其子公司和/或其联营公司实益持有标的公司某一类普通股证券的比例达 1%或以上。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。

评级说明

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力（含此期间的股息回报）相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数），具体如下：

行业评级

- 增持：** 预计行业股票指数超越基准
中性： 预计行业股票指数基本与基准持平
减持： 预计行业股票指数明显弱于基准

公司评级

- 买入：** 预计股价超越基准 15%以上
增持： 预计股价超越基准 5%~15%
持有： 预计股价相对基准波动在-15%~5%之间
卖出： 预计股价弱于基准 15%以上
暂停评级： 已暂停评级、目标价及预测，以遵守适用法规及/或公司政策
无评级： 股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

法律实体披露

中国: 华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格, 经营许可证编号为: 91320000704041011J
香港: 华泰金融控股(香港)有限公司具有香港证监会核准的“就证券提供意见”业务资格, 经营许可证编号为: AOK809
美国: 华泰证券(美国)有限公司为美国金融业监管局(FINRA)成员, 具有在美国开展经纪交易商业业务的资格, 经营业务许可编号为: CRD#:298809/SEC#:8-70231

华泰证券股份有限公司**南京**

南京市建邺区江东中路228号华泰证券广场1号楼/邮政编码: 210019

电话: 86 25 83389999/传真: 86 25 83387521

电子邮件: ht-rd@htsc.com

深圳

深圳市福田区益田路5999号基金大厦10楼/邮政编码: 518017

电话: 86 755 82493932/传真: 86 755 82492062

电子邮件: ht-rd@htsc.com

北京北京市西城区太平桥大街丰盛胡同28号太平洋保险大厦A座18层/
邮政编码: 100032

电话: 86 10 63211166/传真: 86 10 63211275

电子邮件: ht-rd@htsc.com

上海

上海市浦东新区东方路18号保利广场E栋23楼/邮政编码: 200120

电话: 86 21 28972098/传真: 86 21 28972068

电子邮件: ht-rd@htsc.com

华泰金融控股(香港)有限公司

香港中环皇后大道中99号中环中心58楼5808-12室

电话: +852-3658-6000/传真: +852-2169-0770

电子邮件: research@htsc.com

<http://www.htsc.com.hk>**华泰证券(美国)有限公司**

美国纽约公园大道280号21楼东(纽约10017)

电话: +212-763-8160/传真: +917-725-9702

电子邮件: Huatai@htsc-us.com

<http://www.htsc-us.com>

©版权所有2023年华泰证券股份有限公司