

第 1 章 引 言

“基本上，设计是将未知或看不见的想象变成存在的过程。”^[1]

——莎伦·波根波尔（Sharon Poggenpohl）

人机对话交互可以看作一个复杂系统，其中很多部分对我们而言都是未见的、未知的。仅仅使用工程思维，或者仅仅研究算法，无法实现成熟的人机对话交互。只有通过对手上已有的设计材料进行不断研究，去想象新的可能性，通过实验原型系统测试不同的设计效果，逐渐把未知的部分变为已知，这个过程才是本人所认为的人机对话交互中的设计过程，其重点是要为 CUI 的对话行为设计提供实证依据。

1.1 研究背景

1.1.1 对话交互界面

对话交互界面（Conversational User Interface, CUI）的设计研究属于人机交互研究领域。人机交互（Human-Computer Interaction, HCI）实际上是人和计算机交换信息的过程。计算机底层的机器语言由二进制的 0 和 1 代码组成，这和人们对话使用的自然语言之间存在巨大的差距，因此，人机交互设计利用人类的各种感官和行为能力来搭建与计算机交换信息的桥梁。1973 年出现的以鼠标、键盘、显示器组成的图形交互界面（Graphic User Interface, GUI）一直沿用至今，形成了一系列桌面电脑经典的交互范式。移动智能设备的普及将 GUI 推向了以触控为核心的新范

式,输入设备方面舍弃鼠标和键盘,改用手指触控,使得交互更简单自然。但 GUI 仍需要用户对交互方式进行一定学习,交互设计在追求自然的目标上还有很长的道路要走。人工智能技术的发展催化了许多更为自然的交互形式,包括实体交互界面(Tangible User Interface, TUI)、语音交互界面(Voice User Interface, VUI)、对话交互界面等。“界面”(Interface)一词,已经不再仅限于一个可见的实体面板,而是在用户与计算机之间承载交互信息的媒介的总称。

近年来,对话交互界面(CUI)成了热点。对话交互界面指与用户进行仿真对话的计算机交互界面,其特点是允许用户以自然语言与计算机交互,就像人们日常进行的对话一样。对话交互降低了用户对界面的学习难度,使人机交互更为自然。注意本书所指的 CUI 包括多种人机对话的模式和方式,常见的语音助手和文字聊天机器人(Chatbot)也都属于本书定义的 CUI 范畴。为了方便表述,本书中出现的“用户与 CUI 进行对话”的一类说法实际是指用户通过 CUI 与计算机进行对话。

与计算机进行对话是人们长久以来的梦想,对话的交互形式可以改变人与机器的关系。就像科幻电影 *Her*^①(2013 年)男主角 Theodore 每日与陪伴他的人工智能系统 Samantha 对话交谈,最终对 Samantha 产生了复杂的情感。电影探讨了未来人与机器更为密切、更为复杂的社会关系。

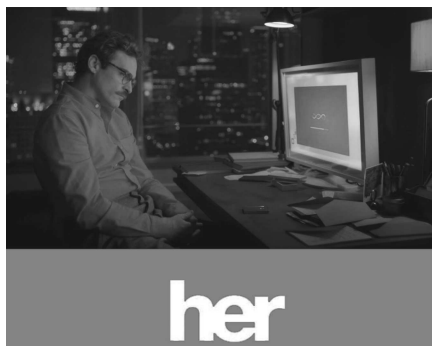


图 1.1 电影 *Her* (2013) 海报(见文前彩插)

实际上, CUI 在几十年前就出现了。从 1966 年第一个聊天机器人 ELIZA^[2] 开始,到 20 世纪末的自动语音应答(Interactive Voice Response,

^① 电影 *Her* 的 imdb 主页 <https://www.imdb.com/title/tt1798709/>

IVR),时至今日,基于人工智能的对话系统已经深入到了人们的日常生活当中。目前市场上几乎每一台智能设备都会搭载语音助手,比如 iPhone 的 Siri、亚马逊 Echo 的 Alexa、小米手机的小爱同学、微软 Windows 的 Cortana 等。除了语音助手,文本形式的聊天机器人也被广泛用于线上购物的客服、博物馆的导览等。对话交互界面的设计近年来也成为热门的研究方向。许多设计研究探索将对话交互界面应用于教育^[3-7]、车内^[8-12]、健康^[13-17]等各种场景的可能性。许多研究也专注于一些 CUI 中的交互细节,例如唤醒机制^[18-21]、多模态对话^[22-24]、对话情绪^[25-27]等。

1.1.2 CUI 设计面临的挑战

本书基于文献研究提出,CUI 设计当前主要面临以下两个方面的挑战:一是缺乏考虑多层次用户体验的设计方法;二是人机对话交互的形式亟待创新和多元化设计。

目前 CUI 的书籍和著作主要以开发经验的总结为主,如 *The Conversational Interface: Talking to Smart Devices*^[28]。这类书籍非常具有参考价值,但其阐述的重点是开发经验,而非以用户体验为目标的设计方法。

相比于 CUI,图形用户界面 GUI 的设计发展得更为成熟,有研究者提出将 GUI 的设计准则迁移到 CUI 上。但 GUI 与 CUI 的许多界面特性有显著区别,无法直接套用,因此 Murad 等人提出将 GUI 中的易用性原则(Usability Heuristics)经过调整后再迁移到 CUI^[29-30]。类似地,Langevin 等人提出将 Nielsen 设计原则应用在 CUI 设计中^[31],Yang 和 Aurisicchio 提出基于自我决定理论(Self-Determination Theory)的 CUI 设计原则^[32]。IBM 研究组的 Moore 等人提出的 Alma 设计框架^[33]则更关注 CUI 的实践,框架包含设计准则、交互模式、内容模板三部分。

但以上的 CUI 设计理论主要还是关注与功能相关的易用性体验,对用户体验的考虑不够多元。CUI 在为用户提供服务时,其交互除了完成基本的对话任务,还具有许多社交属性,这意味着用户体验同样会受到易用性以外的其他方面的影响。例如,教学场景中的 CUI 如果无视用户当前的情绪状态,在后者表现出沮丧情绪时依然强迫用户继续完成课程,就会带来糟糕的用户体验;如果访谈场景中的 CUI 总是提一种类型的问题,

用户就会认为其智能程度太低,失去与之交谈的兴趣。Norman 的情感化设计理论^[34]是人机交互设计领域具有代表性的关注用户体验的理论,该理论强调设计应当考虑用户体验的三个层次,即本能层、行为层、反思层,而目前的 CUI 设计的关注点局限在行为层,即功能性和易用性相关的体验。本书基于 Norman 的理论,认为 CUI 设计需要更多层次地考虑用户体验。

本书还认为,目前 CUI 的人机对话交互形式缺乏创新性和多元性。目前常见的对话交互产品主要遵循“请求-回复”的对话形式,允许用户说出一些常用的“请求”语句来完成咨询或者下达命令,例如,用户向语音助手询问天气、时间,或者下命令来启动家中的空调。在没有具体目的的闲聊中,产品会根据用户的话语回复一句具有一定相关性的话语,但这类闲聊很难形成更深入、可以达成具体目标的对话。而人与人的自然对话形式远远比这两种要丰富多元。

一些 CUI 产品商为第三方设计师和开发者提供了设计指导,例如 Google Assistant 基于对话的合作原则^[35]总结了 CUI 设计准则^①,Amazon Alexa 基于自身的对话能力和功能(Alexa Skill)为开发者提供了设计手册^②。这些手册允许设计师和开发者在既有的对话形式中替换部分对话内容,但对话交互的形式还是只能基于产品所提供的形式。除了商用 CUI 产品的设计手册,许多 CUI 的研究也是基于或面向已有产品的交互形式开展的,例如,针对居家环境中的智能音箱、智能手机中的语音助手的设计研究^[36-41]。这些针对既有产品的设计指导手册和相关研究,也受限于产品的功能和对话交互形式。

埃因霍温理工大学的 Lee 等人对目前 CUI 设计关注功能性的“功利主义”进行了批判,并提出了人机对话交互缺乏多元性的问题,呼吁 CUI 设计突破“请求-回复”的形式,考虑不同类型用户的需求,并考虑对话中的情感等因素进行设计^[42-43]。Völkel 等人的对话激发实验也表明,用户对 CUI 的需求不仅仅在于功能性的对话,还希望 CUI 可以进行更主动、交互更强的对话,可以更了解用户,具备幽默感,可以表达见解等^[44]。

① Google Assistant 的设计手册, <https://developers.google.com/assistant/conversation-design/>

② Amazon Alexa 的设计手册, https://build.amazonalexadev.com/What_is_voice_user_interface_VUI.html

除此以外,研究者还提出,目前 CUI 对用户需求的考虑缺乏包容性,对话管理方式缺乏多样性等^[42,45]。基于此,本书认为 CUI 需要更具创新性和多元性的对话交互形式设计。

1.2 研究问题

根据 CUI 设计所面临的挑战,本书提出以下两个主要的研究问题:

- 研究问题 1: 如何在具体场景的 CUI 设计中多层次地考虑用户体验?
- 研究问题 2: 如何探索人机对话形式的多元性?

本书的研究将围绕探索一种在具体场景中使用的 CUI 设计方法而展开,在不同的层次考虑用户体验,而不仅是注重和功能相关的易用性;同时探索有助于对话交互形式创新的设计方法,促进人机对话交互的多元化。

1.3 研究内容

本书强调理论和实践的结合。在理论层面,本书首先对 CUI 的特性进行分析,归纳出自然人机对话三方面的特性,即不确定性、多元性、系统性。随后提出层次化用户体验的 CUI 设计方法,该方法包括层次化 CUI 用户体验的概念、CUI 交互流程框架,以及将前两者结合的 CUI 设计流程。

本书基于情感化设计理论,提出在 CUI 设计中层次化用户体验的概念,即考虑将 CUI 的用户体验分为三个层次:直觉感受层、对话功能层以及对话认知层。其中,直觉感受层是用户对使用 CUI 直觉性的感受体验,受 CUI 的外观、音色等因素影响;对话功能层是用户在使用 CUI 时与其功能相关的易用性体验,受 CUI 基本的对话理解能力、对话任务完成的质量等因素影响;对话认知层是用户在使用 CUI 一段时间之后,反思和总结其使用体验后形成的认知,包括用户对 CUI 的智能程度和价值认可、用户与 CUI 形成的关系等。在此基础上,本书提出 CUI 的交互流程框架,并对部分设计维度和变量进行层次化用户体验的分析。最后提出层次化用户体验的 CUI 设计流程,包括发现、定义、构思、验证四个阶

段，为 CUI 设计师的理论研究和实践应用提供参考。

在实践层面，本书以所提出的设计方法为基础，进行了两项涉及不同对话场景的研究，分别是智能访谈场景和任务协助场景。选择这两个场景的原因是：（1）场景适合使用 CUI 进行交互；（2）场景需要研究新的对话交互形式；（3）场景相关的其他研究文献缺乏多层次用户体验的考虑。两项研究均涉及多个用户体验层次，体现了考虑多层次用户体验的思想。

第一项研究面向智能访谈的对话场景，CUI 将对用户展开半结构化访谈，收集与话题相关的用户经历和观点。目前已有的访谈 CUI 主要是基于提前预设的固定问题进行访谈，而人类访谈者经常会使用追问行为来获取更多信息，因此该研究将 CUI 的追问行为设计作为核心的设计变量。基于文献中访谈专家总结的追问技巧，研究搭建了可以自动完成访谈和追问的 CUI 原型系统，原型系统可以在访谈中提出三种类型的追问，分别是直接追问、关联追问、通用追问。研究基于该原型系统进行了用户实验测试，实验评估了不同追问设计对多个层次用户体验的影响。实验通过测量被试对于不同类型追问的回答意愿，来评估直觉感受层的用户体验；通过计算不同类型追问获取到的信息量，来评估对话功能层的用户体验；通过收集实验后被试对不同类型追问的流畅性和相关性评价，来评估对话认知层的用户体验。实验结果表明，此次设计的 CUI 追问方法可以有效获取访谈信息，并获得良好的用户评价。具体来说，直接追问和关联追问具有更高的回答意愿和更好的相关度评价（即更好的直觉感受层和对话认知层体验），而通用追问可以引导用户表达出更多信息（即更好的对话功能层体验）。研究基于实验结果提出了相应的优化设计方案。本项研究面向智能访谈场景对本书所提出的设计方法进行了实践，在该场景下对其可行性进行了验证。

第二项研究面向任务协助场景，CUI 将协助用户完成一系列计算任务。CUI 会根据任务中的用户情绪状态调整对话的回复内容，从而让用户感知到 CUI 的情感智能，并以更积极的情绪状态完成任务。研究针对目前 CUI 缺乏闭环情感体验设计的问题，提出将用户情绪的感知和反馈结合的闭环设计。具体来说，就是基于语音情绪识别和共情原则来进行情绪反馈的 CUI 设计。研究还开发了名为 Heard yoUr Emotion，即 HUE 的对话交互系统。HUE 的情绪反馈方式一共有三种，分别是使用共情语

气词、使用情绪调节策略的语句，以及两者结合的方式。根据 HUE 的设计，研究选择可行的对话交互技术开发原型系统，并基于原型系统开展用户实验。实验结果显示，HUE 的情绪反馈设计可以有效提高用户对于 CUI 的情感智能评价，并且减缓用户在任务中产生的负面情绪。具体来说，使用共情语气词来反馈情绪可以更有效地减轻用户的负面情绪（即更好的直觉感受层体验）；而在用户处于焦虑状态中，情绪调节语句的反馈方式获得了更高的情感智能评价（即更好的对话认知层体验）。本项研究面向任务协助场景对本书所提出的设计方法进行了实践，在该场景下对其可行性进行了验证。

上述两项研究旨在探索两类具有创新性和多元性的人机对话交互：CUI 更丰富的提问方式；CUI 对用户情绪状态的感知和反馈，为设计师提供了在具体场景中使用本书所提出的设计方法来多层次考量用户体验的实践参考。

1.4 研究贡献

本书的贡献主要分为理论贡献和实践贡献。理论贡献如下：（1）根据对 CUI 的分析，归纳出三方面自然人机对话特性：不确定性、多元性、系统性；（2）根据情感化设计理论，提出在 CUI 设计中将用户体验分为直觉感受层、对话功能层、对话认知层三个层次的概念；（3）提出了有利于促进人机对话交互形式创新和多元化的 CUI 设计方法，包括层次化用户体验的交互流程框架和设计流程。

本书的实践贡献如下：（1）基于所提出的理论方法进行了两项涉及不同对话场景的研究，对所提出的设计理论方法在具体的场景中进行了可行性验证；（2）所进行的研究为设计师使用本书所提出的设计理论方法提供了实践应用的具体参考。

1.5 研究创新点

目前有关 CUI 的研究过于“功利主义”，更多的是关注易用性，缺少多层次的用户体验考量。设计学领域中 Don Norman 的情感化设计理

论^[34]被广泛地应用到各个场景当中,其对于用户体验层次化考量的概念颇具代表性。将情感化设计理论应用到 CUI 中的思路,在之前的语音交互界面(VUI)相关研究中也出现过。李真真等人通过讨论情感化设计理论在智能移动设备 VUI 中的使用思路,提出了一些设计建议^[46]。刘佳萌则更为系统地将情感化设计层次和目前的智能音箱语音交互设计结合起来,提出了智能音箱语音交互的情感化设计原则和策略^[37]。张雪从用户心智模型的角度,将情感化设计层次与智能音箱的语音交互体验结合^[36]。但这些研究更多是关注设计理论模型,尚未形成具体的设计方法并进行实践研究。相比于已有的研究工作,本书具有以下三个创新点:

首先,本书在 CUI 设计方法中融入了层次化用户体验的概念。已有研究更多关注如何把情感化设计融入设计理论层面,重点在于用户需求和心智模型的研究,主要产出设计准则和理论模型。而本书提取了情感化设计中用户体验层次的思想,并将其融入具体的 CUI 设计方法中。

其次,本书提出了可供实践参考的 CUI 设计方法。本书强调理论与实践的结合,将所提出的设计方法融入 CUI 设计的实践中,形成了在具体场景中具有一定实用性的 CUI 设计方法。此外,本书还记录了设计方法实践过程中的细节,可供设计师和研究者参考。

最后,本书促进了人机对话形式的多元化发展。此前的 CUI 研究主要面向的是智能音箱或语音助手等目前市面上已有的语音交互产品。本书的研究不受限于已有产品的交互形式,所提出的交互流程框架和设计流程,有助于促进对话交互形式的多元化创新。相应地,本书在具体场景中的研究均探索了具有创新性的对话交互形式。

1.6 论文结构

第 1 章为引言,简要介绍本书的研究背景,提出主要的研究问题,概括研究内容、研究贡献、研究创新点和论文结构。

第 2 章为文献综述,界定“对话”在本书中的定义范围,围绕对话交互界面的相关学科研究进展进行介绍。本书首先介绍社会科学中关于对话的基础研究,接着介绍从最早的聊天机器人到最近的对话式人工智能相关的对话交互界面的历史发展,然后介绍对话交互技术的相关背景,

最后介绍人机交互设计领域关于对话交互界面的前沿研究。

第 3 章对 CUI 的特点进行分析,归纳出自然人机对话的重要特性,为第 4 章设计方法的提出做铺垫。分析围绕三个方面展开:CUI 中的用户心理模型、CUI 易用性的特点,以及 CUI 中设计与算法的辩证关系。

第 4 章提出本书最核心的 CUI 设计方法,该方法包括层次化用户体验的概念、CUI 的交互流程框架和 CUI 设计流程。本章的前沿理论是第 5、6 章进行具体场景研究的基础。

第 5 章阐述在智能访谈场景中的追问对话交互设计研究。首先对研究进行整体概述,其次介绍访谈场景中 CUI 的相关背景,然后介绍前期利用对话分析进行的数据库的搭建工作,基于数据分析和文献调研的追问行为设计和用户实验的开展工作,最后对研究中的实验结果和设计变量进行反思讨论。

第 6 章介绍任务协助场景中 CUI 的情绪反馈对话交互设计研究。首先对研究进行整体概述,其次介绍 CUI 情绪反馈设计的相关背景,然后介绍共情的对话原则调研提出的情绪反馈设计,以及基于原型系统的用户实验,最后对实验结果和设计变量进行反思讨论。

第 7 章基于本书的研究工作展开以下方面的讨论:如何在具体场景的设计和研究中使用层次化用户体验的 CUI 设计方法;如何进行多元化的对话交互设计;对话交互技术发展对未来 CUI 设计的影响。

第 8 章是全书的总结,未来本人将在更多的场景中实践本书提出的 CUI 设计理论,进一步探索具有系统性的多元化 CUI 设计理论。

