

人机交互概述

人机交互历史

- 批处理卡片

- 联机终端（命令行）

- 图形用户界面（GUI）

- 未来的人机交互

人机交互的基础知识

- 交互框架

 - 执行 / 评估活动周期 EEC

- 信息处理模型

 - 人类处理机模型

 - 格式塔心理学

 - 人脑记忆特性

交互设计原则与目标

- 可用性目标

- 用户体验目标（这部分是今年新增的，考到可能性很大）

- 体验和可用性的关系

- 简易可用性工程

- 黄金规则

- 十条启发式规则（括号内为和黄金规则的对应关系）

交互式系统的需求

- 用户建模（很重要）

 - 人物角色

- 需求验证：原型

- 层次化任务分析 HTA

评估

- 什么是评估？

- 评估的原则（小题）

- 评估范型

- 实证研究方法（重要，今年新增）

 - 研究假设

 - 因变量和自变量

 - 实验构成

 - 实验设计

 - 组内设计和组间设计

用户测试

- Decide 评估框架

- 小规模实验

询问用户和专家

- 启发式评估（重点）

交互系统中的设计

- 简化交互设计策略

交互设计模型与理论

- 预测模型

 - GOMS

 - KLM 击键层次模型

 - Fitts 定律

人机交互概述

- 人机交互的英文名：HCI, Human-Computer Interaction
- 定义：HCI 是一门与人类使用的交互式计算系统的设计、评估和实施有关的学科，并且研究了他们周围的主要现象。
- 以用户为中心（UCD, user centered design）的设计方法

人机交互历史

有四个阶段，不是后面出现了前面就消失了。

批处理卡片

- 每次只能由一个用户对计算机进行操作
- 编写程序使用以 01 串表示的机器语言
- 缺点：
 - 不符合人的习惯
 - 耗费时间，又容易出错
 - 只有少数专业人士才能运用自如

联机终端（命令行）

- 缺点：
 - 大部分命令语言对用户输入的要求非常严格
 - 命令名称的缩写在一定程度上减轻了用户的使用负担

图形用户界面（GUI）

- 主要特征：直接操纵
- 图形用户界面优于字符界面？
 - 不同的交互方式本身在可用性方面并没有根本性的不同，更重要的是认真对待界面设计的态度。[Whiteside 1985]
- 优点：依赖识别而非记忆

未来的人机交互

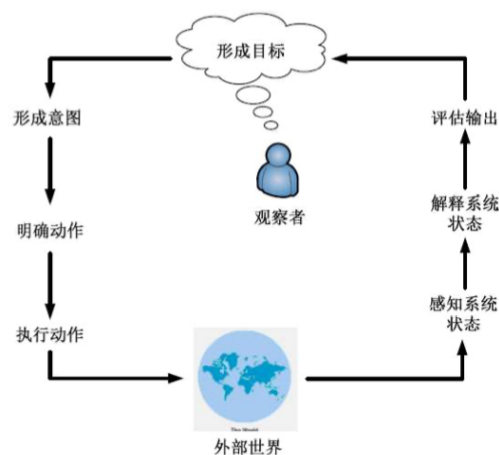
- 多媒体界面
- 多通道交互技术
- 虚拟现实、语音交互、脑机交互

人机交互的基础知识

交互框架

执行 / 评估活动周期 EEC

- 定义了活动的四个组成部分：
 - 目标 (Goal) \neq 意图 (Intention) (单个目标可对应多个意图)
 - 执行 (Execution)
 - 客观因素 (World)
 - 评估 (Evaluation)
- EEC 模型
 - 共有 7 个阶段：1-4 执行阶段；5-7 评估阶段
 - 可解释为什么有些界面的使用存在问题



- 执行隔阂：用户为达目标而制定的动作与系统允许的动作之间的差别
- 评估隔阂：系统状态的实际表现与用户预期之间的差别

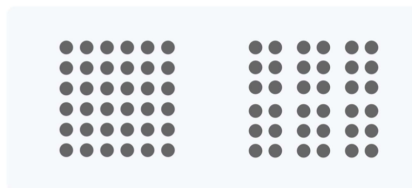
信息处理模型

人类处理机模型

- 包含三个交互式组件：
 - **感知处理器**：信息将被输出到声音存储和视觉存储区域
 - **认知处理器**：输入将被输出到工作记忆
 - **动作处理器**：执行动作
- 存在的问题：
 - 把认知过程描述为一系列处理步骤
 - 仅关注单个人和单个任务的执行过程，忽视了复杂操作执行中人与人之间及任务与任务之间的互动，忽视了环境和其他人可能带来的影响

格式塔心理学

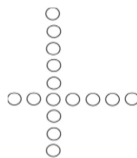
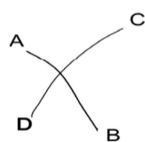
- 相近性原则：空间上比较靠近的物体容易被视为整体



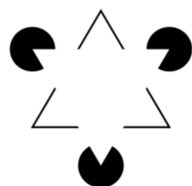
- 相似性原则：人们习惯将看上去相似的物品看成一个整体



- 连续性原则：共线或具有相同方向的物体会被组合在一起



- 完整和闭合性原则：人们倾向于忽视轮廓的间隙而将其视作一个完整的整体



人脑记忆特性

- 感觉记忆：
 - 又称瞬时记忆
 - 在人脑中持续约 1s 钟
 - 帮助我们z把相继出现的一组图片组合成一个连续的图像序列，产生动态的影像信息
- 短时记忆：
 - 感觉记忆经编码后形成
 - 又称工作记忆，约保持 30s
 - 储存的是当前正在使用的信息，是信息系统加工的核心，可理解为计算机的内存
 - 短时记忆的存储能力约为 7 ± 2 个单元
- 长时记忆：
 - 短时记忆中的信息经进一步加工后会变为长时记忆
 - 只有与长时记忆区的信息具有某种联系的新信息才能够进入长时记忆
 - 长时记忆的信息容量几乎是无限的，但会遗忘，是指长时记忆中的信息有时是无法提取，不代表长时记忆区的信息丢失了

交互设计原则与目标

可用性目标

- 易学性
- 易记性
- 高效率
- 安全性
- 效用性: utility、一定程度上该产品提供了正确的功能, 可以让用户做他们需要做的或想做的事情

用户体验目标（这部分是今年新增的，考到可能性很大）

- （感觉就是一些主观的感受，是一些形容词）

用户体验目标



■通常交互式系统包含一个多样性的用户体验目标，涵盖了一系列情绪和感受体验

■观察右表你能发现什么现象？

表 1-1 用户体验好的和不好的方面

好的方面		
满意的	有帮助的	有趣的
愉快的	激励人的	激发性的
迷人的	具有挑战性的	令人惊讶的
令人愉快的	提高社交能力的	有价值的
令人激动的	支持创新	情感上的满足
令人兴奋的	刺激认知	
不好的方面		
幼稚的	使人感到愚蠢的	
不愉快的	矫揉造作的	
使人愧疚的	傲慢的	啰嗦的
恼人的		

在特定的时间和地点使用或 与一个产品交互时，选择传达用户的感受、目前状态、情绪、感觉等最佳词汇的过程，可以帮助设计者了解用户体验的多面性变化性的本质

体验和可用性的关系

- 体验：主观；可用性：客观
- 矛盾性
 - 用塑料锤砸屏幕上的老鼠较用鼠标点击更费劲并且容易出错，但会带来一个更愉快和有趣的体验
- 有些可用性和用户体验目标是不兼容的
 - 如设计一个既安全又有趣的过程控制系统可能是不可能或不可取的
- 认识和理解可用性和其他用户体验目标之间的关系是交互设计的核心

简易可用性工程

四种有效技术

- 用户和任务观察
- 场景
- 边做边说（重要）

- 让真实用户在使用系统执行一组特定任务的时候，讲出他们的所思所想
- **最有价值的单个可用性工程方法**
- 可了解用户为什么这样做，并确定其可能对系统产生的误解
- 实验人员需要不断地提示用户，或请他们事先观摩
- 启发式评估（重要）
 - 好像是涉及到一些定量的问题？
 - 能够发现许多可用性问题
 - 为避免个人的偏见，应当让多个不同的人来进行经验性评估

黄金规则

1. 尽可能保持一致
2. 符合普遍可用性
3. 提供信息丰富的反馈
4. 设计说明对话框以生成结束信息
5. 预防并处理错误
6. 让操作容易撤销
7. 支持内部控制点
8. 减轻短时记忆负担

十条启发式规则（括号内为和黄金规则的对应关系）

1. 系统状态的可见度（3）
2. 系统和现实世界的吻合（2）
3. **用户享有控制权和自主权（6、7）**
4. 一致性和标准化（1、8）
5. 避免出错（5）
6. 依赖识别而非记忆（8）
7. **使用的灵活性和高效性（6）**
8. **审美感和最小化设计（8）**
9. 帮助用户识别、诊断和恢复错误（5、6）
0. 帮助和文档（8）

交互式系统的需求

用户建模（很重要）

人物角色

- 不是真实的人
- 是基于观察到的那些真实人的行为和动机，并且在整个设计过程中代表真实的人
- 是在人口统计学调查收集到的实际用户的行为数据的基础上形成的综合原型，拼凑出几个用户档案是不行的
- 作用：解决产品开发过程中出现的 3 个设计问题
 - 弹性用户
 - 自参考设计
 - 边缘情况设计

人物角色 + 情景 → 需求（需求获取：观察 + 场景）

需求验证：原型

- 低保真原型（多数项目的首选）
 - 与最终产品不太相似的原型
 - 优点是简单、快速、便宜、易于制作和修改
 - 三种方式
 - 草图：每张卡片代表一个屏幕或屏幕的一部分，经常用于网站开发
 - 故事板：一系列草图，展示了用户如何使用该设备完成任务，在设计早期使用，通常与场景一起使用，具有更多细节，并有机会进行角色扮演
 - 绿野仙踪法：用户认为他们是在与计算系统交互，实际上是开发人员同他们进行响应
- 高保真原型
 - 与最终产品更为接近，使用相同的材料
 - 制作时间长，难以修改
 - 风险：用户会认为原型就是系统，开发人员可能认为已找到了一个用户满意的设计

层次化任务分析 HTA

- 把任务分解为若干子任务，再把子任务进一步分解为更细致的子任务。之后，把他们组织成一个“执行次序”，说明实际情形下如何执行各项任务。

评估

评估是设计过程的组成部分。

什么是评估？

- 系统化的数据搜集过程
- 评估侧重系统的可用性和用户体验

评估的原则（小题）

- 评估应该依赖产品的用户，与专业技术人员的水平和技术无关
- 评估与设计应结合进行，仅靠用户最后对产品的一两次评估，是不能全面反映出软件可用性的（评估不是设计过程中一个独立的阶段）
- 评估应在用户的实际工作任务和操作环境下进行，根据用户完成任务的结果，进行客观的分析和评估
- 要选择有广泛代表性的用户，参加测试的人必须具有代表性

评估范型

- 快速评估
 - 可以在任何阶段进行
 - 强调“快速了解”，而非仔细记录研究发现，得到的数据通常是非正式的、叙述性的
 - 是设计网站时常用的方法
 - 基本特性：快速
- 可用性测试（用户测试，DECIDE 评估框架）
 - 评估典型用户执行典型任务时的情况，包括出错次数、完成任务的时间等
 - 基本特征：是在评估人员的密切控制下实行的
 - 主要任务：量化表示用户的执行情况
 - 缺点：
 - 测试用户的数量通常较少
 - 不适合进行细致的数据分析
- 实地研究（小孩子）
 - 基本特征：在自然工作环境中进行
 - 目的：理解用户的实际工作情形以及技术对他们的影响
 - 重难点：如何不对受试者造成影响；控制权在用户，很难预测即将发生和出现的情况
- 预测性评估
 - 研究人员通过想象或对界面的使用过程进行建模，逐步通过场景或基于问题回答的走查法
 - 基本特征：用户可以不在场，整个过程快速、成本较低
 - 启发式评估属于预测性评估

评估泛型（方法）、评估技术

通常在题目里面会问方法（泛型）是什么，技术是什么。

技术有：

- 观察用户
 - 有助于确定新产品的需求
 - 可用于评估原型
 - 挑战：如何在不干扰用户的前提下观察用户，如何分析大量数据
- 询问用户意见

- 问卷调查?
- 询问专家意见
- 测试用户的执行情况
 - 可比较不同设计方案优劣
 - 通常在受控环境中进行
- 基于模型和理论，预测界面的有效性
 - GOMS 模型和 KLM 模型

实证研究方法（重要，今年新增）

研究假设

- 假设是一种可以通过实证研究直接检验的精确问题陈述
- 一个具体的研究假设奠定了一个实验以及统计学显著性检验的基础
- 零假设 Null Hypothesis、备择假设 Alternative Hypothesis
 - 零假设通常指不同的实验条件不会产生差异；而备择假设往往是一个与零假设相反的陈述
 - 实验的目标是找到统计学证据来反驳或否定零假设，以支持备择假设
 - 可以同时研究多对零假设和备择假设
- 一个成功的实验，从一个或多个好的假设开始是至关重要的

因变量和自变量

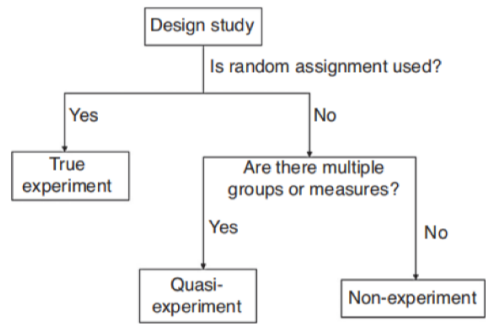
- 自变量 Independent Variables 是指研究者感兴趣的因素或因变量变化的可能原因。
 - Independent 说明该变量与受试者的行为无关，即参与者无法对自变量施加任何影响
 - 一个自变量至少需要 2 个不同的取值，这些取值被称为实验条件 test conditions
 - 人的特性是天然的自变量，如年龄、性别、身高等，但他们不能被“操纵”
- 因变量 Dependent Variables 是指研究者感兴趣的结果或效果
 - 术语“因”说明该变量依赖于受试者的行为或自变量的变化
 - 如：任务完成时间、速度、准确性、错误率、任务重试的次数、按退格键次数等
 - 因变量必须被明确定义

实验构成

- 实验条件：指的是我们需要比较的不同技术、设备或程序
- 实验单位：指的是我们应用实验条件的对象，在人机交互研究领域，实验单位通常是具有特定特征的人类受试者，特定特征例如性别、年龄、计算机使用经验等
- 分配方式：指的是将实验单位分配到不同实验条件的方式（随机化）

实验设计

- 什么是实验、准实验、非实验



组内设计和组间设计

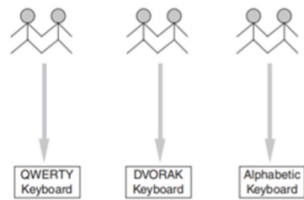


图 3.3 组间设计

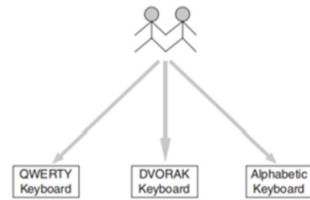


图 3.4 组内设计

- Within-group design 组内设计
 - 优点
 - 样本量小
 - 隔离了个体影响
 - 缺点
 - 学习效果的影响
 - 疲劳问题
- Between-group 组间设计
 - 优缺点与组内设计完全相反
- 选择合适的设计方法：
 - 任务简单，个体差异有限 → 组间设计
 - 当任务简单且认知过程有限时，个体差异较小，如在屏幕上选择一个目标的任务，应该采用组间设计
 - 当任务复杂或涉及重大认知功能时，个体差异更大，如阅读、理解、信息检索和解决问题相关的任务，应该采用组内设计
 - 受学习效果影响较大的任务 → 组内设计
 - 如：在一个比较网站中两种类型菜单导航效果的实验，在一个条件下完成导航任务的参与者会获得大量关于网站架构的知识，必须采用组内设计
 - 目标参与者群体很小的任务 → 组内设计
 - 很难找到和招募合格的参与者是许多 HCI 研究人员经常面临的问题
 - 当采用 **组内设计** 时，应该：
 - 控制学习效果

- 将实验条件的顺序随机化
- 当研究的目标不是与应用的初始交互时，减少学习效果的影响的有效方法是提供充足的培训
-

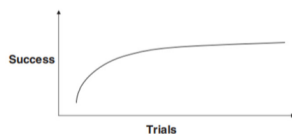


图 3.5 典型学习曲线

- 解决疲劳问题
 - 单个实验的适当长度应该是 60 - 90 分钟或更短，并提供必要的休息机会
- 学习和疲劳是不可避免的：使用拉丁方设计 Latin Square 进行平衡
-

► 假设有4种实验条件“ABCD”和4个测试用户“甲乙丙丁”，如何应用拉丁方

4 x 4 Latin Square

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

4 x 4 Balanced Latin Square

A	B	C	D
B	D	A	C
D	C	B	A
C	A	D	B

用户测试

Decide 评估框架

考试的时候要写具体的内容，比如“决定评估的总体目标是”，不能单纯背步骤。

六个步骤：

1. 决定评估需要完成的总体目标 Determine
2. 发掘需要回答的具体问题 Explore
3. 选择评估范型和技术 Choose
4. 标识必须解决的实际问题，如测试用户的选择 Identify
5. 决定如何处理有关道德的问题 Decide
 1. 应保护个人隐私，可签署一份协议书
 2. 指导原则：
 1. 说明研究的目的及要求参与者做的工作
 2. 说明保密事项，对用户 & 对项目
 3. 测试对象是软件，而非个人
 4. 对测试过程的特殊要求，是否边做边说等
 5. 说明是否对过程进行录像
 6. 用户有随时终止测试的权利
6. 评估解释并表示数据 Evaluate

- 参与者数量：至少 4 ~ 5 位，5 ~ 12 位用户也就足够了

小规模实验

很重要，在实验设计时写上“进行多次小规模实验”一定有分。

询问用户和专家

启发式评估（重点）

- 是一种灵活而又相当廉价的评估方式
- Nielsen 的十条启发式规则
- 由可用性专家完成
- 步骤
 - 彻底检查界面
 - 将界面与启发式规则进行对比
 - 列举可用性問題
 - 应用启发式规则对每一个问题进行解释与确认
- 优点
 - 不涉及用户，所以面临的实际限制和道德问题较少
 - 成本相对较低，不需要特殊设备，而且较为快捷，又被称为“经济评估法”
- 缺点
 - 评估人员需要长时间的训练才能成为专家，理想的专家应同时具备交互设计和产品应用域的知识
 - 可能出现“虚假警报”

用户测试和启发式评估

- 不同方法发现的问题不同
- 启发式评估更廉价，但用户测试才是可用性的 Gold Standard

交互系统中的设计

简化交互设计策略

- 删除、组织、隐藏、转移

交互设计模型与理论

预测模型

- 能够预测用户的执行情况，但不需要对用户做实际测试
- 特别适合于无法进行用户测试的情形

GOMS

- 最著名的预测模型
- 基于人类处理机模型
- 是关于人类如何执行认知—动作型任务以及如何与系统交互的理论模型
- 全称：
 - Goal 目标
 - Operator 操作
 - Method 方法
 - Selection Rule 选择规则：GOMS 认为方法的选择不是随机的
- 优点：
 - 能够容易地对不同的界面或系统进行比较分析
- 局限性：
 - 假设用户完全按一种正确的方式进行人机交互，没有清楚地描述错误处理的过程
 - 只针对那些不犯任何错误的专家用户
 - 任务之间的关系描述过于简单
 - 忽略了用户间的个体差异

KLM 击键层次模型

- 对用户执行情况进行量化预测，仅涉及任务性能的一个方面：时间
- 用途：
 - 预测无错误情况下专家用户在下列输入前提下完成任务的时间
 - 便于比较不同系统
 - 确定何种方案能最有效地支持特定任务
- 没有考虑：
 - 错误、学习性、功能性、回忆、专注程度、疲劳、可接受性.....

Fitts 定律

- 能够预测使用某种定位设备指向某个目标的时间
- 人机交互中，根据**目标大小**及**至目标的距离**，计算指向该目标的时间，可指导设计人员设计按钮的位置、大小和密集程度
- 对图形用户界面设计有明显的意义
- Fitts 定律建议：

- 大目标、小距离具有优势
- 大目标
 - 屏幕元素应该尽可能多的占据屏幕空间
 - 屏幕元素应尽可能利用屏幕边缘的优势
- 小距离
 - 最好的像素是光标所处的像素
- 大菜单，如饼型菜单，比其他类型的菜单使用简单