

认知心理学基础

第九讲 概念

主讲人：刘剑毅
人工智能学院

概念的定义

■ 概念 (Concept)

- 概念是事物本质的反映，是对一类事物进行概括的表征

■ 认知心理学中概念研究的范畴

- **概念形成**：也成为**概念学习**，是指个人掌握概念的过程，获得事物的概括表征
- **概念结构**：概念的表征由哪些因素所构成，研究表征的内部组织

■ 概念研究的早期阶段

- **刺激-反应原则**：共同因素说 (Hull, 1920) ，共同中介说 (Osgood, 1953) ，强调概念形成的被动性，忽略了人的主动性

概念形成（学习）

■ 假设-考验说

- 人在概念形成过程中，需要综合利用即时获得的、和已存贮的信息来主动提出一些可能的假设组成一个**假设库**，通过某种**策略**，从记忆库中选取某种假设，并基于外界的**反馈**进行**考验**，直到形成某个正确的假设，形成概念。

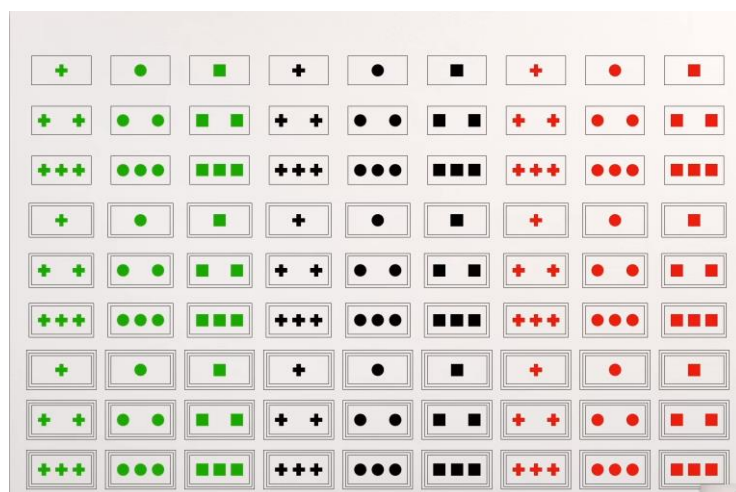
人工概念形成实验 (Bruner, 1956)

■ 影响因素 (IV)

- 形状
- 颜色
- 外框
- 个数

■ 实验范式

- 规定几个维量的属性（如红色圆形）为人工概念，但并不告诉被试
- 向被试呈现一个正实例
- 要求被试从右图自主选择正实例，并给予对或错的二值反馈
- 直到被试发现概念



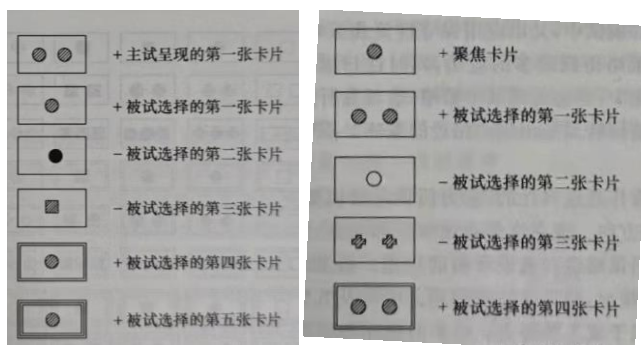
人工概念形成实验

■ 实验结果

- 被试的选择不是杂乱无章的，而是有着某种顺序；**部分假设**基于正实例的一个属性建立未知概念的假设，而**总体假设**聚焦于正实例的所有属性

■ 策略选取

- **同时性扫描**：同时进行多个概念的假设和考验——记忆负担重
- **继时性扫描**：一次只考验一个假设——记忆负担轻但易重蹈覆辙
- **保守性聚焦**：排除法，每次更换一个焦点——优于上述两种策略
- **博弈性聚焦**：一次更换两个以上焦点——有可能快速发现概念



保守性聚焦的两个范例

强化学习 (Reinforcement Learning)

● 一个例子：清扫机器人

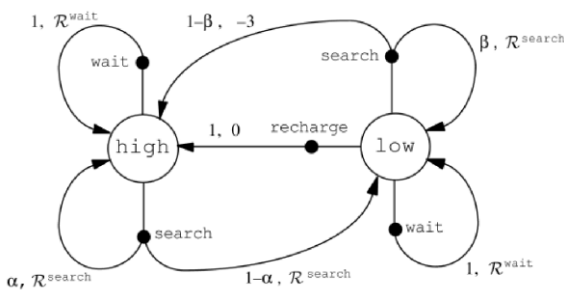
任务描述：尽力搜索收集场区内的垃圾，并择机自动充电，以避免电池耗尽。

强化学习的四个要素：

- **状态**： $\mathcal{S} = \{\text{high}, \text{low}\}$
- **动作**： $\mathcal{A}(\text{high}) = \{\text{search}, \text{wait}\}$
 $\mathcal{A}(\text{low}) = \{\text{search}, \text{wait}, \text{recharge}\}$
- **迁移概率**： $\mathcal{P}_{ss'}^a = \Pr\{s_{t+1}=s' \mid s_t=s, a_t=a\}$
- **奖励**： $\mathcal{R}_{ss'}^a = E\{r_{t+1} \mid s_t=s, a_t=a, s_{t+1}=s'\}$

Sutton, Barto, 《强化学习导论》，2005

● 状态迁移图



$s = s_t$	$s' = s_{t+1}$	$a = a_t$	$\mathcal{P}_{ss'}^a$	$\mathcal{R}_{ss'}^a$
high	high	search	α	$\mathcal{R}^{search}$
high	low	search	$1 - \alpha$	$\mathcal{R}^{search}$
low	high	search	$1 - \beta$	-3
low	low	search	β	$\mathcal{R}^{search}$
high	high	wait	1	$\mathcal{R}^{wait}$
high	low	wait	0	$\mathcal{R}^{wait}$
low	high	wait	0	$\mathcal{R}^{wait}$
low	low	wait	1	$\mathcal{R}^{wait}$
low	high	recharge	1	0
low	low	recharge	0	0

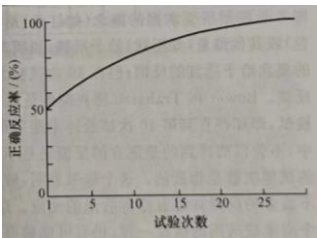
规则设计的目的：

- $\mathcal{R}^{search} > \mathcal{R}^{wait} > 0$ ：鼓励积极搜索而非消极等待
- 惩罚失败：电池耗尽。

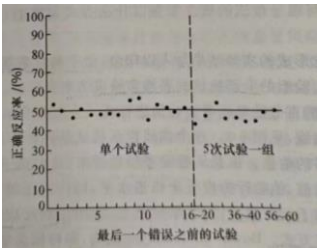


概念形成研究中的实验范式设计

- 问题：概念形成的过程是渐进式还是阶跃式？
- 渐进式概念形成的实验证据：
- 对渐进实验结果的反驳：
单个被试的阶跃式概念掌握实质被大量被试求平均的表象所掩盖
- Bower, Trabasso 1964的后向学习（Backwards Learning）经典实验
- 实验范式设计：以被试最后一次错误反应来回溯编排实验序号，统计正确反应率
- 实验结果：
 - ✓ 正确率始终波动于50%，即机遇水平
 - ✓ 被否定的假设仍会被再次应用——替代采样，无记忆现象
- 实验结论：概念形成是以全或无（阶跃）的方式进行的



渐进式猜想的证据



逆向学习曲线实验



假设考验说的发展

Levine 1966, 空白试验法

- 实验目的：直接揭示被试每次是否应用假设，以及应用什么假设
- 实验范式设计

- 4个维度：字母、左右、黑白、大小
- 一次给被试呈现一正一负一对刺激，无反馈（因此称空白试验），要求被试判断正负，并在若干轮后发现概念

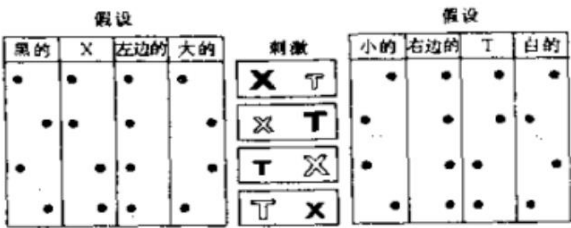


图 8-6 Levine 的空白试验的刺激及对应 8 种假设的反应模式

空白试验法

(Levine 1966)

实验范式设计

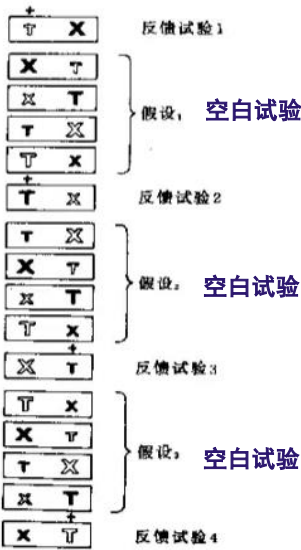
- 空白试验与反馈试验交替进行

范式原理

- 因空白试验中无反馈，被试连续的选择可以隐式（implicitly）反映出被试的概念假设
- 通过反馈试验，使被试逐步掌握概念的信息，直至发现概念

实验结果

- 92.4%的被试在空白试验中均呈现出8种假设模式之一，验证了假设行为的存在，并揭示了其具体假设
- 一个被反馈试验否定的假设，再次被作为假设的概率仅2%，验证了记忆的作用，反驳了“替代采样”策略（应为12.5%）



8-7 空白试验的 16 次试验程序

概念结构

- 研究目标
 - 揭示概念表征由哪些因素构成，以及这些因素间的相互关系
- 两种彼此对立的学说
 - 特征表说 (Bourne et.al, 1979)
基本观点：从一类个体具有的**共同重要特征**来说明概念
 - 原型说 (Rosch, 1975)
基本观点：概念是以**原型 (即最佳实例)** 表征出来的，人们总是从能最好说明一个概念的实例来理解该概念的

特征表说 (Bourne et.al, 1979)

- 继承了**层次网络模型** (Collins, 1969)、**集理论模型** (Meyer, 1970)、**特征比较模型** (Rips, 1973) 等中对概念的语义特征表征方式

概念

$C=R(X, Y, \dots)$

例：鸟=合取 (羽毛, 动物)

概念规则：整合特征的规则

定义性特征：描述一类个体的共同特征

表 8-1 着眼于两个特征的刺激总体分组的概念规则

名称	符号表达	语言描述
肯定	R	所有红色的图形均为肯定实例
合取	$R \cap S$	所有红色的方块均为肯定实例
包含性析取	$R \cup S$	所有红色的图形或方块或既是红色又是方块的均为肯定实例
条件	$R \rightarrow S$	如果图形是红色的, 则它必须是方块, 才能成为肯定实例
双重条件	$R \iff S$	红色的图形能为肯定实例, 如果而且仅仅如果它是方块
否定	$\bar{R}$	所有不是红色的图形均为肯定实例

原型说

■ Rosch 1975 认为，概念主要是以原型，即它的最佳实例表征的，人们能从最好地说明一个概念的实例来理解该概念

- 最高等级的实例即为该概念的原型（最佳实例）
- 范畴成员代表程度：表明同类个体的容许变异性，即其他个体偏离原型的容许距离

表 8-2 4 个语义范畴的实例的优良程度评定的常模

家具	等级	水果	等级	机动车	等级	武器	等级
椅子	1.5	橙子	1	汽车	1	枪	1
沙发	1.5	苹果	2	旅行汽车	2	手枪	2
睡椅	3.5	香蕉	3	卡车	3	左轮手枪	3
桌子	3.5	桃子	4	小汽车	4	机关枪	4
安乐椅	5	梨	5	公共汽车	5.5	步枪	5
梳妆台	6.5	杏	6.5	出租汽车	5.5	弹簧刀	6
转椅	6.5	柑橘	6.5	吉普车	7	刀	7
咖啡桌	8	梅子	8	急救车	8	匕首	8
摇椅	9	葡萄	9	摩托车	9	猎枪	9
双人椅	10	油桃	10	街车	10	剑	10
柜橱	11	草莓	11	大篷货车	11	炸弹	11.5
写字台	12	葡萄柚	12	Honda 牌汽车	12	手榴弹	11.5
床	13	浆果	13	缆车	13	原子弹	13.5
						剃刀	13.5

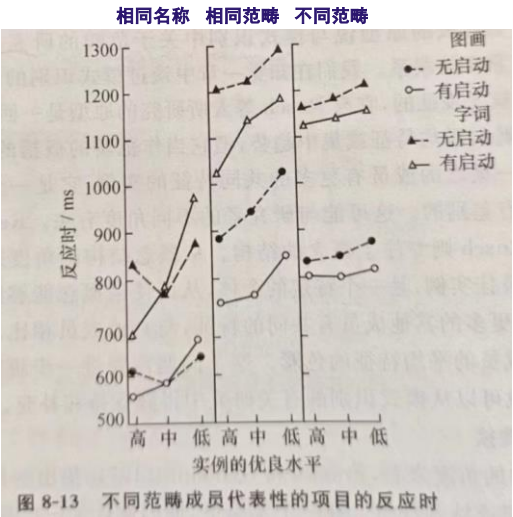
(引自 Rosch, 1975)

原型的编码

- Rosch 1975 匹配试验：
给被试同时呈现一对名词，要求被试尽快说出其范畴是否相同，记录反应时
- 影响变量：
 - 范畴成员代表程度：高（知更鸟）、中（猫头鹰）、低（企鹅）
 - 模态：图画材料、字词材料
 - 启动（Priming）有或无：呈现刺激前，将其所涉及的范畴（鸟）告诉被试
 - 范畴相同与否：名称相同（知更鸟-知更鸟）、范畴相同（知更鸟-麻雀）、范畴不同（知更鸟-苹果）

原型的编码

- 实验结果：
 - 范畴成员代表程度高时，其反应时更短，说明更接近原型的名称具有更大的易化效应
 - 图画材料快于字词材料，说明原型可能是以图画（表象）来编码的
 - 启动技术明显加速反应时
 - 不同范畴实验结果有着与同一范畴相似的趋势



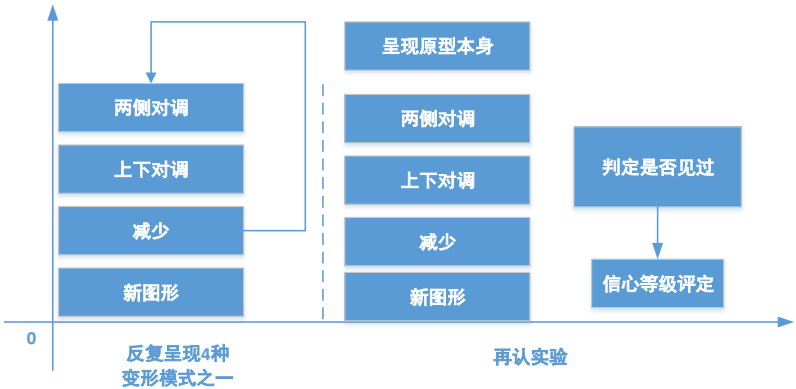
原型转换试验

(Franks 1971)

实验范式设计

基本模式	
两侧对调	
上下对调	
减少	
3种转换结合	

图 8-14 基本模式及其转换的实例



原型转换试验

(Franks 1971)

■ 实验结果

- 第一阶段并未见过的原型以最大概率被再认，并得到最高信心评定
- 距离原型越远的转换，其再认等级越低

■ 实验结论

- 被试从属于一个范畴的刺激变形中，抽象出了最佳实例（原型）

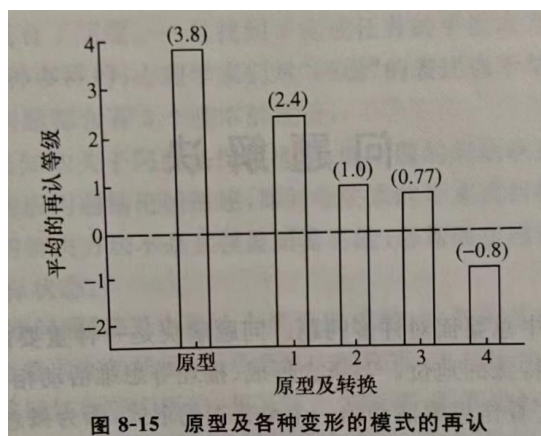


图 8-15 原型及各种变形的模式的再认