

## 第六章 机器学习

学习 → 行为的改变 → 西蒙

知识的获取 → 认知主义

构造或修改表示 → 迈克尔斯基

发展史 → 通用的学习系统 → 神经网络

基于符号表示的概念学习和研究 → 示例归纳 Winston

基于知识的各种学习系统研究 → 专家系统

联接学习和符号学习的深入研究 → 神经网络  
开发各种学习方法  
生成和选择学习任务的能力  
遗传算法

分类 → 学习策略

表示形式

应用领域

综合考虑

学习策略 → 机械学习 → 事先编好

示例学习 → 从环境获取信息

演绎学习 → 演绎推理 (归纳推理)

类比学习 → 从源域、目标域寻找相似性

基于解释的学习 → 通过概念、例子、理论、准则归纳为一个解释

归纳学习 → 正反例中归纳

简单  
↓  
复杂

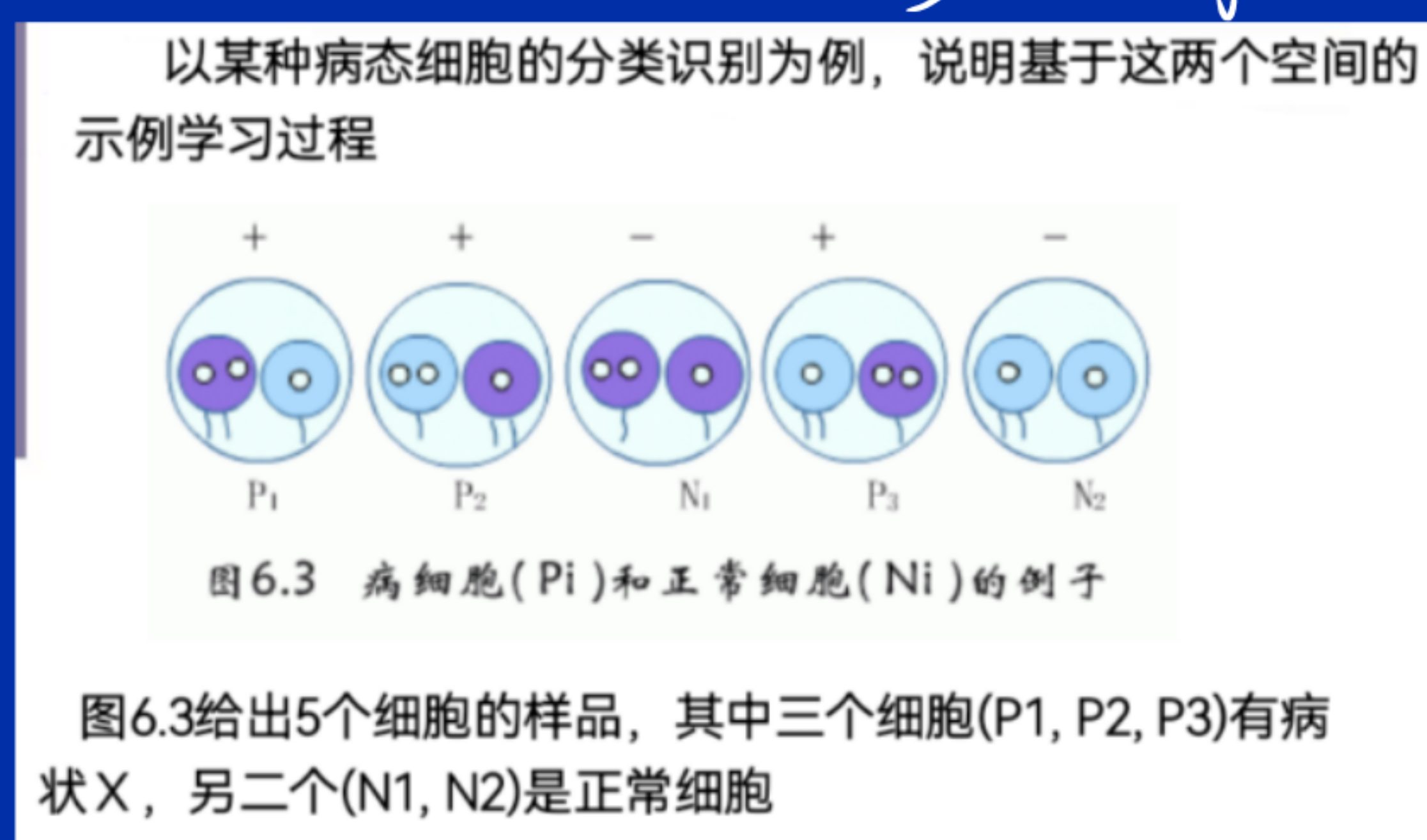
示例学习

具体事例 → 一般性理论

覆盖所有正例: 完全描述

不包括任何反例: 一致描述

包含正反例, 且可获得一致的概念描述: 解释描述



泛/特化: 反对称, 可传递的关系, 自底向上/自顶向下

假设空间为半序集 (同层假设间无关)

逐步泛化

H: 当前假设集, 初始为第一个观察的正例 P

N: 已观察到的反例集, 初始值为空集 {}



下一个例子为  $i$

若  $i$  为正例

① 对每个不覆盖  $i$  的假设  $h \in H$ ,

用能覆盖  $i$  和  $h$ , 且泛化程度最低的假设代之。

移去  $H$  中过于泛化的假设

移去  $H$  中能覆盖反例的假设

若  $i$  为反例

① 把  $i$  加入反例集合  $N$

② 从集  $H$  中移去能覆盖  $i$  的假设

逐步特化:

从最泛化的策略开始, 依靠反例生成特化假设

$H$ : 初始为最泛化的假设

$P$ : 正例集, 初始为  $\{i\}$

若下一个例子  $i$  为反例

① 对每个覆盖  $i$  的假设  $h \in H$ , 用可被  $H$  覆盖但排斥  $i$  的特化程度最低的假设取代

② 移去过于特化的假设

③ 移去不覆盖正例的假设

双向学习: 遇到正例做泛化, 遇到反例做特化, 双向合一学习成功  
决策树构造法:

归纳推理产生的结果是一棵决策树。

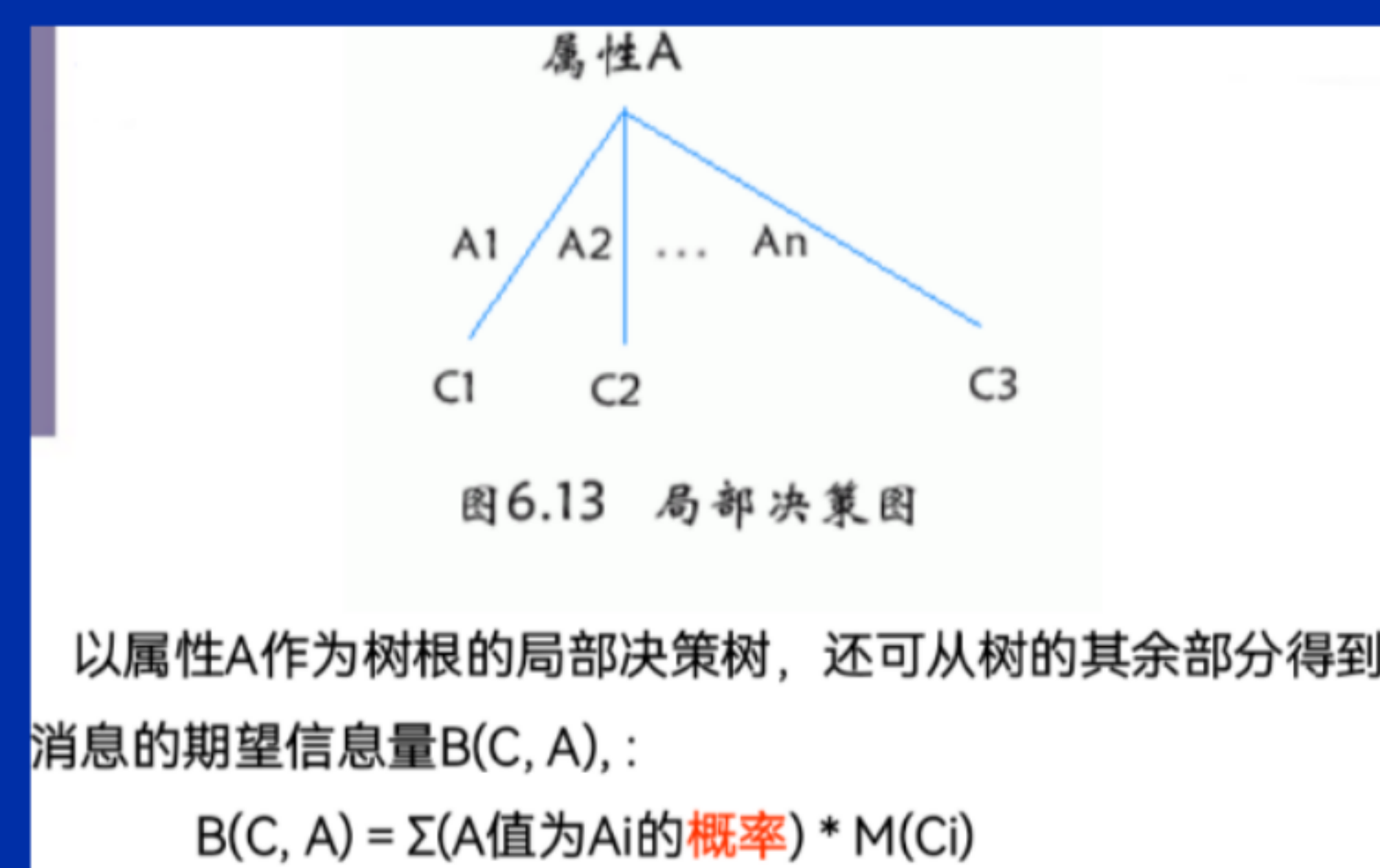
用它可区分所有给定例子的类属

香农信息论中的方法使分类时期望的测试次数最小

$$-P^+ \log_2 P^+ - P^- \log_2 P^-$$

$P^+$ ,  $P^-$ : 分类中 "+" 和 "-" 的对象所占的比例

期望信息量:  $M(C)$ ,  $M(\{ \}) = 0$



我们希望:  $M(C) - B(C, A)$  为最大值。

基于解释的学习

① 基于解释的泛化

目标概念: 对所学概念的一个初始描述。

训练例子: 目标概念的一个正例。

领域理论: 解释为何训练例子是正例的规则和事实集合。

可操作准则: 学到的知识所需遵从的表示形式。

遗传算法:

① 将群体作为处理对象。

② 用二进制串表示个体



④ 第一近邻表示个体

⑤ 计算适应值, 高的所表示的特征有更大可能出现在下一代

⑥ 遗传操作: 交换和突变

SGA (简单遗传算法)

流程: begin

1) 生成初始群体

2) 评估群体

3) while (未达到要求)  
do

begin

① 选择作为下一代群体的各个体

② 执行交换和突变操作

③ 评估群体

end

end

若适应值为  $f_i$ , 个体  $i$  在下一代的比例将为

$$\frac{f_i}{\sum f_i}$$

个体总数为  $n$  (不变), 出现个数为  $n \times \frac{f_i}{\sum f_i}$

交换:

$A_1 = 0001$     若交换点为 4     $A_1' = 0000$

$A_2 = 0010$      $A_2' = 0011$

突变: 1, 0 变换

SGA中: 群体大小为 30~200

交换率为 0.5~1

突变率为 0.001~0.05