

Solving Problems by Searching 阅读笔记

1 文献信息

洪坤焯 19211231

作者：斯图尔特·罗素（Stuart Russell）、诺文（Peter Norvig）

论文题目：Solving Problems by Searching

发表途径：《人工智能：一种现代的方法（第3版）》

（《Artificial Intelligence: A Modern Approach（Third Edition）》）

发表时间：2010. 8. 1（本文笔记依据第三版）

2 问题意义

搜索是指在确定性的、可观察的、静态的和完全可知的环境下,Agent 可以用来选择行动的方法。在这种情况下,Agent 可以构造行动序列以达到目标。在 Agent 可以开始搜索解之前,必须对目标和良定义的问题加以形式化。

很多问题的求解不能通过一步完成,而是要采用一组行动序列以达到目标。因此研究搜索以求解问题很有必要。

3 思路方法

本章的思路如图 3.1.1 所示:

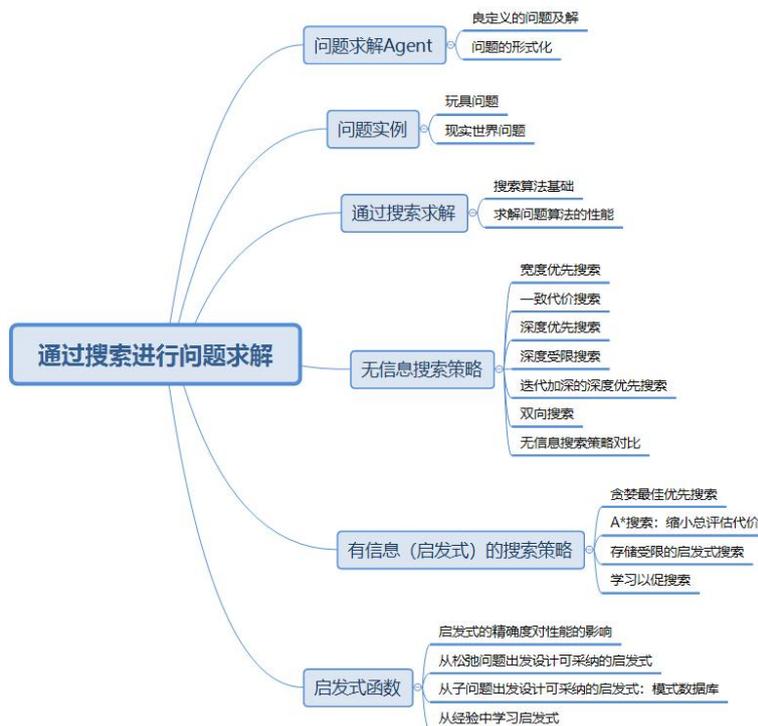


图 3.1.1 《通过搜索进行问题求解》的撰写思路

3.1 问题求解 Agent

文章对于人工智能的问题求解做了相关的介绍。包括解（达到目标的动作序列）、过程（寻找该动作序列的过程称为搜索）、问题形式化（给定一个目标，决定要考虑的动作与状态）、搜索的原因（对于某些 NP 完成或者 NP-hard 问题，只能通过搜索解决）、问题求解智能体（一种基于目标的智能体，通过搜索解决问题）。

其中，问题形式化是搜索问题的关键，有五个要素，分别为初始状态、动作、转换模型、目标测试、路径代价。初始状态是智能体出发时的转态；动作指的是描述该智能体可执行的动作；转换模型是描述每个动作做什么；目标测试是确定一个给定状态是否是目标状态；路径代价是每条路径所分配的一个数值代价。

3.2 问题实例

在举例时，文章将搜索举例分为玩具问题和现实世界问题，报告中主要总结“真空吸尘器世界”问题。吸尘器的状态是由智能体的位置和灰尘的位置决定的，如图 3.2.1 所示。链接表示动作如 L、R、S 分别为左移、右移、吸尘。

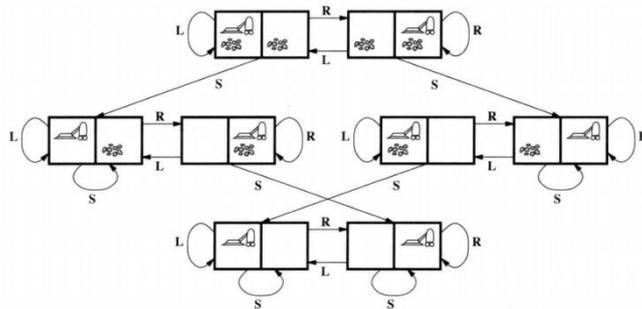


图 3.2.1 吸尘器状态图

其问题形式化的五个要素如表 3.2.1 所示

表 3.2.1 吸尘器问题形式化的五个要素

要素	说明
初始状态	任何转态都可以作为初始转态
动作	每个状态有三个动作：左移、右移以及吸尘
转换模型	除了在最左（右）端左（右）移和在清洁区吸尘外其他动作应达到预期效果
目标测试	是否所有的区域内都干净
路径代价	等于路径的步数（每一步的代价）

3.3 通过搜索求解

搜索算法的本质就是考虑各种可能的行动序列。其结构大多数为一种树搜索算法，区别主要在于搜索策略，也就是如何选择将要扩展的状态。

搜索算法需要一个数据结构来记录搜索树的构造过程。对树中的每个结点定义的数据结构包含四个元素：

n.STATE:对应状态空间中的状态;

n.PARENT:搜索树中产生该结点的结点(即父结点);

n.ACTION:父结点生成该结点时所采取的行动;

n.PATH-COST:代价，一般用 $g(n)$ 表示，指从初始状态到达该结点的路径消耗。

评价算法的性能要考虑完备性、最优性、时间复杂度、空间复杂度四个方面。完备性是指当问题有解时，这个算法是否能保证找到解。最优性是指搜索策略是否能找到最优解。时间复杂度是指找到解需要花费多长时间。空间复杂度是指在执行搜索的过程中需要多少内存。

3.4 无信息搜索

无信息搜索(盲目搜索)是指该搜索策略没有超出问题定义提供的状态之外的附加信息,可以完成的是生成后继节点,并且区分一个目标转态或者一个非目标转态。包括宽度优先、深度优先和一致代价搜索等搜索策略。对于一个无信息搜索策略,可以用 b (搜索树的最大分支因子)、 d (最浅解的深度)、 m (搜索树的最大深度) 来评价。

(1) 宽度优先搜索的搜索策略是扩展最浅的未扩展节点,使用 FIFO 队列实现,也即新的后继节点放后面。其时间复杂性和空间复杂性分别为 $b + b^2 + b^3 + \dots + b^d = O(b^d)$ 、 $O(b^d)$ 其中 b 为分支因子, d 为最浅解的深度。宽度优先搜索不能解决除小分支因子外的指数复杂性问题。

(2) 一致代价搜索的搜索策略是扩展最低代价的未扩展节点,其实现方法是通过队列按照路径代价排序,最低优先。其时间复杂性和空间复杂性均为 $O(b^{1+\lceil C^*/\epsilon \rceil})$ 其中 b 为分支因子, C^* 为最优解的代价, ϵ 为至少每个动作的代价。

(3) 深度优先搜索的搜索策略是扩展最深的未扩展节点,使用 LIFO 队列实现,其时间复杂性和空间复杂性分别为 $O(b^m)$ 、 $O(bm)$ 其中 b 为分支因子, m 为任一节点的最大深度。

(4) 深度受限搜索是对深度优先搜索的改进。如果状态空间是无限的,那么深度优先搜索就会发生失败,为了解决这一问题,可以采用一个预定的深度限制,也即深度 L 以外的节点被视作没有后继节点。但是深度限制的缺点也很明显:如果最浅的目标在深度限制之外,也即 $L < d$,那么该方法会出现额外的不完备性;若 $L > d$,则深度受限的方法也将是非最优的。

(5) 迭代加深搜索结合了深度优先和宽度优先的优势。逐步增加深度限制反复运行直到找到目标。它以深度优先搜索相同的顺序访问搜索树的节点,但先访问的累计顺序是宽度优先。

(6) 双向搜索同时进行两个搜索,一个是从初始状态向前搜索,而另一个则从目标向后搜索,当两者在中间相遇时停止。该方法可以通过一种剩余距离的启发式估计来导向。

3.5 有信息(启发式)搜索

有信息搜索(启发式搜索)采用超出问题本身定义的、问题特有的知识,因此能够找到比无信息搜索更有效的解。一般方法使用如下函数中的一个或两者:评价函数,记作 $f(n)$,用于选择一个节点进行扩展;启发式函数,记作 $h(n)$,作为 f 的一个组成部分。

最佳优先搜索采用 $f(n)$ 搜索策略:一个节点被选择进行扩展是基于一个评价函数 $f(n)$ 。大多数的最佳优先算法还包含一个启发式函数, $h(n)$ 。实现方法:与一致代价搜索相同。然而,最佳优先搜索使用 $f(n)$ 代替 $g(n)$ 来整理优先队列。其特例是贪婪搜索和 A* 搜索。

(1) 贪婪搜索的搜索策略是试图扩展最接近目标的节点。其评价函数

$$f(n) = h(n)$$

仅使用启发式函数对节点进行评价。 $h(n)$ 为从 n 到最接近目标的估计代价。由于每一步它都试图得到能够最接近目标的节点,故称为“贪婪”。其时间复杂性和空间复杂性均为 $O(b^m)$

其中 b 为分支因子, m 为搜索空间的深度。

(2) A*搜索的搜索策略是避免扩展代价高的路径, 使总的估计求解代价最小化。其评价函数

$$f(n) = g(n) + h(n)$$

为到达该节点的代价, $h(n)$ 为从该节点到目标的估计代价, A*搜索是最优的。

(3) 迭代加深 A*搜索是迭代加深深度优先搜索的变种。它从 A*搜索算法借鉴了这一思想, 即使用启发式函数来评价到达目标的剩余代价。因为它是一种深度优先搜索算法, 内存使用率低于 A*算法。但是, 不同于标准的迭代加深搜索, 它集中于探索最有希望的节点, 因此不会去搜索树任何处的同样深度。

(4) 迭代加深深度优先搜索:使用搜索深度作为每次迭代的截止值。迭代加深 A*搜索:使用信息更丰富的评价函数, 即

$$f(n) = g(n) + h(n)$$

其中 $g(n)$ 为到达该节点的代价, $h(n)$ 为该节点到目标的估计代价。

3.6 启发式函数

(1) 启发式的精确度对性能的影响

启发式的设计对有效分支因子的值有影响。

有效分支因子 b^* : $N + 1 = 1 + b^* + (b^*)^2 + (b^*)^3 + \dots + (b^*)^d$, 所得 b^* 越小则算法性能越好;

一般来说启发式使用值更大的函数是更好的。

(2) 从松弛问题出发设计可采纳的启发式

松弛问题是指减少了行动限制的问题, 状态空间是原有状态空间的超图, 多了很多边。一个松弛问题最优解的代价, 是原问题的可采纳启发式。通过去掉经过形式化的原问题的一个或几个条件可以形成松弛问题。最优松弛问题是指可以比较生成的松弛问题的最优解的代价, 选择最大的最优解代价的松弛问题。除此之外还可以通过从子问题除法设计可采纳的启发式和从经验中得到启发式等

4 实验结论

(1) 求解一个问题就是一系列动作, 并且搜索是为达到目标寻找这些动作的过程。

(2) 一个问题由五个部分组成:初始状态,行动集合,转移模型描述这些行动的结果,目标测试函数和路径代价函数。问题的环境用状态空间表示。状态空间中从初始状态到达目标状态的路径是一个解。

(3) 无信息搜索方法只能访问问题的定义。基本算法包括:宽度优先搜索、一致代价搜索、深度优先搜索、双向搜索;深度受限搜索在深度优先搜索上加了深度限制。迭代加深搜索在不断增加的深度限制上调用深度受限搜索直到找到目标。

(4) 有信息搜索可能需要访问启发式函数 $h(n)$ 来估算从 n 到目标的解代价。一般的最佳优先搜索算法根据评估函数选择扩展结点。贪婪最佳优先搜索和 A*搜索是两个特例。

(5) 启发式搜索算法的性能取决于启发式函数的质量。好的启发式有时可以通过松动问题的定义来构造,将子问题的解代价记录在模式数据库中,或者通过对问题类的经验学习得到

5 习题

对以下问题给出完整的形式化。选择的形式化方法要足够精确以便于实现。

- 只用四种颜色对平面地图着色，要求每两个相邻的地区不能具有相同的颜色。
- 屋子里有只 3 英尺高的猴子，离地 8 英尺的屋顶上挂着一串香蕉。猴子想吃香蕉。屋子里有两个可叠放、可移动、可攀爬的 3 英尺高的箱子。
- 有这样一个程序，当输入一个包含很多记录的文件时会输出消息“不合法的输入记录”。每个记录的处理都是独立的。请找出报错的是哪个记录。
- 有三个水壶，容量分别为 12 加仑、8 加仑和 3 加仑，还有一个放液嘴。可以把水壶装满或者倒空，从一个壶倒进另一个壶或者倒在地上。请量出刚好 1 加仑水。

解：a 初始状态：没有着色区域。目标测试：所有区域都有色，并且相邻区域中没有相同的颜色。后续函数：指定一种颜色到一个区域。成本函数：任务数。

b 初始状态：猴子和香蕉。目标测试：猴子有香蕉。后续函数：跳上板条箱从板条箱上跳下来；把板条箱从一个地方推到另一个；从一处走到另外一处；抓住香蕉（如果站在板条箱的话）。成本函数：操作数。

c 初始状态：考虑所有输入记录。目标测试：考虑单个记录，并给出“非法输入消息。后续函数：对前半部分的记录再次运行；对后半部分记录再运行。成本函数：试验次数。

d 初始状态：壶有值 $[0, 0, 0]$ 。后续函数：给定的值 $[x, y, z]$ ，生成 $[12, y, z]$, $[x, 8, z]$, $[x, y, 3]$ (通过填充); $[0, y, z]$, $[x, 0, z]$, $[x, y, 0]$ (对于任意两个具有当前值 x 和 y 的罐子，将 y 倒入 x 中；这会将水壶 x 更改为最小的 $x+y$ 和水壶的容量，用 y 减去水壶第一个水壶所得的量)。成本函数：操作数。

6 启发思考

(1) 几种搜索算法各有利弊，结合需要求解的问题选择合适搜索算法。如宽度优先搜索算法是完备的，在单位代价的情况下是最优的，但是具有指数级别的空间复杂度，适用于分支因子小的问题求解；深度优先搜索既不是完备的也不是最优的，但它具有线性的空间复杂度。适用于分支因子相对较大的问题求解。

(2) 新算法的发展是结合已有算法优点的过程。如深度受限搜索在深度优先搜索上加了深度限制。迭代加深搜索在不断增加的深度限制上调用深度受限搜索直到找到目标。

(3) 评价算法的性能需要提出多维的指标。如文章提出的完备性、最优性、时间复杂度、空间复杂度四个方面标准定量地确定了各种搜索算法的性能，为比较算法各方面的性能明确了标准。

7 附: 课程对应课本《罗素和诺维格。人工智能: 一种现代方法》章节后习题

1 (第三章: 通过搜索解决问题 3.1) 解释为什么问题的形式化必须在目标的形式化之后答: 在制定目标时, 我们决定世界的哪些方面是我们感兴趣的, 哪些可以忽略或抽象掉。然后在制定问题时, 我们决定如何处理重要的方面(忽略其他方面)。如果我们先制定问题我们不知道该包括什么, 该省略什么。也就是说, 在目标制定、问题制定和问题解决之间会有一个迭代周期, 直到一个人找到一个足够有用和有效的解决方案。

2 (第三章: 通过搜索解决问题 3.5) 用 3.2.1 节给出的“高效的”增量形式化方法处理 n 皇后问题。解释为什么状态空间至少有 $\sqrt[n]{n!}$ 个状态, 估算穷举探索可行的 n 的最大值(提示: 考虑一个皇后在每列中能够攻击到的最大方格数目, 从而得出分支因子下界)。

解: 对于 n 个皇后, 注意到一个皇后在任意给定的列中最多占有三个方格, 也就是在放置下一个皇后时要注意避开这三个方格, 因此有 (n-3) 种选择。同理再下一个皇后在放置时有 (n-6) 种选择

$$S \geq n(n-3)(n-6) \dots \quad 2) \quad S^3 \geq n^3(n-3)^3(n-6)^3$$

进行放缩: $S^3 \geq n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-1) \cdot \dots = n!$ 也即 $S \geq \sqrt[3]{n!}$

3 (第五章: 约束满足问题 5.7) 证明下面的断言: 对于每棵博弈树, MAX 使用极小极大算法对抗次优招数的 MIN 得到的效用值不会比对抗最优招数的 MIN 得到的效用值低。你能否找出一棵博弈树, 使得 MAX 用次优策略依然要好于次优 MIN 时的策略。

证明: 考虑一个 MIN 节点的子节点是终端节点。如果 MIN 发挥次优, 则节点的值大于或等于如果 MIN 发挥最优时的值。因此, 作为 MIN 节点的父节点的 MAX 节点的值只能增加。这个论点可以通过一个简单的归纳一直延伸到根源。如果 MIN 的次优策略是可预测的, 那么可以比 Minimax 策略做的更好。例如, 如果 MIN 总是落入某种陷阱并输了, 那么设置陷阱就能保证赢, 即使实际上 MIN 有毁灭性的反应。

4 (第五章: 约束满足问题 5.18) 证明即使在有机会结点的情况下, 对叶结点值的正线性变换(例如把值 x 变换成 ax+b, 其中 a>0)不会影响在博弈树中对行棋的选择。

解: 考虑使用归纳法。用归纳的深度归纳法。不妨假设节点的子数为 x_1, x_2, \dots, x_n 。并且变换是 $ax+b$, 其中 a 是已知的, 则

$$\min(ax_1+b, ax_2+b, \dots, ax_n+b) = a \min(x_1, x_2, \dots, x_n) + b$$

$$\max(ax_1+b, ax_2+b, \dots, ax_n+b) = a \max(x_1, x_2, \dots, x_n) + b$$

故 $P_1(ax_1+b) + P_2(ax_2+b) + \dots + P_n(ax_n+b) = a \sum_{i=1}^n P_i x_i + b$

因此问题的节点为一个单层树, 其中树对 b 的值是一个线性变换

即当 $a > 0$ 时 $x > y \Rightarrow ax+b > ay+b$ 因此根处的最佳选择将与原树的树中的最佳选择相同

5 (第六章: 对抗性搜索和游戏 6.1) 图 6.1 所示的地图着色问题共有多少个解? 如果是四色有多少个解? 如果只有两色呢?

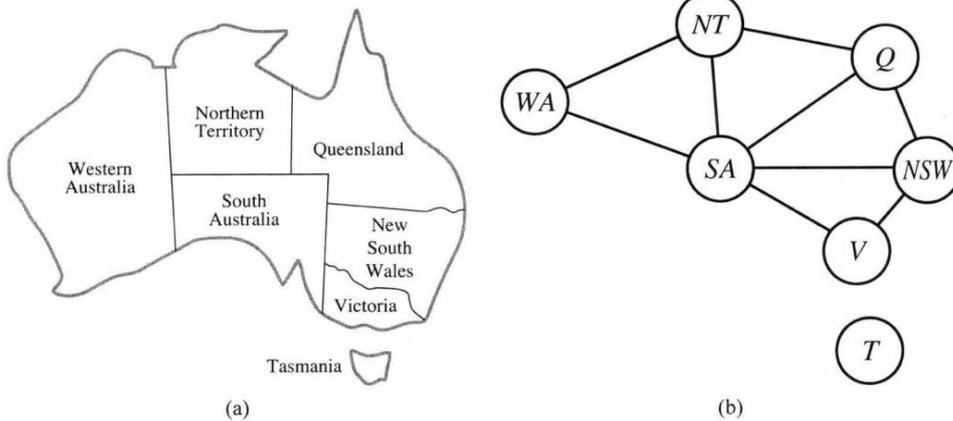


图 6.1 着色问题

(a) 澳大利亚的州和行政区。视此地图着色问题为约束满足问题 (CSP)。目标是对每个区域分配颜色, 使得相邻的区域不同色。(b) 将地图着色问题表示成约束图

用三种颜色着色澳大利亚有 18 种方案。从 SA 开始, 它可以有三种颜色中的任意一种。顺时针方向移动 WA 可以有另外两种颜色中的任意一种, 其他的都是严格决定让大陆有 6 种可能, 塔斯曼尼亚有 3 种可能, 共计 18 种可能。

四种颜色: 从 SA 开始, 它可以有四种颜色中的任意一种。顺时针方向移动 WA 可以有另外三种颜色中的任意一种, 其他的都是两种, 塔斯曼尼亚有 4 种可能, 共计 768 种可能。两种颜色无解。

6 (第六章: 对抗性搜索和游戏 6.9) 解释为什么在 CSP 搜索中, 一个好的启发式选择变量的时候应该选择约束最多的变量, 而选择值的时候应该选择受到约束最少的。

答: 最受约束的变量是有意义的, 因为它选择了一个变量 (在其他条件相同的情况下) 可能导致失败而且越早失败越有效率 (从而排除掉大部分最小约束值启发式是有意义的, 因为它为将来的分配提供了最大的机会, 以避免冲突。

7 (第十二章: 量化不确定性 13.1) 根据基本原理证明: $P(a|b \wedge a) = 1$ 。

13.1 首先定义条件概率 $P(X|Y) = \frac{P(X \cap Y)}{P(Y)} = \frac{P(XY)}{P(Y)}$

再定义逻辑连接词如下:

$$P(A|B \cap A) = \frac{P(A \cap (B \cap A))}{P(B \cap A)} = \frac{P(B \cap A)}{P(B \cap A)} = 1$$

因此: $P(A|B \cap A) = 1$ 成立

8 (第十二章: 量化不确定性 13.2) 使用概率公理证明: 一个离散随机变量的任何概率分布, 总和等于 1

13.2 已知公理 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ 对于离散型随机变量 X

不妨假设 a 为事件 $X = x_1$, b 为事件 $X = \text{其他任意值}$, 则有

$$P(X = x_1 \cup X = \text{其他任意值}) = P(X = x_1) + P(X = \text{其他任意值}) + 0$$

因为 a 和 b 是对立的, $P(A \cap B) = 0$.

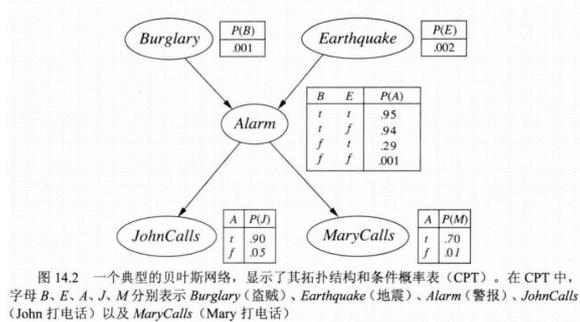
推广到更多事件的情况 $P(X = x_1 \cup \dots \cup X = x_n) = P(X = x_1) + \dots + P(X = x_n)$

左边是必然事件, 等号右边值为 1. 同理右边和也为 1.

因此, 一个离散型随机变量的任何概率分布, 总和为 1.

9 (第十三章: 概率推理 14.4) 考虑图 14.2 的贝叶斯网络。

a. 如果没有观察到证据, Burglary 和 Earthquake 是相互独立的吗? 从数值语义和从拓扑语义进行证明。b. 如果观察到 Alarm=true, Burglary 和 Earthquake 是相互独立的吗? 通过计算所涉及的概率是否满足条件独立性而证明你的结论。



14.4 (a) 是的, 可以求得 $P(B \cap E) = P(B)P(E)$, 因此没有观察到证据时 Burglary 和 Earthquake 相互独立

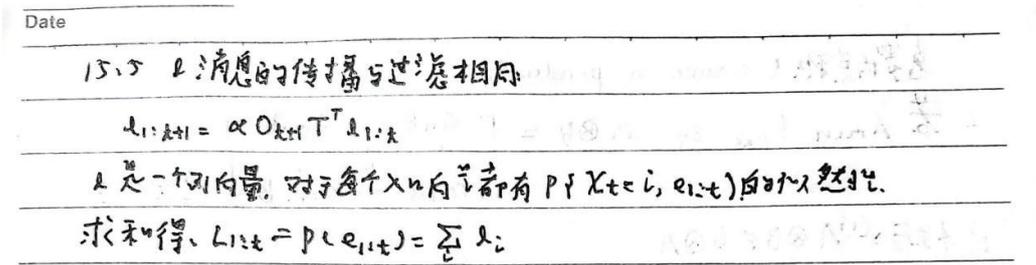
(b) 问题: $P(B, E|A) = \begin{cases} \alpha \times 0.95 \times 0.001 \times 0.002 & B=b, E=e \\ \alpha \times 0.94 \times 0.01 \times 0.998 \times 0.94 & B=b, E \neq e \\ \alpha \times 0.999 \times 0.002 \times 0.29 & B \neq b, E=e \\ \alpha \times 0.999 \times 0.1998 \times 0.001 & B \neq b, E \neq e \end{cases}$

$\alpha \times 0.008$ $B=b, E=e$
 $\alpha \times 0.13728$ $B=b, E \neq e$
 $\alpha \times 0.12703$ $B \neq b, E=e$
 $\alpha \times 0.19962$ $B \neq b, E \neq e$

α 是一个常数, 而 $P(b, e|A) \neq P(b|A)P(e|A)$

因此 B 和 E 不是条件独立的

10 (第十四章: 时间上的概率推理 15.5) 公式(15.12)描述了矩阵形式的 HMM 的滤波过程。请给出公式(15.7)所描述的似然计算的类似公式。



11 (第十五章: 做出简单决定 16.2) Chris 考虑从 4 辆二手车中买一辆具有最大期望效用的车。Pat 考虑从 10 辆车中做同样的事情。其他条件是相同的, 两个人中哪个更可能买到更好的车? 哪一个更可能会对其买到的车的质量失望? 多大程度(使用期望质量的标准偏差)?

答: Pat 比 Chris 更有可能拥有一辆更好的车, 因为他有更多的信息可供选择。然而, 如果他仅仅从表面价值上去理解最好的汽车的预期效用, 他更有可能失望。可以通过数值积分计算出期望失望值约为标准差的 1.54 倍。

12 (第十六章: 做出复杂决定 17.9) 考虑图 17.14(b) 的 101*3 的世界。在开始状态, Agent 可以从两种确定性的行动中选择, Up 或者 Down, 但在其他状态, Agent 只有一个确定性的行动 Right。假设一个折扣回报函数, Agent 在什么样的折扣 γ 下选择 Up, 又在什么样的折扣下选择 Down? 每种行动的效用计算为 γ 的函数(注意, 这个简单的例子实际上反映了许多真实世界的情形, Agent 必须评估马上要采取的行动的值, 与潜在的接下来的长期的后果进行对比, 例如选择将垃圾倒入湖中)。

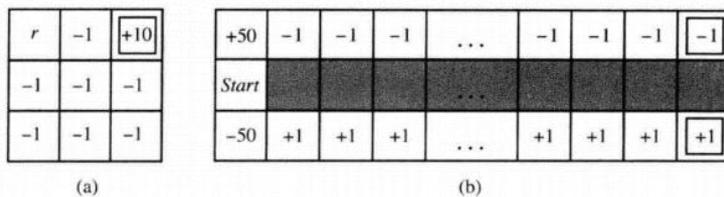


图 17.14
(a) 习题 17.8 的 3x3 世界, 其中给出了每个状态的回报。右上角的方格是终止状态。
(b) 习题 17.9 的 101x3 世界(中间省略了相同的 93 列)。开始状态的回报为 0

