Chapter9 动态规划(Dynamic Programming)

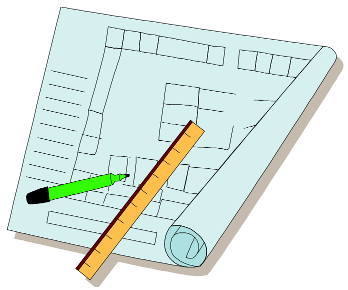
本章主要内容：

1 动态规划实例

2 动态规划的基本概念

3 动态规划的基本思想与基本原理

4 逆序解法与顺序解法



引 言

□动态规划是解决**多阶段决策过程**最优化的一种方法。

□该方法是由美国数学家贝尔曼（**R. E. Bellman**）等人在**20**世纪**50**年代初提出的。并成功地解决了生产管理、工程技术等方面的许多问题，从而建立了运筹学的一个新的分支，即动态规划。**Bellman**在**1957**年出版了《**Dynamic Programming**》一书，是动态规划领域中的第一本著作。

引 言

□动态规划与其他规划方法的不同之处在于：

动态规划是求解某类问题（多阶段决策问题）的一种方法，是考察问题的一种途径，而不是一种特定算法。

因此，它不像线性规划那样有一个标准的数学表达式和明确定义的一组(算法)规则，而必须对具体问题进行具体分析处理。因此，学习动态规划时，除对基本概念和基本方法正确理解外，还应在一定经验积累基础上，以丰富的想像力去建立模型，用创造性的技巧去求解。

动态规划实例

实例1 机器负荷分配问题（时间阶段问题）

◎设有某种机器设备，用于完成两类工作A和B。若第k年初完好机器的数量为 x_k，若以数量 u_k 用于A，余下的（x_k-u_k）用于工作B，则该年的预期收入为 $g(u_k)+h(x_k-u_k)$。这里$g(u_k)$和 $h(x_k-u_k)$是已知函数，且 $g(0)=h(0)=0$。

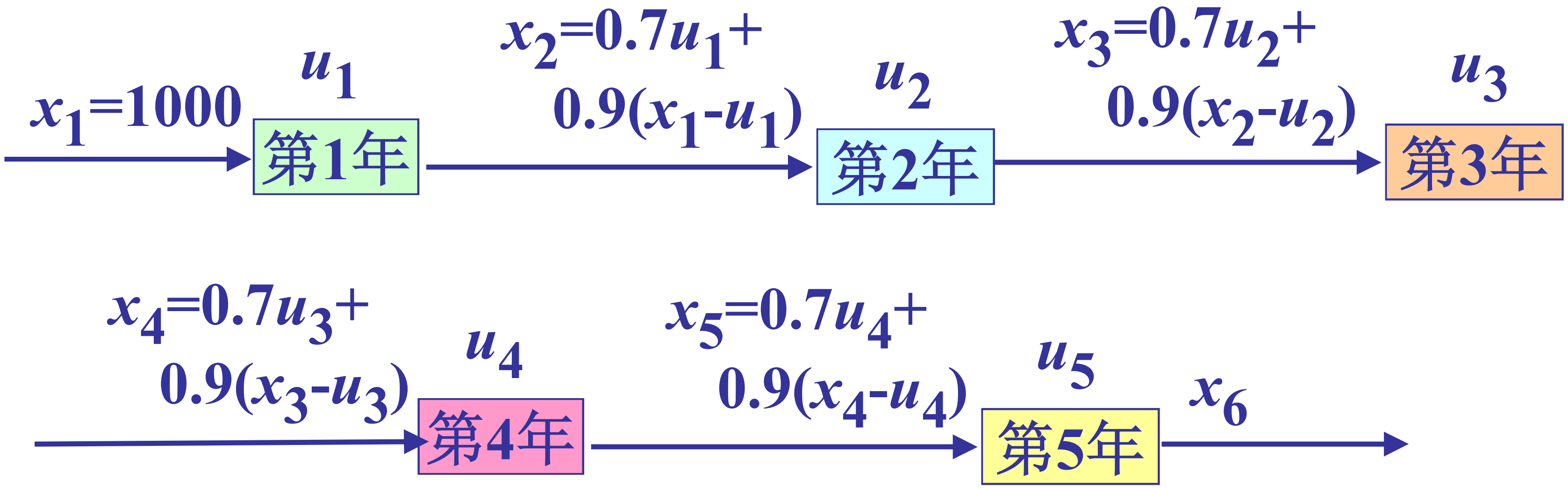
◎又机器设备在使用中会有损坏，设机器用于工作A时，一年后能继续使用的完好机器数占年初投入量的70%；若用于工作B时，一年后能继续使用的完好机器数占年初投入量的90%。则在下一年初能继续用于A、B工作的设备数为 $x_{k+1}=0.7u_k+0.9(x_k-u_k)$

◎设第1年初完好的机器总数为1000台，问在连续5年内每年应如何分配用于A、B两项工作的机器数，使5年的总收益为最大。

动态规划实例

□这是一个**多阶段决策过程**。

□该过程可以分为相互联系的若干阶段，每一阶段都需作出决策，从而形成全过程的决策。 第k阶段收益$g(u_k)+h(x_k-u_k)$

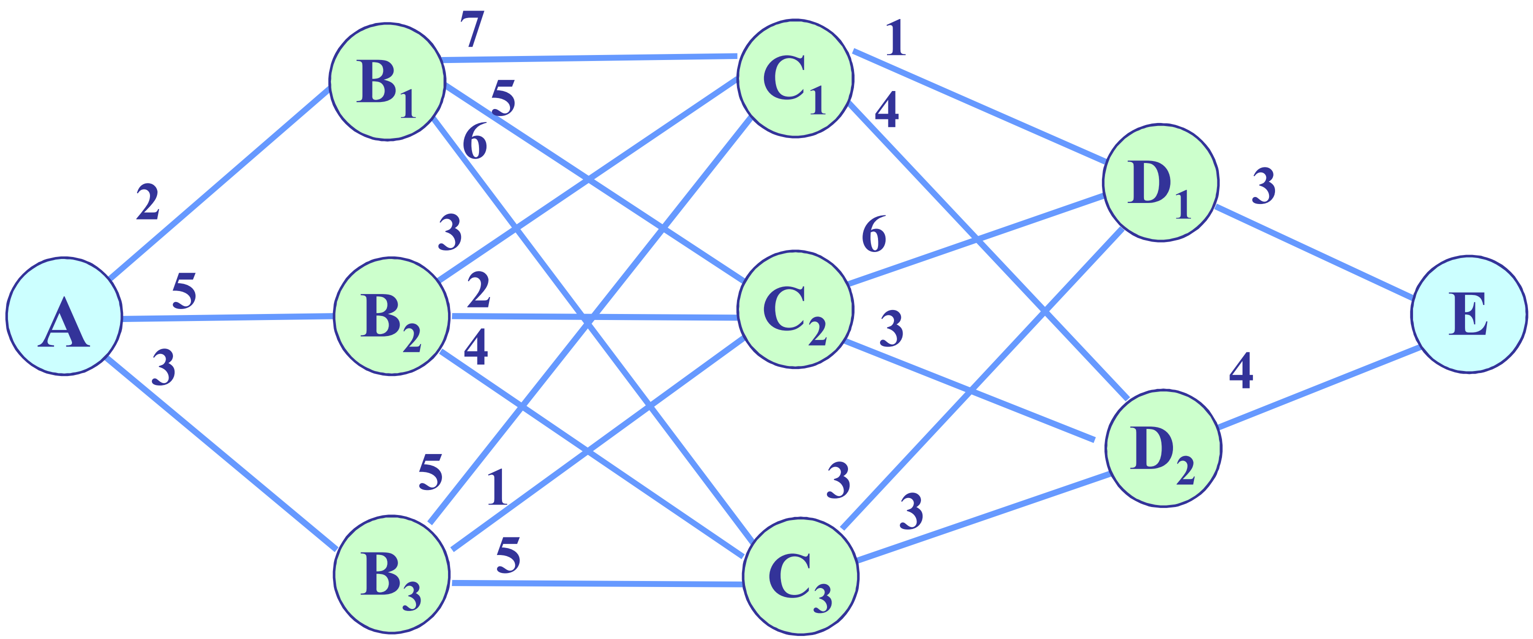


□相应的问题称为**多阶段决策问题**。

动态规划实例

实例2 最短路线问题（空间阶段的例子）

设有一个旅行者从图中A点出发，途中要经过B、C、D等处，最后到达终点E。从A到E有很多条路线可以选择，各点之间的距离如图所示，问该旅行者如何选择路线，使总的路程为最短。



- 可看成**4阶段**的决策问题。

动态规划实例

□从以上两个例子，可以知道

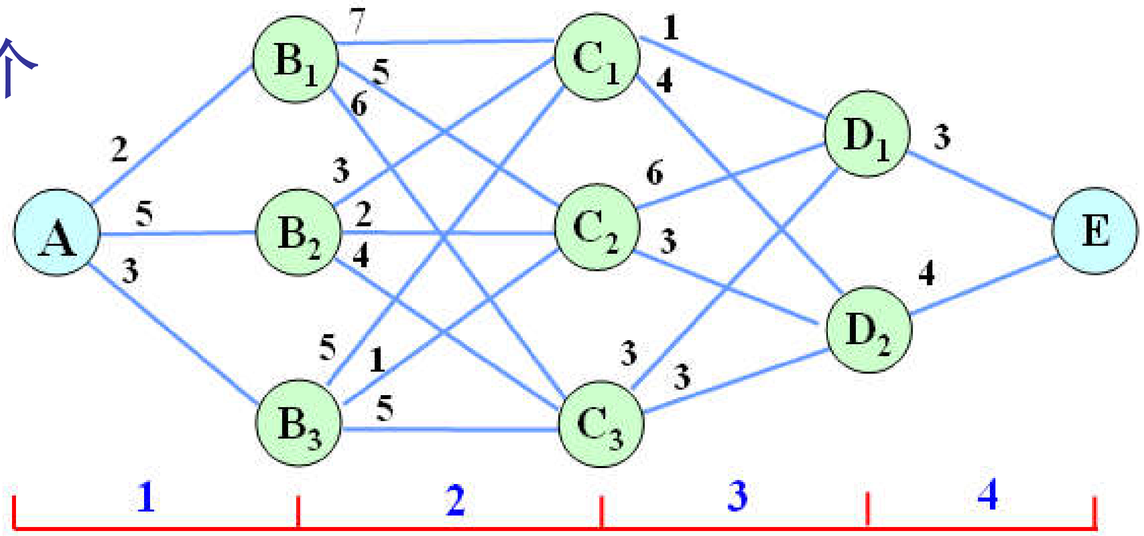
所谓**多阶段决策问题**是指这样的决策问题：其过程可分为若干个相互联系的阶段，每一阶段都对应着一组可供选择的决策，每一决策的选定既依赖于当前面临的状态，又影响以后总体的效果。

当每一阶段的决策选定以后，就构成一个决策序列，称为一个**策略**，它对应着一个确定的效果。**多阶段决策问题就是寻找使此效果最好的策略。**

动态规划的基本概念

（1）阶段（stage）

- 为了便于求解和表示决策及过程的发展顺序，而把所给问题恰当地划分为若干个相互联系又有区别的子问题，称之为多段决策问题的**阶段**。一个阶段，就是需要作出一个决策的子问题。
- 通常，阶段是按决策进行的**时间或空间**上先后顺序划分的。
- 描述阶段的变量称为**阶段变量**，常记为k，$k=1,2, \ldots,n$。
- 如本例可按空间分为4个阶段来求解，$k=1, 2, 3, 4$。

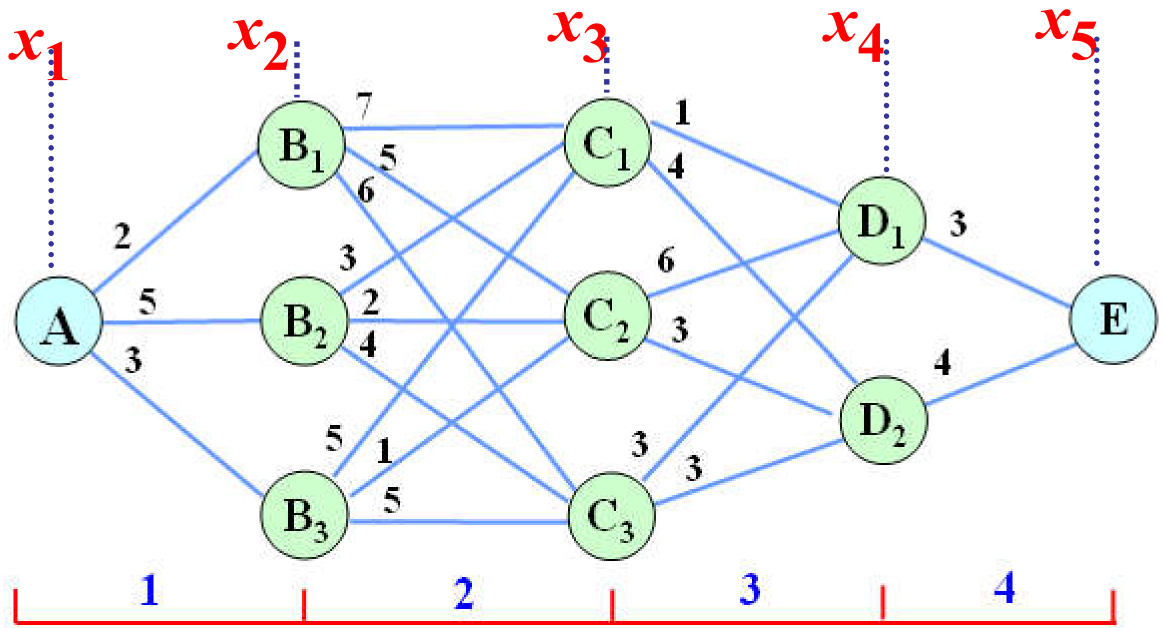


动态规划的基本概念

（2）状态（state）

□**状态**：每阶段初的客观条件。描述各阶段状态的变量称为**状态变量**，常用x_k表示第k阶段的状态。

□例1中，**状态**就是某阶段的出发位置。



□按状态变量的取值是连续还是离散，可将动态规划问题分为**离散型**和**连续型**。

动态规划的基本概念

□动态规划中的**状态应满足无后效性（马尔科夫性）**：

所谓**无后效性**指系统到达某个状态前的过程的决策将不影响到该状态以后的决策。［指系统从某个阶段往后的发展，仅由本阶段所处的状态及其往后的决策所决定，与系统以前经历的状态和决策（历史）无关。过程的过去历史只能通过当前的状态去影响它未来的发展］

□例1中，当某阶段的状态已选定某个点时，从这个点以后的路线只与该点有关，不受该点以前的路线的影响，所以满足状态的无后效性。

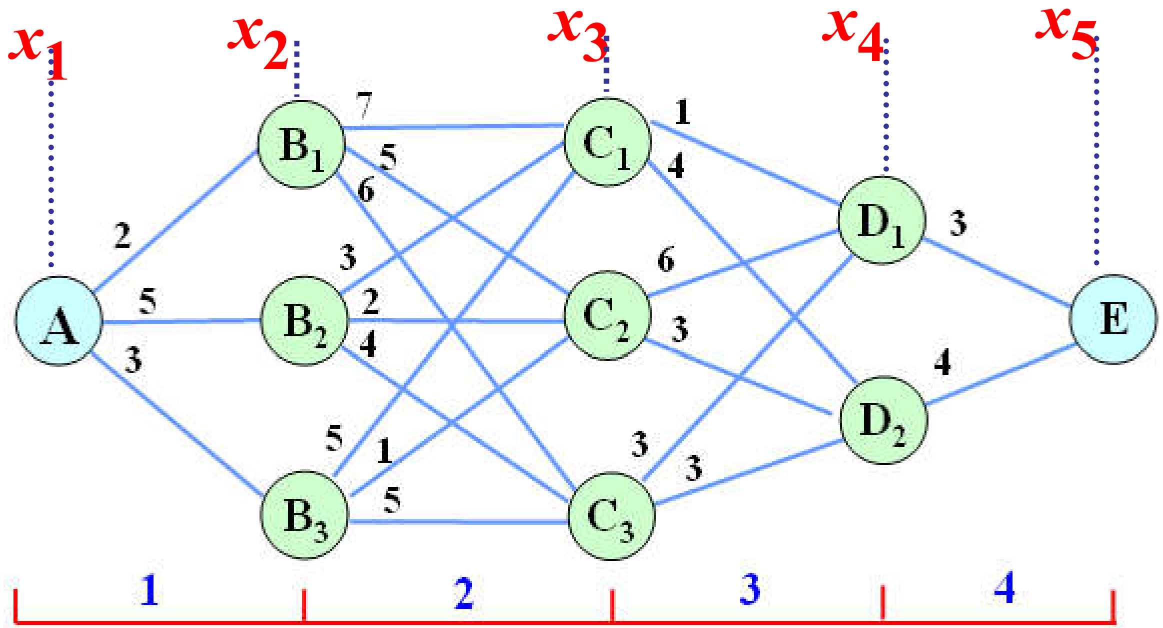
动态规划的基本概念

- **状态集合**：状态变量 x_k 的取值集合称为状态集合，状态集合实际上是关于状态的约束条件。
- 通常用 S_k 表示状态集合，$x_k \in S_k$。

- 第1阶段

  $S_1=\{A\}$；
- 第2阶段具有3个状态B1、B2和B3，故

  $S_2=\{B1, B2, B3\}$。
- ……



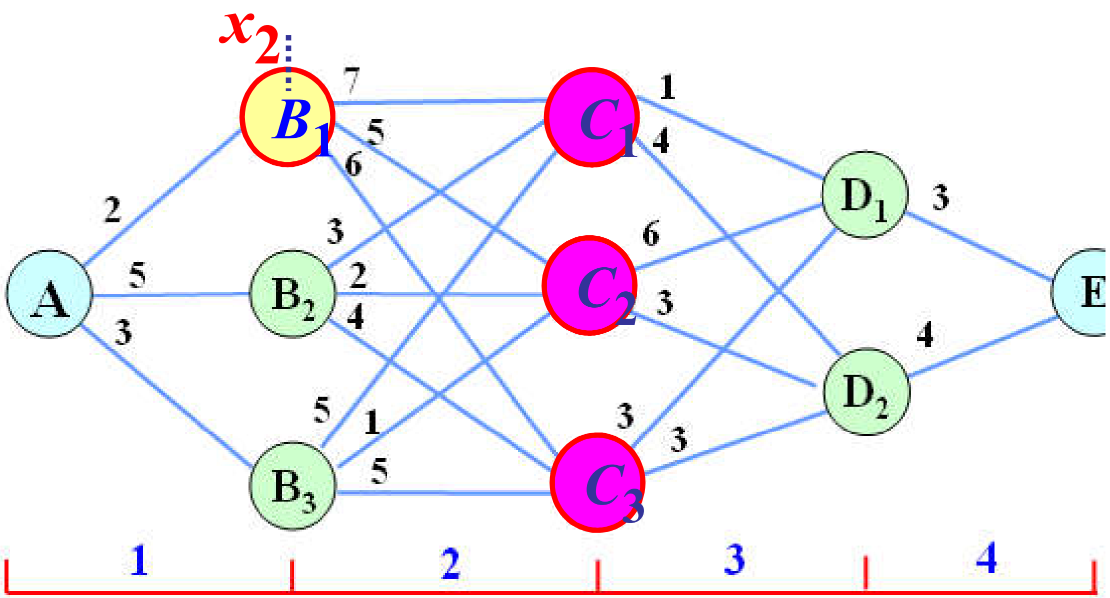
动态规划的基本概念

（3）决策（decision）

□当过程处于某一阶段的某状态时，可以做出不同的决定，从而确定下一阶段的状态，这种决策称为**决策**。

□描述决策的变量称为**决策变量**，常用$u_k(x_k)$表示第k阶段当状态处于x_k时的决策变量，它是状态变量的函数。

□例1中，从第2阶段的状态B1出发，可以选择下一阶段的C1、C2、C3。
□如我们决定选择C1，则可表示为：$u_2(B1) = C1$。

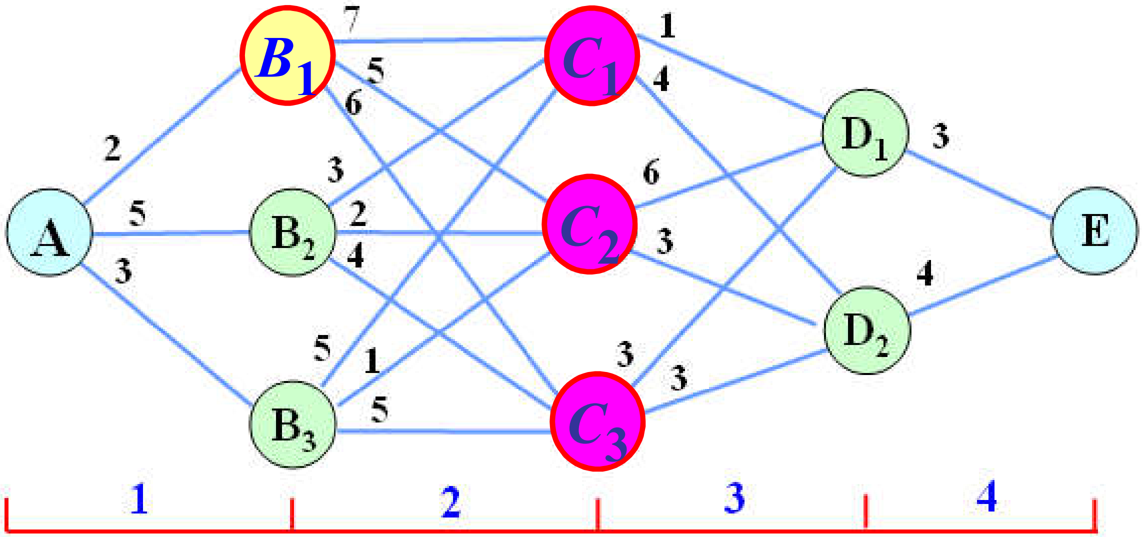


动态规划的基本概念

□**决策集合**：第k阶段当状态处于x_k时决策变量$u_k(x_k)$的取值范称为**决策集合**，常用$D_k(x_k)$表示。

□例1中，从第2阶段的状态B1出发，可以选择下一阶段的C1、C2、C3。

□即 $D_2(B_1)=\{$ C1、C2、C3 $\}$；



□决策集合实际上是决策的约束条件，$u_k(x_k)\in D_k(x_k)$。

动态规划的基本概念

□小结

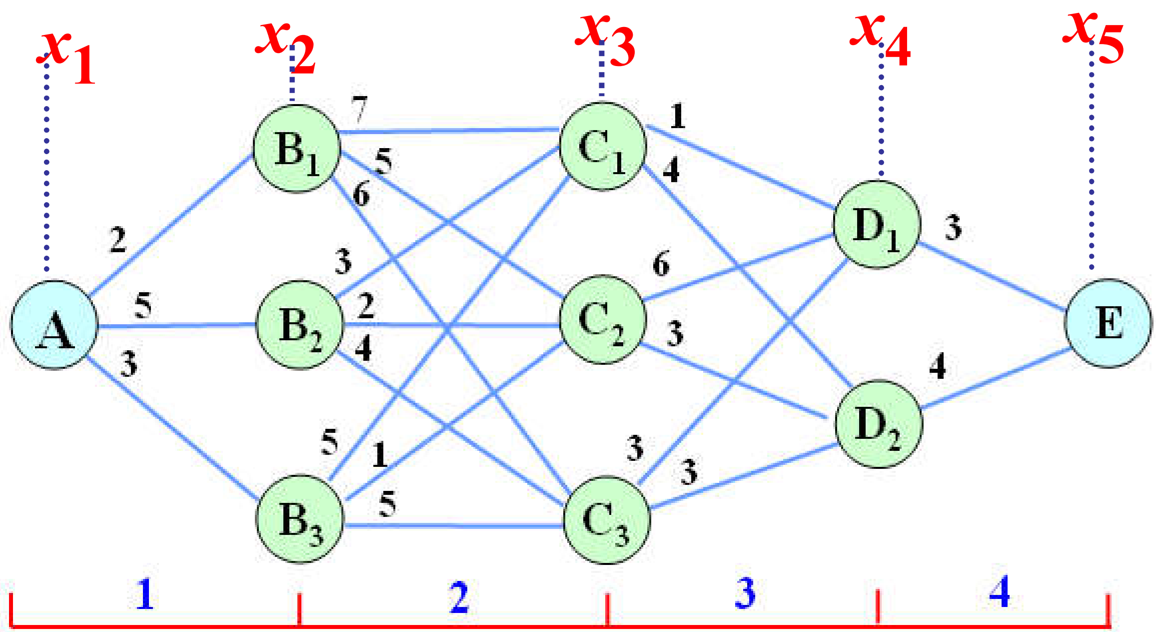
阶段 k

状态 x_k

状态集合 S_k

决策 $u_k(x_k)$

决策集合 $D_k(x_k)$。

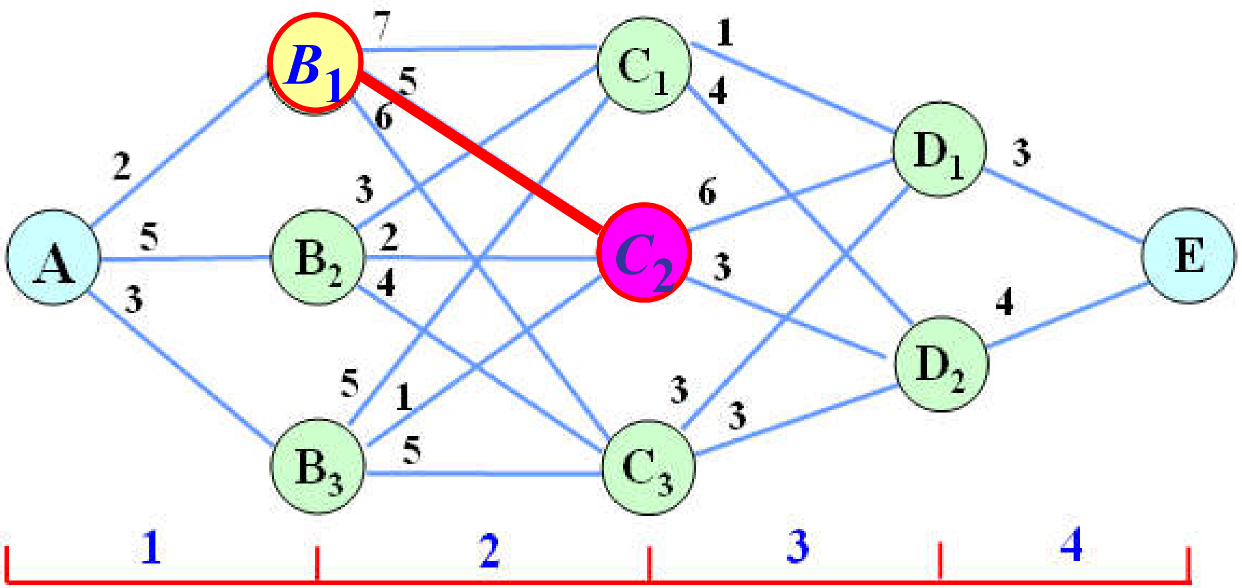


动态规划的基本概念

（4）状态转移律（方程）

□**状态转移律**：从x_k的某一状态值出发，当决策变量$u_k(x_k)$的取值决定后，下一阶段状态变量x_{k+1}的取值也随之确定。描述从x_k转变为x_{k+1}的规律称为状态转移规律（方程）。

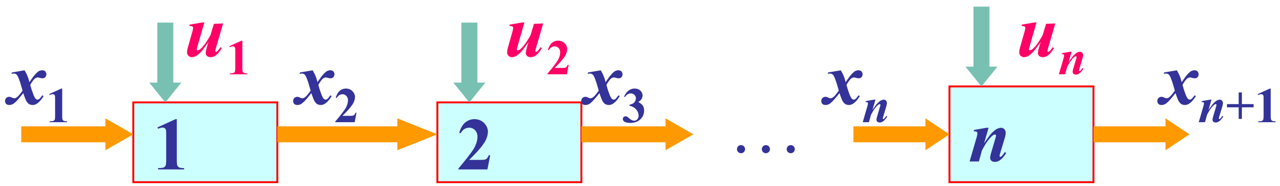
□从第2阶段的状态B1出发，如我们决定选择C2（也即确定了下一阶段的状态）。



动态规划的基本概念

□一般来说，下一阶段状态变量x_{k+1}的取值是上阶段的某一状态变量x_k和上阶段决策变量$u_k(x_k)$的函数，记为

$$x_{k+1}=T_k(\, x_k,\, u_k(x_k)\,)$$

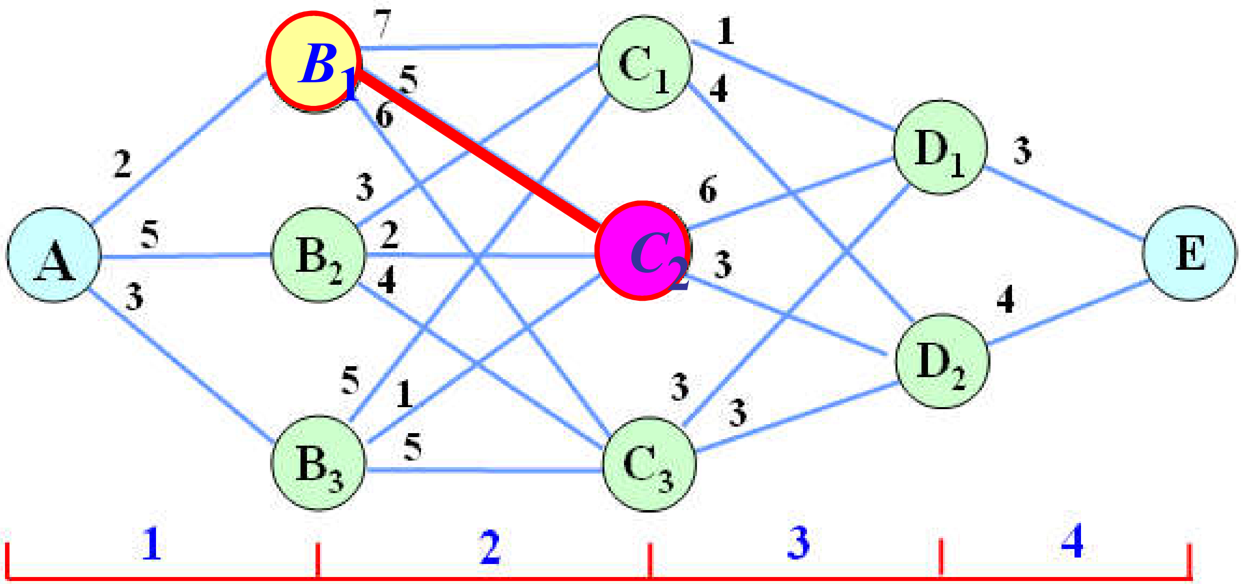


□上例中，

$u_2(\text{ B1 }) = \text{C2}$

□状态转移律为：

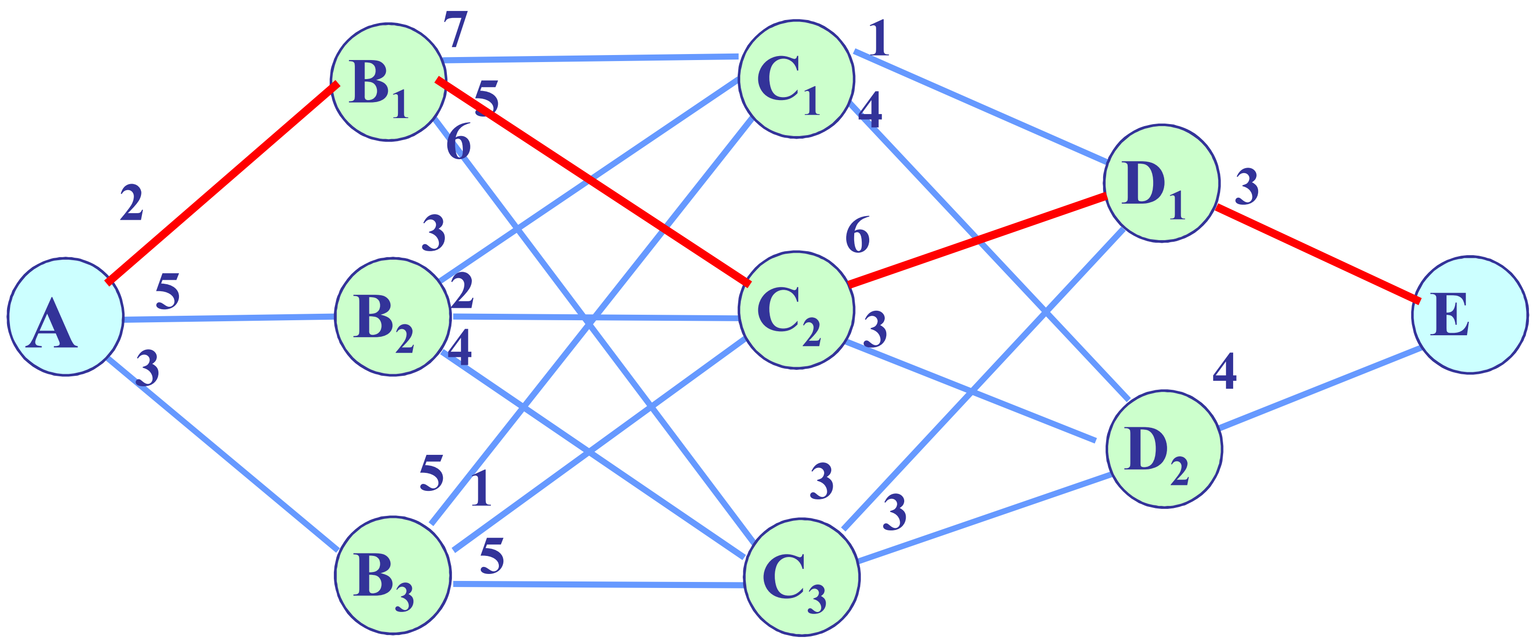
$$x_{k+1} = u_k(\, x_k \,)$$



动态规划的基本概念

（5）策略（policy）和子策略（subpolicy）

□**策略**：由依次进行的n个阶段决策构成的决策序列就构成一个策略，用 $p_{1n}\{ u_1(x_1), u_2(x_2), \ldots, u_n(x_n) \}$ 表示。

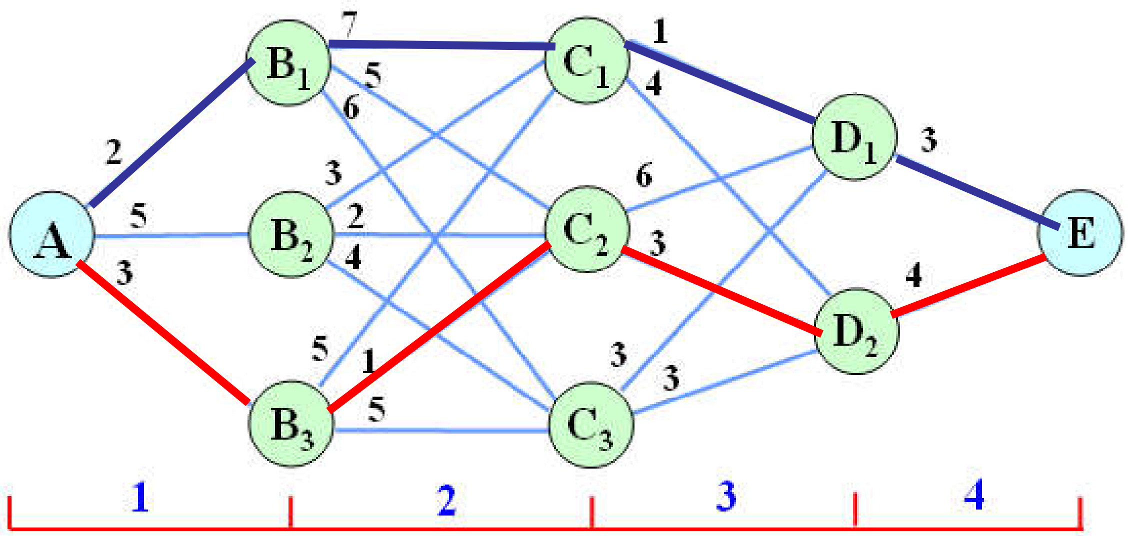


□本例中，如$p_{14}\{ u_1(A)=B_1, u_2(B_1) = C_2, u_3(C_2) = D_1, u_4(D_1) = E \}$表示其中一个策略，其总距离为2+5+6+3=16。

动态规划的基本概念

□**策略集合：**在实际问题中，由于在各个阶段可供选择的决策有许多个，因此，它们的不同组合就构成了许多可供选择的决策序列（策略），由它们组成的集合，称为策略集合，记作 P_{1n}。

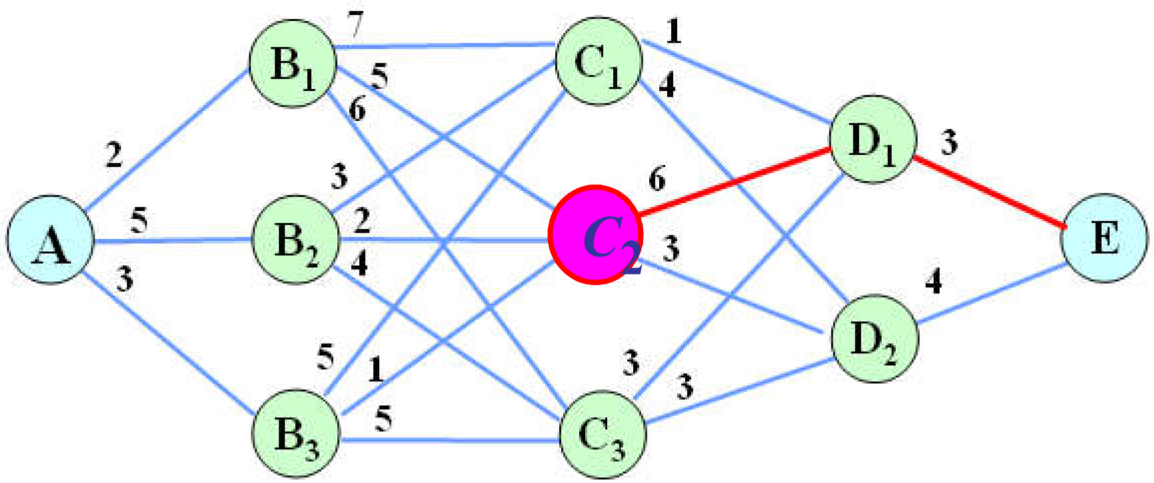
□从策略集合中，找出具有最优效果的策略称为**最优策略**。



动态规划的基本概念

□**子策略**：从k阶段到第n阶段，依次进行的阶段决策构成的决策序列称为k部子策略，表示为

$$p_{kn} = \{ u_k(x_k), u_{k+1}(x_{k+1}), \ldots, u_n(x_n) \}$$



□如从第3阶段的C2状态开始的一个子策略可表示：

$$p_{34}=\{u_3(C_2) = D_1, \; u_4(D_1) = E \}$$

动态规划的基本概念

（6）指标函数

□用来衡量策略或子策略或决策的效果的某种**数量指标**，就称为**指标函数**。

□它是定义在全过程或各子过程或各阶段上的确定数量函数。对不同问题，指标函数可以是诸如费用、成本、产值、利润、产量、耗量、距离、时间、效用，等等。

阶段

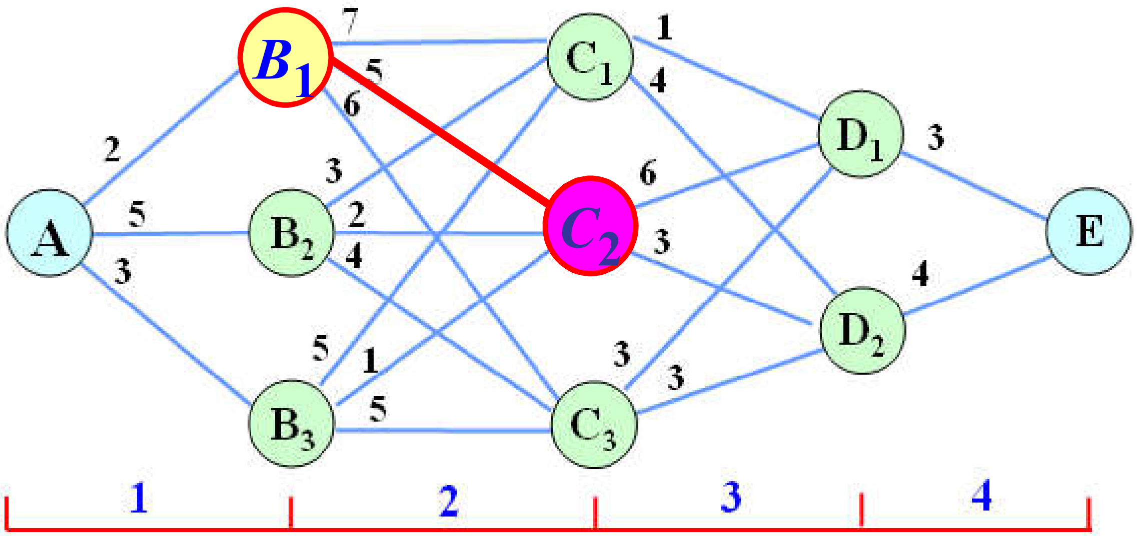
指标函数

过程

指标函数

动态规划的基本概念

①**阶段指标函数**：是指第 k 阶段从状态 x_k 出发，采取决策 u_k 时产生的效益，用 $v_k(x_k, u_k)$ 表示。



□例1中，指标函数是**距离**。

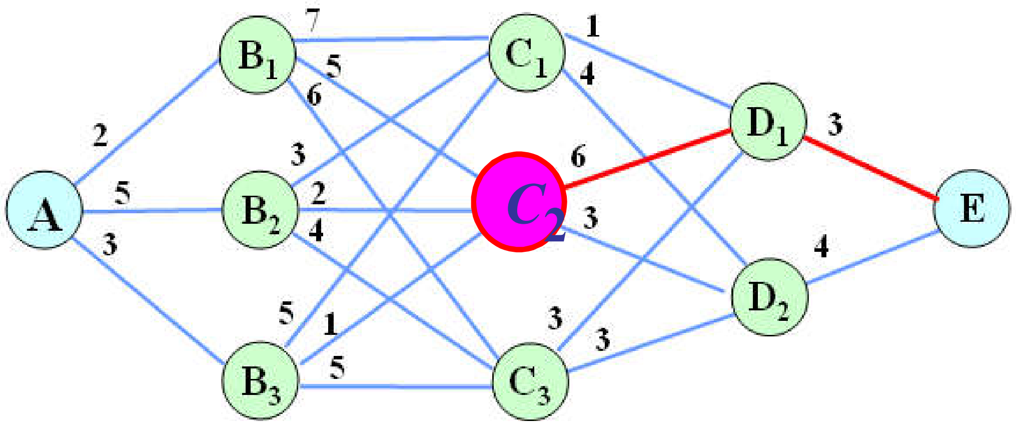
□如 $v_2(B_1, C_2)$ 表示由 B_1 出发，采用决策到 C_2 点的两点间距离，即

$$v_2(B_1, C_2) = 5。$$

动态规划的基本概念

②**过程指标函数**：是指从第 k 阶段的某状态 x_k 出发，采取子策略 p_{kn} 时所得到的效益，记作

$$V_{kn}(x_k, u_k, x_{k+1}, u_{k+1}, \ldots, x_n)$$



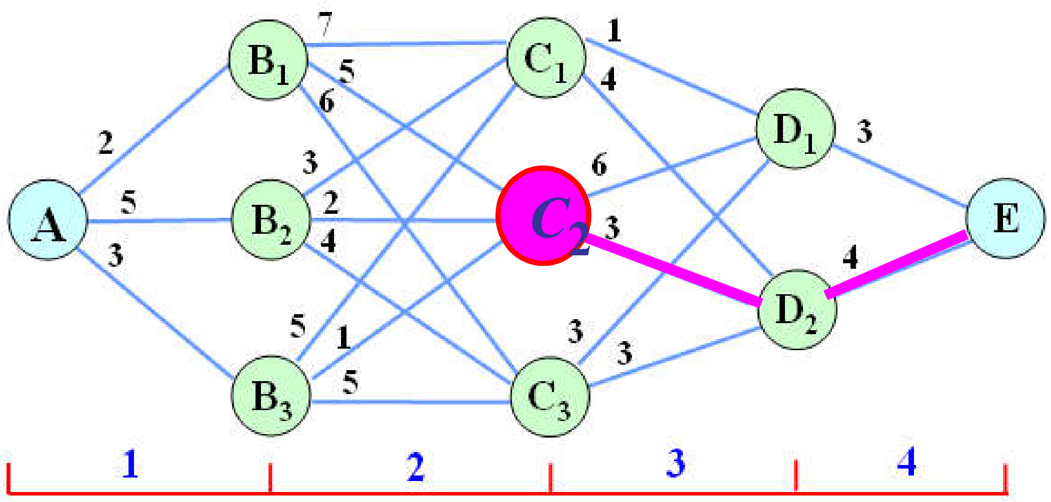
□例1中，如 $V_{34}(C_2, u_3(C2)=D_1, D_1, u_4(D_1)=E, E) = 6+3 = 9$

动态规划的基本概念

③**最优指标函数**：表示从第 k 阶段状态为 x_k 时采用最优策略 p_{kn}*到过程终止时的最佳效益值。记为 $f_k(x_k)$。

$$f_k(x_k) = V_{kn}(x_k, p_{kn}^*) = \underset{p_{kn} \in P_{kn}}{\text{opt}} V_{kn}(x_k, p_{kn})$$

其中 opt 可根据具体情况取*max* 或*min*。



□例1中，如 $f_3(C_2) = 3+4 = 7$。

动态规划的目标？

◎最优解：最优策略 p_{1n}

◎最优值：最优指标 $f_1(A)$

动态规划的基本概念

□多阶段决策问题的数学模型

$$f_k(s_k) = \underset{\substack{x_k \in D_k(s_k), \\ k=n,n-1,\cdots,2,1}}{opt} \{v_k(s_k, x_k) + f_{k+1}(s_{k+1})\}$$

$$f_{n+1}(s_{n+1}) = 0$$

最优指标函数

$$st. \begin{cases} x_{k+1} = T_k(x_k, u_k(x_k)) \\ u_k \in D_k \\ x_k \in S_k \\ k=1,2,\ldots,n \end{cases}$$

状态转移方程

决策变量(函数)

状态变量

阶段变量

动态规划举例

例1 机器负荷分配问题（时间阶段问题）

◎设有某种机器设备，用于完成两类工作A和B。若第k年初完好机器的数量为x_k，若以数量u_k用于A，余下的（$x_k - u_k$）用于工作B，则该年的预期收入为$g(u_k)+h(x_k - u_k)$。其中

$g(u_k) = 8u_k$，$h(x_k - u_k) = 5(x_k - u_k)$。

◎又机器设备在使用中会有损坏，设机器用于工作A时，一年后能继续使用的完好机器数占年初投入量的70%；若用于工作B时，一年后能继续使用的完好机器数占年初投入量的90%。则在下一年初能继续用于A、B工作的设备数为 $x_{k+1}=0.7u_k+0.9(x_k - u_k)$。

◎设第1年初完好的机器总数为1000台，问在连续5年内每年应如何分配用于A、B两项工作的机器数，使5年的总收益为最大。

动态规划举例

1.划分阶段　　按年度来划分阶段，k=1,2,3,4,5

2.正确选择状态变量　　状态变量x_k为第k年度初拥有的完好机器数量

3.确定决策变量及允许决策集合

决策变量u_k为第k年度中分配于A工作的机器数量，则x_k-u_k为用于B工作的机器数量。

第k阶段决策集合$D_k(x_k)=\{u_k \mid 0\le u_k\le x_k\}$

◎这里x_k和u_k均取连续变量，它们的非整数值可以这样理解，如x_k=0.6，就表示一台机器在第k年度中正常工作时间只占6/10；u_k=0.3，就表示一台机器在该年度只有3/10的时间能正常用于A工作。

动态规划举例

4.确定状态转移方程

状态转移方程为 $x_{k+1}=0.7u_k+0.9(x_k-u_k)$

5.确定阶段指标函数和最优指标函数，建立动态规划基本方程

阶段指标函数为

$$V_k(x_k, u_k) = g(u_k) + h(x_k - u_k)$$

令最优指标函数$f_k(x_k)$表示由资源量x_k出发，从第k年开始到第5年结束时所取得的最大预期收入。因而有：

$$f_k(x_k) = \max_{0 \le u_k \le x_k}\{V_k(x_k, u_k) + f_{k+1}(x_k)\}$$

$$f_6(x_6)=0$$

动态规划举例

□机器负荷分配问题 动态规划模型

$f_k(x_k)$ 表示由资源量x_k出发，从第k年开始到第5年结束时所取得的最大预期收入

$$f_k(x_k) = \max_{0\le u_k\le x_k}\{V_k(x_k, u_k) + f_{k+1}(x_k)\}$$

$$f_6(x_6)=0$$

$$st.\begin{cases} x_{k+1}=0.7u_k+0.9(x_k-u_k) & \text{状态转移方程为} \\ D_k(x_k) = \{u_k \mid 0\le u_k\le x_k\} & \text{第}k\text{阶段决策集合} \\ 0\le x_k \le 1000 \quad k=1,2,\ldots,n & \text{第}k\text{阶段状态集合} \end{cases}$$

阶段指标函数为 $V_k(x_k, u_k) = g(u_k) + h(x_k-u_k)$