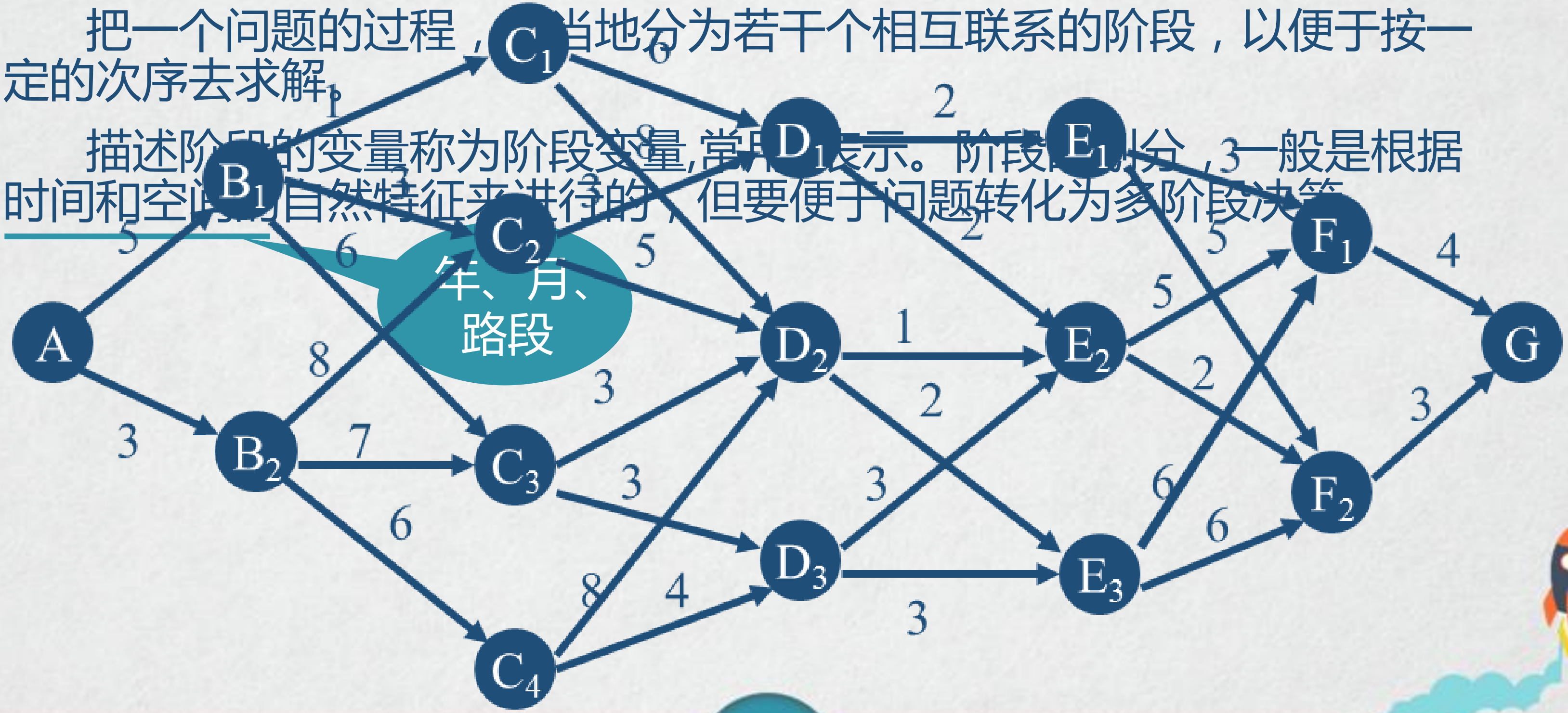
动态规划的基本概念

（一）、基本概念

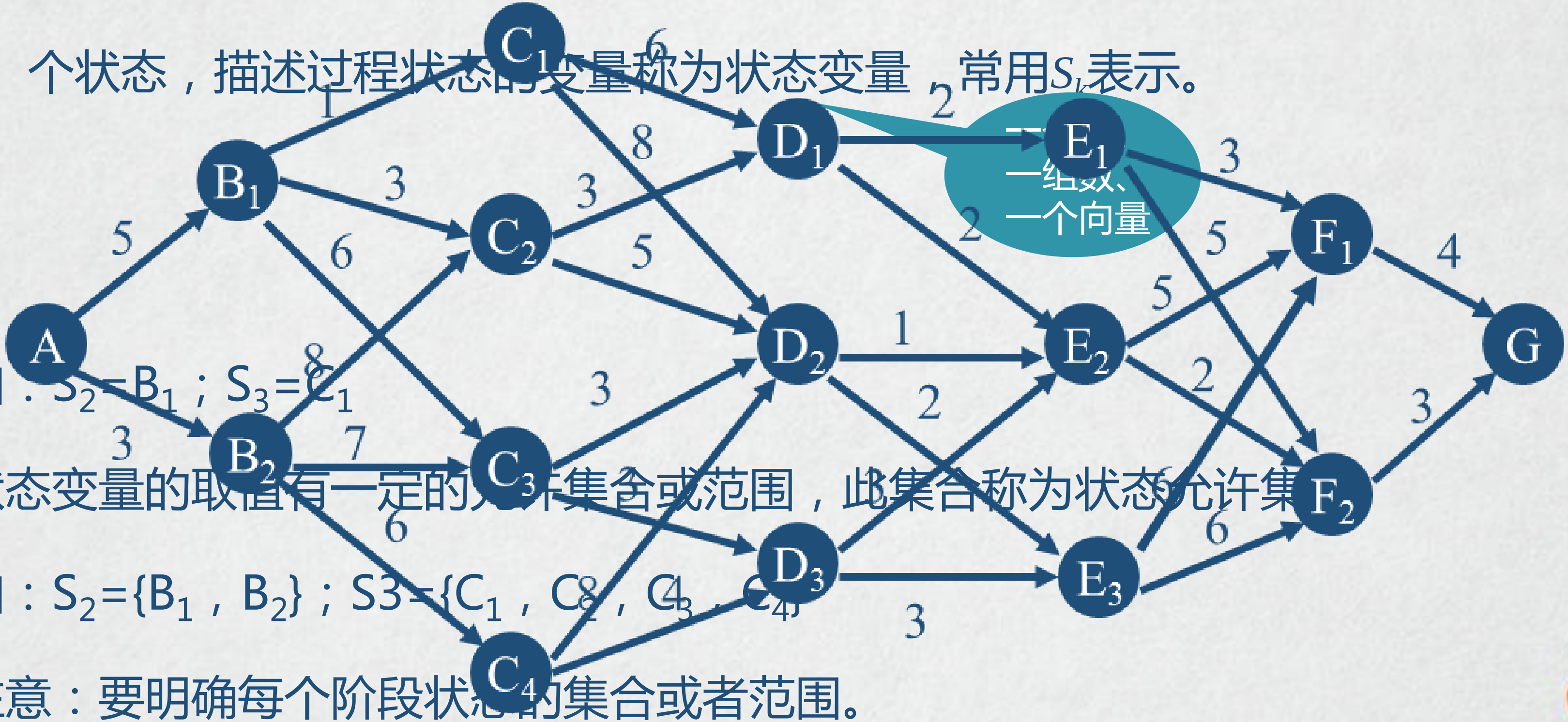
1、阶段：

把一个问题的过程，恰当地分为若干个相互联系的阶段，以便于按一定的次序去求解。

描述阶段的变量称为阶段变量,常用k表示。阶段的划分，一般是根据时间和空间的自然特征来进行的，但要便于问题转化为多阶段决策。



2、状态：表示每个阶段开始所处的自然状况或客观条件。通常一个阶段有若干个状态，描述过程状态的变量称为状态变量，常用S_k表示。



如：$S_2=B_1$；$S_3=C_1$

状态变量的取值有一定的允许集合或范围，此集合称为状态允许集合。

如：$S_2=\{B_1，B_2\}$；$S3=\{C_1，C_2，C_3，C_4\}$

注意：要明确每个阶段状态的集合或者范围。

“状态”具有 “无后效性”（ “马尔科夫性” ）：如果某阶段的状态给定后，当前的状态是以往历史的总结，则在这阶段以后过程的发展不受这阶段以前各阶段的影响。

3、决策：表示当过程处于某一阶段的某个状态时，可以作出不同的决定，从而确定下一阶段的状态，这种决定称为决策。

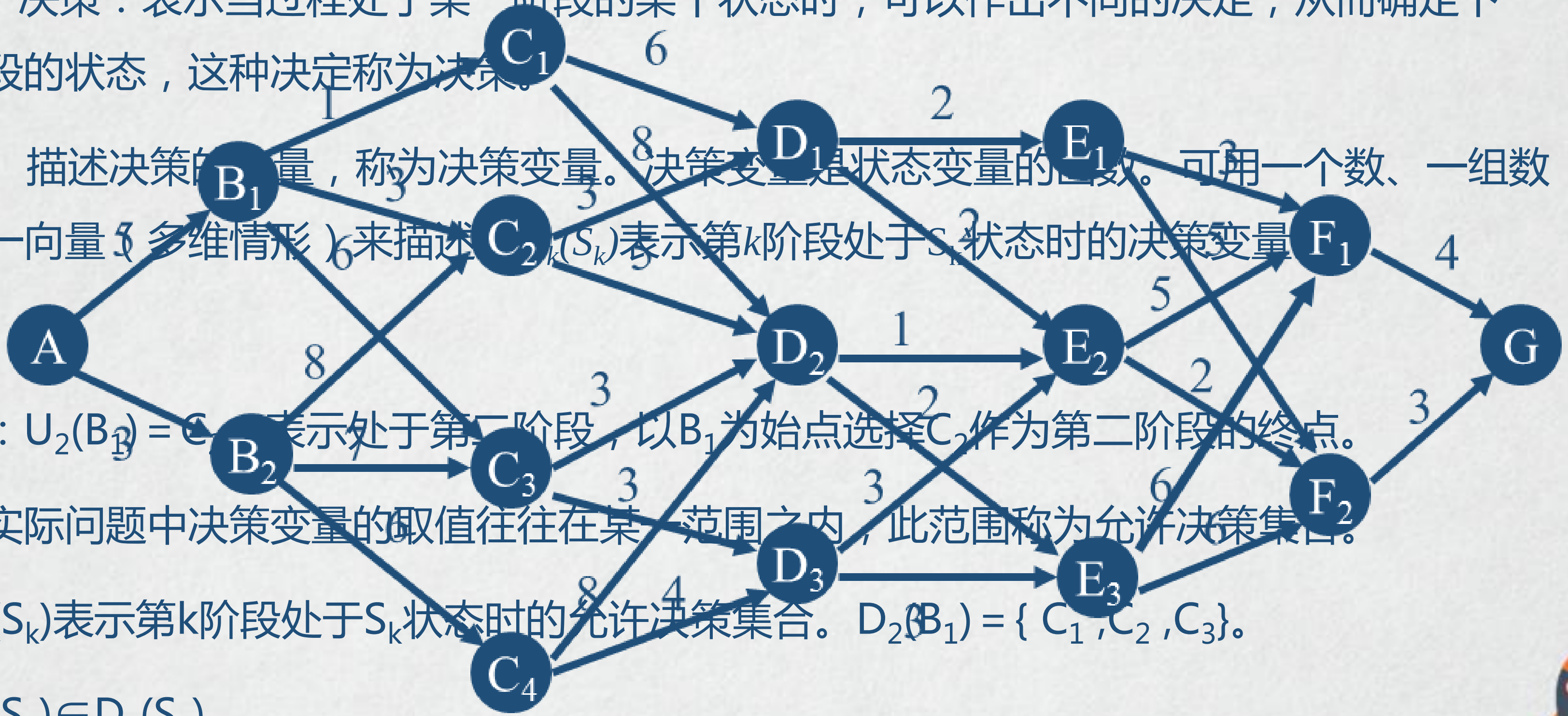
描述决策的变量，称为决策变量。决策变量是状态变量的函数。可用一个数、一组数或一向量（多维情形）来描述。$u_k(S_k)$表示第k阶段处于S_k状态时的决策变量。

如：$U_2(B_1)=C_2$，表示处于第二阶段，以B_1为始点选择C_2作为第二阶段的终点。

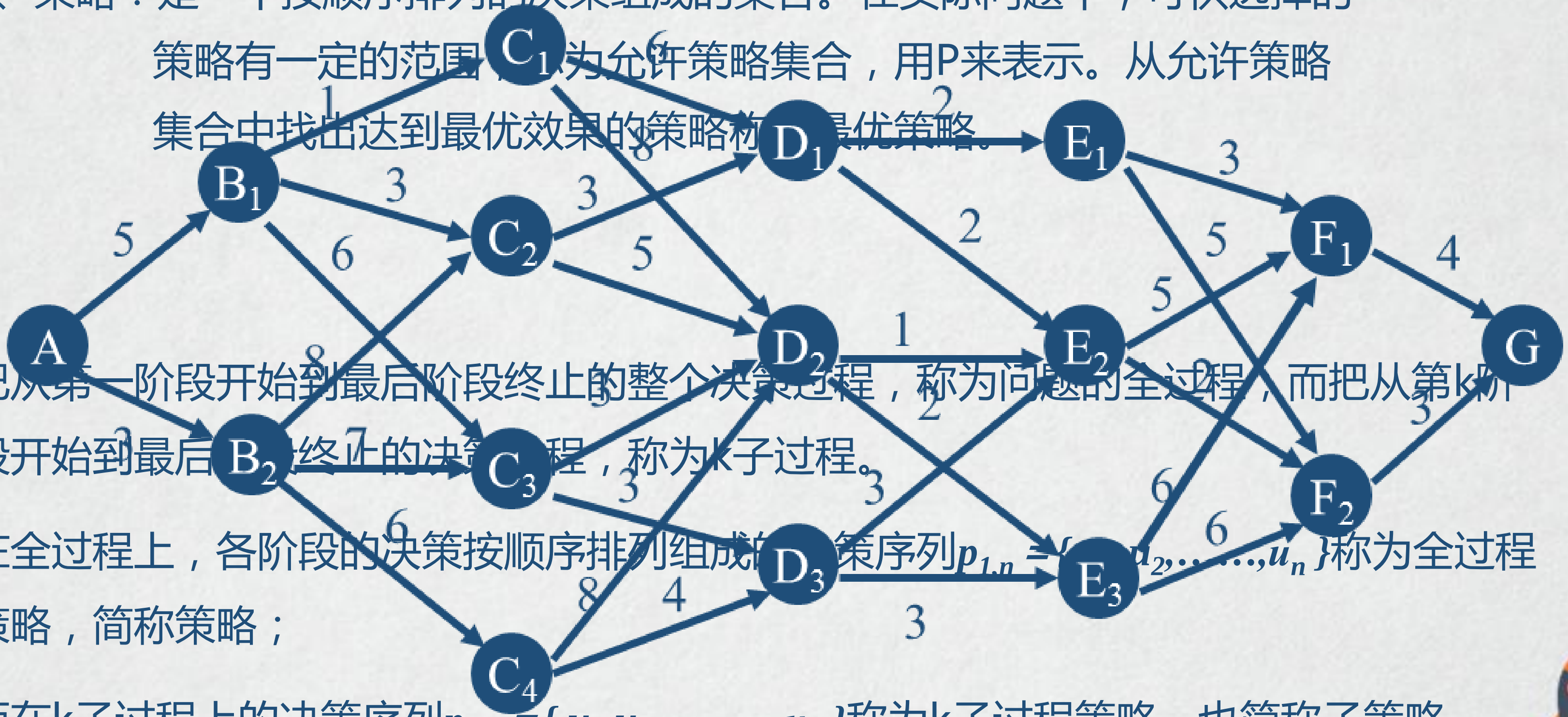
在实际问题中决策变量的取值往往在某一范围之内，此范围称为允许决策集合。

$D_k(S_k)$表示第k阶段处于S_k状态时的允许决策集合。$D_2(B_1)=\{C_1, C_2, C_3\}$。

$U_k(S_k)\in D_k(S_k)$。



4、策略：是一个按顺序排列的决策组成的集合。在实际问题中，可供选择的策略有一定的范围，称为允许策略集合，用P来表示。从允许策略集合中找出达到最优效果的策略称为最优策略。

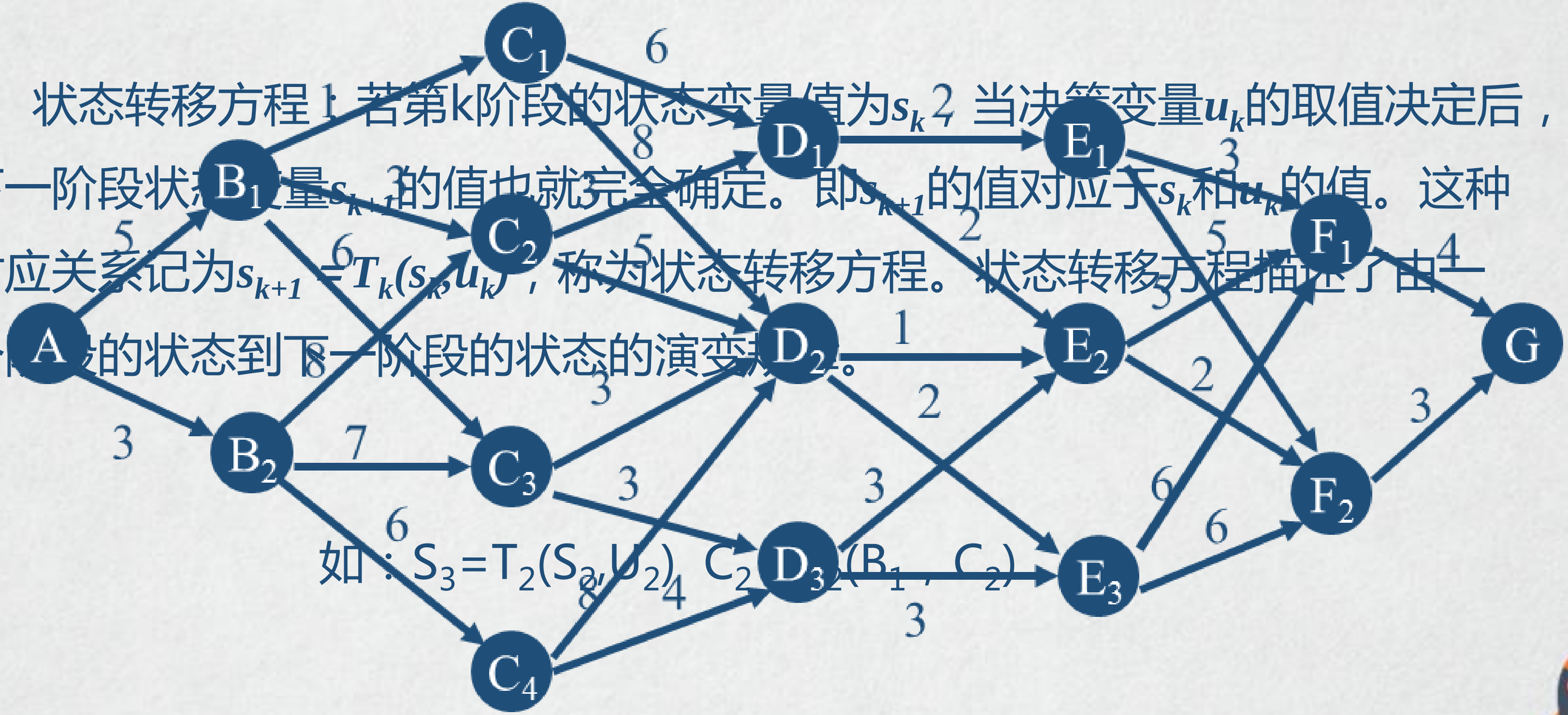


把从第一阶段开始到最后阶段终止的整个决策过程，称为问题的全过程，而把从第k阶段开始到最后阶段终止的决策过程，称为k子过程。

在全过程上，各阶段的决策按顺序排列组成的决策序列$p_{1,n}=\{u_1,u_2,\ldots\ldots,u_n\}$称为全过程策略，简称策略；

而在k子过程上的决策序列$p_{k,n}=\{u_k,u_{k+1},\ldots\ldots,u_n\}$称为k子过程策略，也简称子策略。

5、状态转移方程：若第k阶段的状态变量值为s_k，当决策变量u_k的取值决定后，下一阶段状态变量s_{k+1}的值也就完全确定。即s_{k+1}的值对应于s_k和u_k的值。这种对应关系记为$s_{k+1}=T_k(s_k,u_k)$，称为状态转移方程。状态转移方程描述了由一个阶段的状态到下一阶段的状态的演变规律。



如：$S_3=T_2(S_2,U_2)$，$C_2=T_2(B_1，C_2)$

状态变量要满足无后效性的要求；

如果状态变量不能满足无后效性的要求，应适当地改变状态的定义或规定方法。

状态具有无后效性的多阶段决策过程的状态转移方程如下

$$s_2 = T_1(s_1, u_1)$$
$$s_3 = T_2(s_2, u_2)$$
$$\cdots\cdots$$
$$s_{k+1} = T_k(s_k, u_k)$$

动态规划中能处理的状态转移方程的形式。

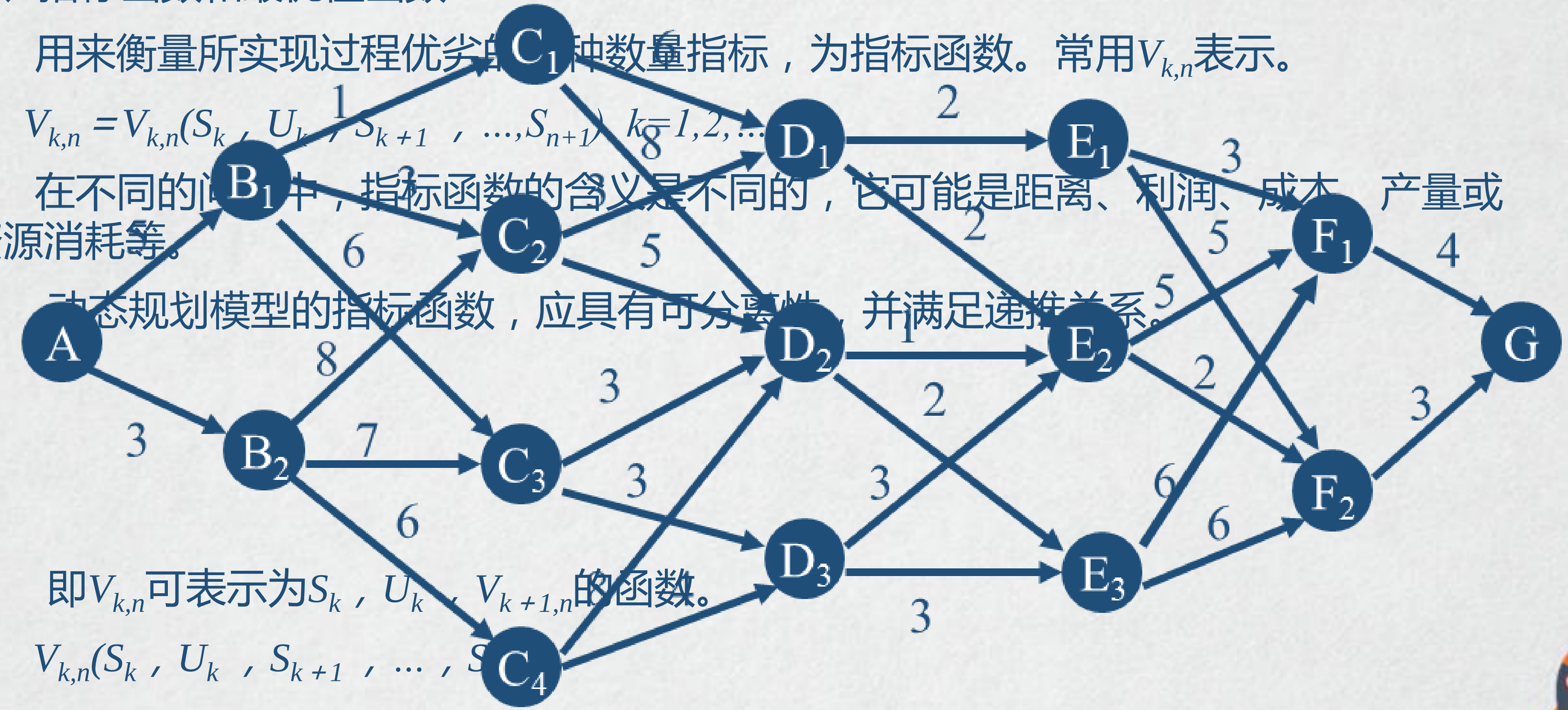
6、指标函数和最优值函数：

用来衡量所实现过程优劣的一种数量指标，为指标函数。常用$V_{k,n}$表示。

$V_{k,n}=V_{k,n}(S_k, U_k, S_{k+1}, \dots, S_{n+1})$ $k=1,2,\dots$

在不同的问题中，指标函数的含义是不同的，它可能是距离、利润、成本、产量或资源消耗等。

动态规划模型的指标函数，应具有可分离性，并满足递推关系。



即$V_{k,n}$可表示为S_k，U_k，$V_{k+1,n}$的函数。

$V_{k,n}(S_k, U_k, S_{k+1}, \dots, S_{n+1})$

$=\psi k[S_k, U_k, V_{k+1,n}(S_{k+1}, U_{k+1}, \dots, S_{n+1})]$

常见的指标函数有：

1)整个过程和它的任一子过程的指标函数是它所包含的各阶段的指标的和。

$$V_{k,n}\ (S_k,\ U_k,\ S_{k+1},\ \ldots,S_{n+1})=\sum_{j=k}^{n} v_j(s_j,u_j)$$

$$V_{k,n}\ (S_k,\ U_k,\ S_{k+1},\ \ldots,S_{n+1})$$

$$=v_k(s_k,u_k)+V_{k+1,n}\ (S_{k+1},\ U_{k+1},\ S_{k+2},\ \ldots,S_{n+1})$$

2)整个过程和它的任一子过程的指标函数是它所包含的各阶段的指标的乘积，即：

$$V_{k,n}\ (S_k,\ U_k,\ S_{k+1},\ \ldots,S_{n+1})=\prod_{j=k}^{n} v_j(s_j,u_j)$$

$$V_{k,n}(S_k,\ U_k,\ S_{k+1},\ \ldots,S_{n+1})$$

$$=v_k(s_k,u_k)\times V_{k+1,n}(S_{k+1},\ U_{k+1},\ S_{k+2},\ \ldots,S_{n+1})$$

指标函数的最优值称为最优值函数，记为$f_k(s_k)$。 表示从第k阶段的状态s_k开始到第n阶段的终止状态的过程，采取最优策略所得到的指标函数值。

如：$f_1(A)=18$，$f_2(B_1)=13$。

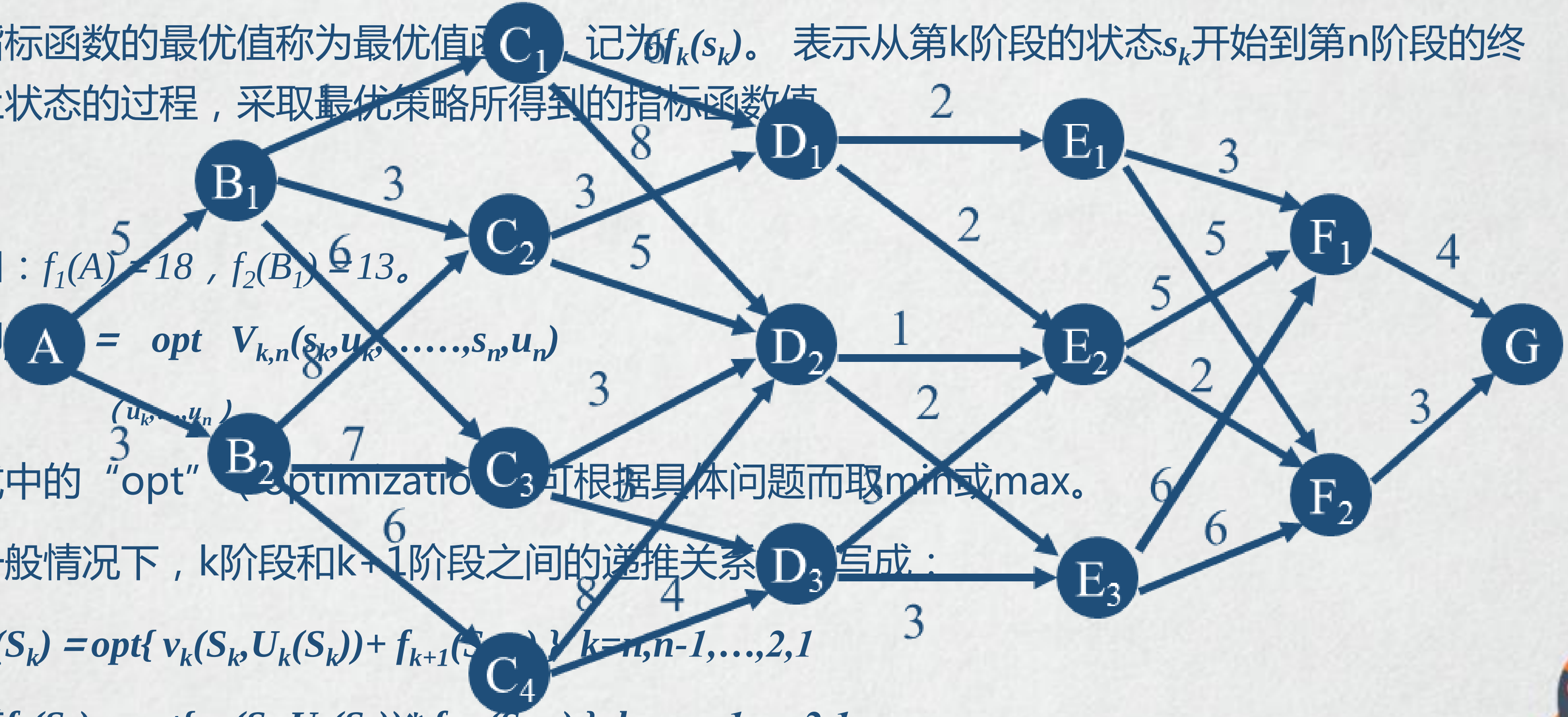
即$f_k(s_k) = \underset{(u_k,\ldots,u_n)}{opt}\ V_{k,n}(s_k,u_k,\ldots\ldots,s_n,u_n)$

式中的“opt”（optimization）可根据具体问题而取min或max。

一般情况下，k阶段和k+1阶段之间的递推关系式可写成：

$f_k(S_k) = opt\{ v_k(S_k,U_k(S_k)) + f_{k+1}(S_{k+1}) \}\ \ k=n,n-1,\ldots,2,1$

或$f_k(S_k) = opt\{ v_k(S_k,U_k(S_k)) * f_{k+1}(S_{k+1}) \}\ \ k=n,n-1,\ldots,2,1$



小结

动态规划本质上是多阶段决策过程;

无后效性

概念：阶段变量k、状态变量s_k、决策变量u_k;

方程 :状态转移方程 $s_{k+1} = T_k(s_k, u_k)$

指标: $V_{k,n} = V_{k,n}(s_k, u_k, s_{k+1}, u_{k+1}, \cdots, s_{n+1})$

效益

$$f_k(s_k) = \underset{\{u_k, \cdots, u_n\}}{opt} V_{k,n}(s_k, u_k, \cdots, s_{n+1})$$

$$V_{k,n}(s_k, u_k, s_{k+1}, u_{k+1}, \cdots, s_{n+1})$$

可递推

$$= \psi_k[s_k, u_k, V_{k+1,n}(s_{k+1}, u_{k+1}, \cdots, s_{n+1})]$$

指标函数形式:和、积

解多阶段决策过程问题，求出

最优策略，即最优决策序列

$$\{u_1^*, u_2^*, \cdots, u_n^*\}$$

优轨线，即执行最优策略时的状态序列

$$\{s_1^*, s_2^*, \cdots, s_n^*$$

最优目标函数值

从 k 到终点最优策略

子策略的最优目标函数值

$$V_{1,n}^* = V_{1,n}(s_1^*, u_1^*, \cdots, u_n^*)$$

$$f_k(s_k) = \underset{\{u_k, \cdots, u_n\}}{opt} v_{k,n}(s_k, u_k, \cdots, s_{n+1})$$