五、前缀码问题:

通讯中常用二进制码表示字母,4 位二进制码可以表示 $2^4$=16 个不同字母,因而表示 26 个英文字母必须用五位二进制码,这是等字长码,便有的字母使用频率很高例 e——13.1%,t——10.5%,还有 a,i,r,s 等,而有的英文字母使用频率很低例 j,z,q 只使用 0.1%,是否可以用不等长字符串的码,频率高的用的码字符串短些,频率低的字长可略长些,从而使电文的总长会有所降低,但不等长字符串有时会造成二义性。

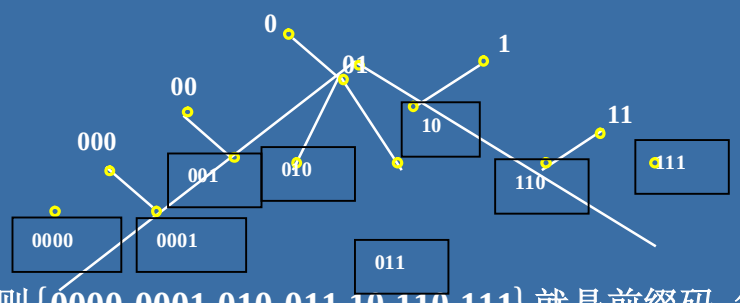
例:

| a | b | c | d | e |
|---|---|---|---|---|
| 00 | 110 | 010 | 10 | 01 |

把集合{00,110,010,10,01}称为码,如收到电文是 010010,就有二义性,理能为 01,00,10 电文应是 ead 如理能为 010,010 就是 CC,造成二义性的原因是"e"01 是"c"对序列 010 的前缀,如把 C 改为 111,则集合{00,110,111,10,01}没有那个序列是另一个序列的前缀,就不会造成二义性,这种码称为前缀码。

定义 18:设 $a_1a_2...a_n$ 是字长为 n 的字符串,称子串 $a_1$, $a_1a_2$, $\cdots$, $a_1a_2...a_{n-1}$ 分别为该字符串的长度为 n 的前缀,若 A=$\{b_1,b_2,...b_m\}$ 为一符号串集合,对于任意的两个字符串 $b_i,b_j$ Î $A(i$ ¹ $j)$, $b_i$ 与 $b_j$ 互不为前缀,则称 A 为前缀码,若符号串 $b_i(i=1,...,m)$ 中只出现 0,1 两个符号,则称 A 为二元前缀码。

定理 21:任意一棵二叉树，均可以产生的一个前缀码。




则$\{0000,0001,010,011,10,110,111\}$就是前缀码,每个码的前缀均应是分枝点。

定理 22:任何一个前缀码均对应一棵二叉树。

从图中可以看出前缀码的每个字符串长度正好对应该叶子的层数。如果电文长度为 10000 个字母,字母$a_1,...,a_n$出现的次数为$w_1,w_2,...,w_n$,$a_i$代码字符串长度为$L(w_i)$正好为该叶的层数,则 10000 个字母的电文的总字符长度正好是 å $w_iL(w_i)$正好是这棵二叉树的权。因此求电文总长最小就转化为求最优树问题。

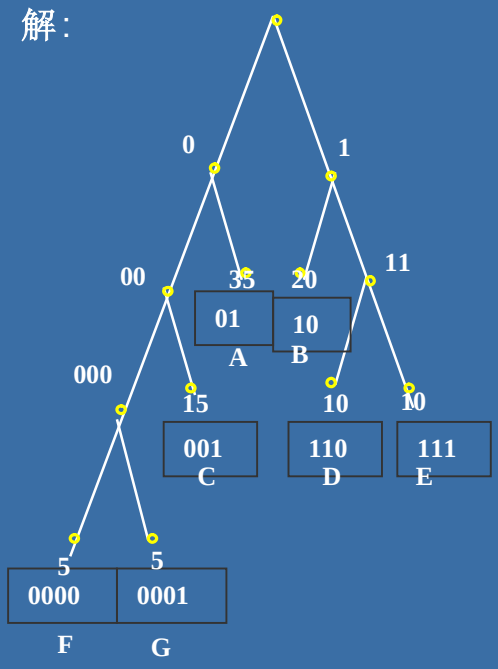
例:已知 ABCDEFG 出现的频率分别为

A-35%,B-20%,C-15%,D-10%,E-10%,F-5%,G-5%。

(1) 求带权 5,5,10,10,15,20,35 的最优数。

(2) 求 T 对应的前缀码。

解:




该图表示最优数

$w(T)$=(5+5)´4+

(10+10+15)´3+

(35+20)´2=265

传输 1000 个这些字母的

电文共需用二进制数字个

数为 1000´265%=2650 个

如用等长码需 3000 个

§6.5 二分图

一、二分图的定义:

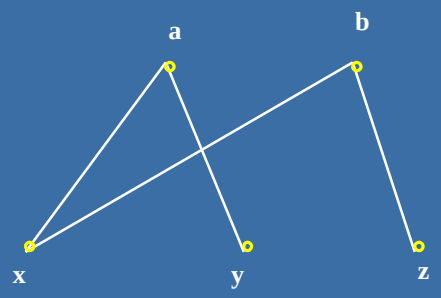
定义 19:如果无向图 $G=<V,E>$ 的结点集 V 分成两个子集 $V_1$, $V_2$ (即满足 $V_1 \cap V_2=\Phi$, $V_1 \cup V_2=V$), 使得 G 中任意一边的两个端点分属于 $V_1$ 和 $V_2$,

则称 G 为<u>二分图</u>或偶图(或称二部图)，$V_1$和$V_2$称为互补结点子集，常记图为$G=<V_1,V_2,E>$。

若$V_1$中的每个结点与$V_2$的每个结点都有且只有一个一条边相关联，称 G 为<u>完全二分图</u>，记为$K_{n,,m}$，其中$|V_1|$=n，$|V_2|$=m。

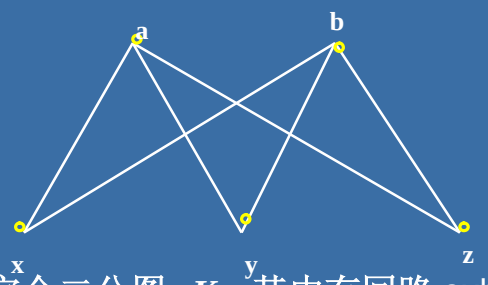
书上说“二分图没有回路”这是错的。

例：



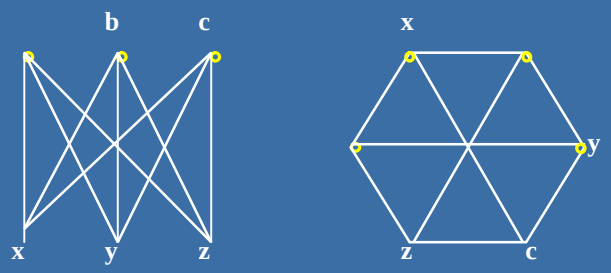

$V_1=\{a,b\}$，$V_2=\{x,y,z\}$任意$e$ Î $E$，e 的两个端点没有全在$V_1$中，也没有全在$V_2$中。

例：




这是完全二分图，$K_{3,2}$其中有回路 axbya。

例：



a

此两图均是 $K_{3,3}$ 因而是同构的。

定理 23:无向图 G 是二分图的充要条件是 G 的所回路的长度均为偶数。

证明:必要性(G 是二分图证明其任何回路长度均为偶数)

设 $G=<V_1,V_2,E>$ 是二分图,$C=V_0V_1....V_kV_0$ 是 G 的任一个回路,其长度为 $K+1$。如 $V_0$ Î $V_1$,由定义

$(V_0,V_1)$ Î $E$,\ $V_1$ Î $V_2$,$(V_1,V_2)$ Î $E$,\ $V_2$ Î $V_1$,....,

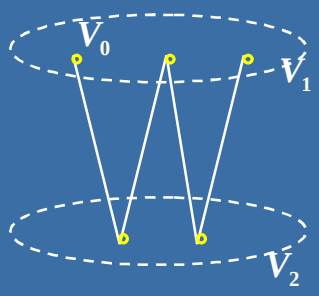
由此得 $V_{2i}$ Î $V_1$,$V_{2i+1}$ Î $V_2$,因此

$(V_k,V_0)$ Î $E$,\ $V_k$ Î $V_2$,\ $k$ 是奇数,\ $k+1$ 是偶数,

\ 此回路的长度是偶数。

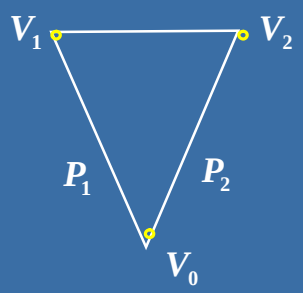
充分性(G 的每个回路长度均为偶数,证明 G 为二分图)

设 G 是连通图,任取 $V_0$ Î $V$。




定义 V 的两个子集 $V_1,V_2$,$V_1=\{V\mid V_0$ 到的通路长度为偶数$\}$,$V_2=V-V_1$,现证明,$V_1$ 中任何两点之间没有边。

反证,如 $V_1$ 中存在两点 $v_1,v_2$ Î $V_1$,而 $v_1,v_2$ 之间有边,由 $V_1$ 的定义,$v_0$ 到 $v_1$ 有一条通路 $P_2$ 长度为偶数,则回路 $v_0P_1v_1v_2P_2v_0$,其长度为 $P_1$ 长度加上 $P_2$ 长度再加上 1(即边 $(v_1,v_2)$)与 G 的任何回路长度为偶数矛盾。



而 G 是连通图，故$V_2$就是到$v_0$点通路长度为奇数的全体,用同样的方法可以证明$V_2$内任何两点间也没有边。∴G 是二分图，互补结点子集就是$V_1$,$V_2$。

对于非连通,可以对每个连通分枝进行以上处理,推论：任何无回路的图均是二分图。

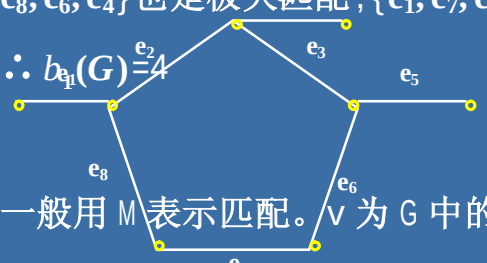
二、边匹配和边独立集，

定义 20：在图 G=<V，E>中，若存在 $E^*$ Í E，使得 $E^*$中的任何两边均不相邻，称 $E^*$为 G 的一个匹配，或称 $E^*$为边独立集。

若 $E^*$是 G 的一个匹配，$E^*$中再加上 G 中的任何一条边 e，则 $E^*$∪{e}就不是 G 的匹配了，则称 $E^*$为 G 中的极大匹配，G 中边数最多的匹配称为 G 的最大匹配。G 的最大匹配的边数称为 G 的匹配数记为 $b_1(G)$。

图中$\{e_8, e_3\}$是一个极大匹配，其他任一条边加入均会有边相邻了，但它不是最大匹配。

$\{e_8, e_6, e_4\}$也是极大匹配，$\{e_1, e_7, e_5, e_4\}$是极大匹配同时也是个最大匹配。∴ $b_1(G)=4$

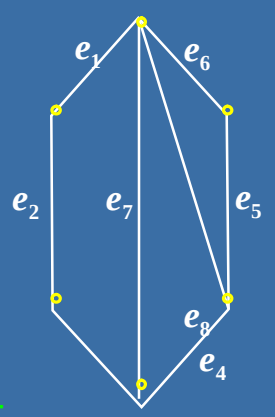



一般用 M 表示匹配。v 为 G 中的点，若存在 M 中的边与 v 相关联，称 v 为 M 的饱和点，否则称 v 为 M 的非饱和点。若 G 中每个点均在 M 中饱和，称 M 是 G 的完美匹配。

显然，若 M 是 G 的一个匹配，则在 M 中饱和点的个数等于 M 中边的个数的两倍，因匹配中两条边不可能同时与一个点关联。

因而一个图存在完美匹配的必要条件是其点数为偶数。完美匹配一定是最大匹配，而最大匹数未必是完美匹配。

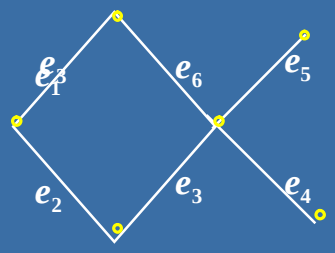
例：




$\{e_1,e_4\}$是极大匹配,$\{e_3,e_8\}$是极大匹配,$\{e_1\ e_3\ e_5\}$是最大匹配又是完美匹配,3 条边,6 个点,每个点均饱和了。

$\therefore b_1(G)=3$

例：




$\{e_1,e_3\},\{,e_6\},\{e_1,e_5\}$均是极大匹配,也是最大匹配,但均不是完美匹配,它不存在完美匹配 ,尽管 6 个点,不存在 3 条边的匹配, $\therefore$ $b_1(G)=2$

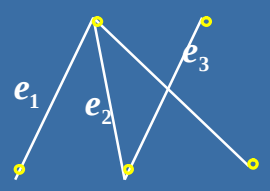
三、二分图与完备匹配

定义 21:已知二分图 $G=<V_1,V_2,E>$,M 是 G 的一个匹配则称 M 为二分图的匹配,若 M 为 G 的一个匹配,且$|M|=\min(|V_1,|V_2|)$,则称 M 为 G 中的一个完备匹配。

若$|V_1|\leqslant|V_2|$称此为 $V_1$到 $V_2$的一个匹配。此时完备匹配必是一个最大匹配,因 G 的每个端点分属 $V_1,V_2$,因而 G 的最大匹配的边数均等于 $V_1,V_2$分别在 M 中的饱和点数,反之最大匹配未必是饱和匹配。

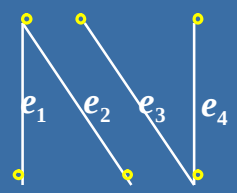
如果二分图 G 存在一个完美匹配,则必有$|V_1|=|V_2|$,反之$|V_1|=|V_2|$未必该二分图存在完美匹配。

例 1:



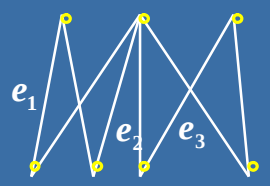

$\{e_2\}$是一个极大匹配。$\{e_1, e_3\}$是个最大匹配，也是完备匹配,称$V_1$到$V_2$的一个匹配, $V_1$的点均饱和了但$V_2$点均未饱和。

例 2:



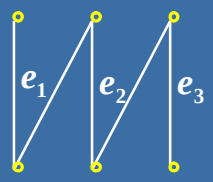

$\{e_1, e_3\}$是个最大匹配,但它不是完备匹配,也不存在完美匹配,尽管$|V_1| = |V_2| = 3$, $b_1(G)$=2

例 3:




其中$\{e_1, e_2, e_3\}$是一个完备匹配,
因$|V_1|$ 1 $|V_2|$不可能存在完美匹配,
$b_1(G)$=3

例 4:




其中$\{e_1, e_2, e_3\}$是完备匹配,也是一个完美匹配。 $b_1(G)$=3