图论应用

2004数学建模竞赛辅导

复旦大学计算机科学与工程系

吴永辉　　博士

2004年7月12日

基础知识

- 离散数学（图论）
- 程序设计语言
- 数据结构
- 算法设计与分析

主要参考书目

1 周咏基, 王建德. 金牌之路——竞赛解题指导. 陕西师范大学出版社. 2001.

2 吴文虎, 王建德. 实用算法的分析与程序设计. 电子工业出版社. 1998.

3 吴文虎, 王建德. 图论的算法与程序设计. 清华大学出版社. 1997.

4 朱洪等. 离散数学教程. 上海科技文献出版社. 1966.

5 雷功炎. 数学模型讲义. 北京大学出版社. 2000.

6 Thomas H. Coremen, etc. Introduction to Algorithm（算法导论）. 高等教育出版社, The MIT Press. 2002

7 Kenneth H. Rosen. Discrete Mathematics and its Applications（离散数学及其应用）. (4th Edition). 2002. 机械工业出版社, McGraw-Hill. (中、英文版)

8 Bela Bollobas. Modern Graph Theory (现代图论). 科学出版社 & Springer. 2001

图论综述

- **1 图的基本概念**

  图的基本概念、路与回路、欧拉图、哈密顿图、最短路
- **2 平面图**

  平面图与欧拉公式
- **3 图的着色**

  顶点着色、平面图的着色、边的着色
- **4 树**

  树的性质、生成树与割集、树的计数、有根树与二分树、最优树
- **5 网络流**

  连通度与块、网络最大流、图与二分图的匹配、独立集和覆盖、Petri网

对应策略

- 将问题A对应另一个便于思考或有求解方法的问题B，化繁为简，变未知为已知。
- 将现实生活中问题转化为图问题

1）经典图论问题解析

2）图论问题解析

一、经典图论问题解析

- 将现实生活中问题转化为图问题的经典问题：
- ***1 中国邮递员问题***
- ***2 4×4的黑白格棋盘跳马***
- ***3 过河问题***

1 中国邮递员问题

- 中国邮递员问题（中国数学家管梅谷）

  背景：

  一个邮递员从邮局出发投递信件，他必须在所管辖范围内的所有街道至少走一次，最后回到邮局。

  问题：

  选择一条最短的路线完成投递任务。

1）建立数学模型 $\Rightarrow$ 用图描述问题

$G=(V, E, \omega)$为一边带权的图。

V中元素为街道交叉点；

E为街道集合；

街道的长度为街道对应边的权，显然所有权均大于0。

2）问题分析 $\Rightarrow$ 如何选择路线？

- 设$G=(V, E, \omega)$为一边带权的图，所有权都非负。邮递员问题就转化为：
- 在G中，从某点v出发求一条经过每条边至少一次的闭路径，使该闭路径所带的权最小。
- 满足条件的闭路径称为最优投递路线

3）算法分析（解决问题的方法）：图论，分而治之

- 若G为欧拉图，最优投递路线就是从指定顶点出发的一条欧拉回路。
- 若G不是欧拉图，则最优投递路线必须要有重复边出现，而要求重复边权之和达到最小。

G必有奇顶点，为了消除奇顶点，必须加若干条重复边，使得重复边的权与原边的权相同，设所得图为G^*，则最优投递路线等价于求G^*的一条欧拉回路，使得重复边权之和 $\sum_{e \in F} \omega(e)$ 最小，其中$F=E(G^*)-E(G)$。

4）G不是欧拉图的最优投递路线算法

当G中只有两个奇顶点u, v时，可先用标号法求出从u到v的最短路，然后将最短路上的边连其权各重复一次，得欧拉图，再走一条欧拉回路，就是G中的一条最优投递路线。

5）理论基础

- 定理1：

　　C是带权无向连通图$G=(V, E, \omega)$中的最优投递路线当且仅当对应的欧拉图G^*应满足：

(1) G的每条边在G^*中至多重复出现一次；

(2) G的每个圈在G^*中重复出现的边的权之和不超过该圈权的一半。

- 证明：必要性
- （1）设C是最优投递路线，即C是G^*中的欧拉回路，满足$\sum w(e)$最小，$F=E(G^*)-E(G)$。设G中边e在G^*中重复度为$m(e)$，即在G中的e的两个端点u, v之间添加了$m(e)-1$条重复边，若$m(e)\geq 3$，在G^*中u, v之间的边中随便删除两条，不改变u, v度数的奇偶性，因而所得图G^{**}仍为欧拉图，由于G中各边的权均为正，因而$W(F_2)<W(F_1)$，其中$F_2=E(G^{**})-E(G)$，而$F_1=E(G^*)-E(G)$，这与C是最优投递路线矛盾。

- 定理2：
- 设带权无向连通图$G=(V, E, W)$，V'为G中奇度顶点集，设$|V'|=2k\ (k\geq 0)$，$F=\{\ e \mid e\in E \wedge$ 在求G的最优回路时加了重复边$\}$，则F的导出子图$G[F]$可以表示为以V'中顶点为起点与终点的k条不交的最短路径之并。

6）算法步骤： **（J. Edmonds和E. L. Johnson70年代, O(n^4)）**

- 1）若G中无奇顶点，令$G^*=G$，转2，否则转3。
- 2）求G^*中的欧拉回路，结束。
- 3）求G中所有奇度顶点对之间的最短路径。
- 4）以G中奇度顶点集V'为顶点集，$\forall v_i, v_j \in V'$，边(v_i, v_j)的权为v_i, v_j之间最短路径的权，得完全带权图K_{2k} $(2k=|V'|)$。
- 5）求中K_{2k}最小权完美匹配M。
- 6）将M中边对应的各最短路径中的边均在G中加重复边，得欧拉图G^*，转2。

2 4×4的黑白格棋盘跳马

- 在4×4的黑白格棋盘（四分之一国际象棋盘）上跳马，使得它经过每个格一次并且仅一次，最后又回到出发点。能否办到？为什么？

- ***1）构造数学模型：***
- 将棋盘转化为无向图，作无向图$G=(V, E)$。16个格中各放一个顶点顶点集V。马跳“日”字，若马能在v_i和v_j之间走一步，则v_i和v_j相邻，于是就组成了边集。

- *2）算法思想：*
- G中是否存在哈密顿回路。

- 定理（***必要条件***）

若图G是哈密顿图，则对于顶点集V的每一个非空真子集S，均成立

$$\omega(G-S)\le |S|$$

其中$|S|$表示S中的顶点数，$G-S$表示G中删去顶点子集S后得到的图。

- ***3）解:***
- 删除中心四个点，得6个连通分支，由上述定理，图中不存在哈密顿回路，所以无解。

- 在8×8的黑白格棋盘（国际象棋盘）上跳马，使得它经过每个格一次并且仅一次，能最后又回到出发点。

3 过河问题

- 一个人带了一只狼、一只羊和一棵白菜过河。河边有一条小船，每次除了人以外，只能带一样东西。如果人不在，狼要吃羊，羊要吃白菜。

1）构造数学模型

- 变量表示：M代表人，W代表狼，S代表羊，V代表白菜，$\varnothing$为空集。
- 左岸初始情况：$MWSV$
- 左岸可能出现的16情况：

  $MWSV, MWS, MWV, MSV, WSV, MW, MS, MV, WS, WV, SV, M, W, S, V, \varnothing$
- 其中不能出现的6情况：

  MWV, MW, MV, WS, SV，$M,$

构造图$G(V, E)$，V中顶点表示剩下的10中情况，如果经过一次渡河，情况甲变成情况乙，在情况甲和情况乙表示的顶点间加一条边，作为E中的边。

- ***2）算法分析***
- 在图$G(V, E)$中找一条从$MWSV$到$\varnothing$的最短路。

- **3）解：7次摆渡**
- $MWSV \Rightarrow WV \Rightarrow MWV \Rightarrow V \Rightarrow W \Rightarrow S \Rightarrow MS \Rightarrow \varnothing$
- $MWSV \Rightarrow WV \Rightarrow MWV \Rightarrow W \Rightarrow MSW \Rightarrow S \Rightarrow MS \Rightarrow \varnothing$

二、图论问题解析

- 1　项链——欧拉图
- 2　旅馆经营——网络流
- 3　图的着色
- 3.1　考试安排——顶点着色
- 3.2　时间表——边着色
- 4　舞会——二分图匹配
- 5　正方体着色——图的基本概念
- 6　关键路径

练习1 *项链*

- 有一条用彩色珠子做的漂亮项链。每个珠子由两种颜色组成，相继的两个珠子在邻接处共享一种颜色：

--(green | red) (red | white) (white | green) (green | blue)--

有一天，项链线断了，珠子撒了一地。收集了地上的珠子，但无法肯定是否收齐。想知道用目前收集的珠子是否能够联成项链。

- 数据实例1
- 输入

5

1 2

2 3

3 4

4 5

5 6

- 输出

Some beads may be lost

- 数据实例2
- 输入 输出

输入		输出	
5		2	1
2	1	1	3
2	2	3	4
3	4	4	2
3	1	2	2
2	4		

- 建立数学模型：图的模型
- 关键问题：什么为顶点？什么为边？

- 1)　　珠子为顶点，珠子串接的可能为边
- ***<u>无法将问题的内容表达完全</u>***
- 如实例2, 在图中可以遍历(1, 2) – (2, 3) – (2, 2), 得到的项链不符合题意.

- 2) 作$G=(V, E)$，每种颜色对应一个顶点，$V=\{v_1, v_2, \ldots\ldots, v_m\}$，每颗珠子对应一条边，若有珠子$(c_1, c_2)$, 就有无向边$v_{c1}, v_{c2}$。
- ⇒找欧拉回路

练习2　*旅馆经营*

问题描述：

Praha是一座美丽的旅游城市，每年假期都有许多游客来Praha度假，Hotel U Veze坐落在Praha郊区，有R间客房。

假期将要来临，宾馆收到许多旅游公司的客房订单，但毕竟宾馆客房数有限，不能容纳所有的游客。

现在假设你是Hotel U Veze的经理，请你有选择地接受一部分订单，使Hotel U Veze可以从这些订单中得到最大的收益。每份订单给出它的起止时间和预定房间数。一旦接受某份订单，在订单规定的时间内，就必须给旅客提供房间。

- 输入输出格式
- 输入：

第一行为假期的总天数n和客房总数R，第2行到第n行是一个邻接矩阵的上三角（不含对角线），c_{ij}表示第i天开房，第j天退房的订单预定的房间总数。

- 输出：

输出总收益天数m；接下来的$m-1$行中有一个邻接矩阵的上三角（不含对角线），每个元素f_{ij}表示对预定日期为第i天开房，第j天退房的订单接受的房间总数。

输入输出样例

输入	输出
12 2	12
0 0 0 0 1 0 0 0 0 0 0	0 0 0 0 1 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0
0 1 0 0 0 0 0	0 0 0 0 0 0 0
0 0 2 0 0 0	0 0 1 0 0 0
0 0 0 0 0	0 0 0 0 0
0 0 0 1	0 0 0 1
0 0	0 0
0	0

- *建立数学模型：网络流解题*
- 构造一个能准确描述问题原型、又便于用标准算法求解的网络模型。用网络流解题。

数学模型

构造图$G(V, E)$。第i天对应顶点i。顶点i到顶点$i+1$存在有向边$e(i=1, 2, \ldots\ldots, n-1)$，$e$的流量上限$c(e)=+\infty$，费用$\omega(e)=0$。若第$i$天开房，第$j$天退房的订单预定的房间总数为$c_{ij}$，且$c_{ij}>0$，则顶点$i$到顶点$j$存在有向边$e$，$e$的流量上限$c(e)=c_{ij}$，费用$\omega(e)=-(j-i)$，即天数的相反数。

算法分析

一个流量代表一份订单，一个流的费用为相应订单的收益天数的相反数，f个流的费用为相同起讫日期的订单之总收益天数的相反数。另一方面，房间数为R，故总流量不得超过R。所以要获得最大收益，就要求G的流量为R的最小费用流。

最小费用最大流方法

- ***练习3　图的着色***
- 3.1　考试安排——顶点着色
- 3.2　时间表——边着色

- ***3.1 考试安排***
- 设学校共有n门课程要进行期终考试, 因为不少同学不止选修一门课, 所以不能把一个同学选修的两门课安排在同一场次进行考试. 问学期的期终考试最少需要多少场次才能完成?

- *构造数学模型: 图的顶点着色*
- 　　设$G=(V, E)$, 以每门课程为一个顶点，当且仅当两门课被同一个学生选修时，在相应两个顶点之间连一条边。得图G。
- 　　将n门功课划分成K个划分$U_1, U_2, \ldots, U_k$, 每个$U_i(1\leq i\leq k)$中顶点两两不相邻，要求划分数K必须最少。

- *<u>算法分析:</u>*
- 对每一个子集内的顶点上同一颜色，同色顶点对应的课程可以同时考试。
- 同色顶点集是一个独立集

- *算法步骤：*

（1）求图的所有极大独立集

（2）求出一切若干极大独立集的和含所有顶点的方案

（3）从中挑选所用极大独立集个数最小者

贪心

- *3.2　时间表*
- 　　学校有m位教师，$X_1, X_2, \ldots\ldots, X_m$和$n$个班级$Y_1, Y_2, \ldots\ldots, Y_n$, X_i老师为Y_j班每天上课，排一个课程表，使所排课时数目尽可能地少。

- 构造数学模型：
- 令顶点集$X=\{X_1, X_2, \ldots\ldots, X_m\}$，$Y=\{Y_1, Y_2, \ldots\ldots, Y_n\}$，顶点$X_i$和$Y_j$之间有边$P_{ij}$相连当且仅当$X_i$老师为$Y_j$班上一个课时的课。
- $\Rightarrow$ 形成一个二分图
- 每一学时，教师最多为一个班上课，每个班至多接受一个教师授课。
- $\Rightarrow$ 二分图用尽可能少的颜色对边着色，相邻边颜色不同

- 理论基础
- 定理　　如果G是二分图，则$x'(G)=\Delta(G)$。（边色数=顶点最大度数）

- *练习4 舞会*
- 一次舞会，共有n位男士和n位女士参加，已知每位男士至少认识两位女士，而每位女士至多认识两位男士。问能否将男士和女士分配为n对，使每对中男士和女士彼此认识？

- *1）构造数学模型：二分图*
- n位男士和n位女士分别用顶点表示；
- 若某位男士与某位女士认识，在代表他们的顶点之间连一条线。

- 定理：

设$G=(V_1, V_2, E)$是二分图，如果

（1）V_1中每个顶点至少关联t条边；

（2）V_2中每个顶点至多关联t条边；

则G中存在V_1到V_2的匹配，其中t是正整数。

- 所以，可分配。

- *练习5　正方体着色*
- 　　有4个立方体，用红、黄、蓝、绿4种颜色将它们面着色，使得4种颜色在每个立方体的6个面中都至少出现1次。着色方案如下图.证明：可将这4个立方体堆成四棱柱，使它每个侧面都呈现4种颜色。

R	R	G	B
RYGB	RYBG	BBRY	GYRG
R	Y	G	Y
(1)	(2)	(3)	(4)

- 1）建立数学模型：

　　用图表示正方体的着色方式：

　　R(红), Y(黄), B(蓝), G(绿)表示4个顶点；

　　顶点相邻当且仅当正方体有两个相对的面分别桌的是这两种颜色；令G_i表示正方体i的着色方式，$i=1, 2, 3, 4$；$G=G_1 \cup G_2 \cup G_3 \cup G_4$。

- 2）算法分析
- 存在可行方案的充要条件是G中存在两个生成子图，均含4条边，且标码为1，2，3，4的边各只出现一次，每个顶点关联两条边。

先按G调整好棱柱的前后两个侧面，使每个侧面都具有不同的颜色，如下图1所示；再按G调整左右两个侧面，前后侧面颜色不变，使左右两个侧面的颜色如下图2所示。

前	后	左	右
G	R	G	Y
R	B	Y	B
Y	G	B	R
B	Y	R	G

6 关键路径

- **1）PERT**（Program Evaluation and Review Technique）. 计划评审技术, 1956年。

- **2）PERT图**

设D是n个结点的有向简单带权图，若满足：

（1）D中无回路；

（2）D中有一个顶点的入度为0，记为v_1，称v_1为***发点***，有一个顶点的出度为0，记为v_n，称v_n为***收点***；

（3）任意的$v \in V-\{v_1, v_n\}$，则v在某条从v_1到v_n的路径上，则称D是**PERT图**。

- 在PERT图中，每条边表示一道工序或一种活动，若有向边$(v_i, v_j), (v_j, v_k)$相邻，则表示工序(v_j, v_k)必须在(v_i, v_j)结束后才能开始。
- $\forall v_i \in V$, v_i表示一种状态，它表示关联到它的工序都结束后，关联于它的工序才能开始。
- v_1表示工程开始
- v_n表示工程结束。

- ***3）关键状态，关键工序/关键活动***

  在**PERT**图中的*关键路径*是从发点v_1到收点v_n的最长（按权计算）的路径。处于关键路径上的顶点称为*关键状态*，处在关键路径上的边称为*关键工序*或*关键活动*。

- *要使工期缩短，缩小关键路径上边的权。*

4）最早完成时间/最晚完成时间

最早完成时间

设**PERT**图为有向图D，任意的$v_i \in V(D)$，称从发点v_1沿最长的路到达v_i所需要的时间为v_i的最早完成时间，记作$TE(v_i)$。

$TE(v_i)$是以v_i为起点的各工序的最早可能开工时间，因而称为v_i的最早完成时间，它是v_1到达v_i的最长路径的权。

$$\begin{cases} TE(v_1) = 0 \\ TE(v_i) = \max\limits_{v_j \in \Gamma^-(v_j)} \{TE(v_j) + \omega_{ji}\}, i \neq 1 \end{cases}$$

其中，$\Gamma^-(v_j)$为v_i的先驱元集，ω_{ji}为边(v_j, v_i)的权。

最晚完成时间

在保证收点v_n的最早完成时间$TE(v_n)$不增加的条件下，从v_1最迟到达v_i所需要的时间，称为$\mathbf{v_i}$的最晚完成时间，记作$TL(v_i)$。

$TL(v_i)$为$TE(v_n)$与v_i沿最长（按权计算）路径到达v_n所需时间之差，$TL(v_i)$是关联于v_i的各道工序所允许的最迟开工时间。

$$TL(v_n) = TE(v_n)$$

$$TL(v_i) = \min_{v_j \in \Gamma^+(v_i)} \{TL(v_j) - \omega_{ij}\}, i \neq n.$$

其中，$\Gamma^+(v_i)$为v_i的后继元素，ω_{ij}为边(v_i, v_j)的权。

- 5）缓冲时间/松弛时间

$TL(v_i)$-$TE(v_i)$为v_i的缓冲时间或松弛时间，记为$TS(v_i)$。

$$TS(v_i)\geq 0，i=1, 2, \ldots, n.$$

- 定理：

  $TS(v_i)=0$当且仅当v_i处在关键路径上。

三、图论难题思考

- 1 *图的基本概念（遍历）*

  拯救大兵瑞恩

- 2 *网络流标号法*

  航天实验

  家园

1　拯救大兵瑞恩

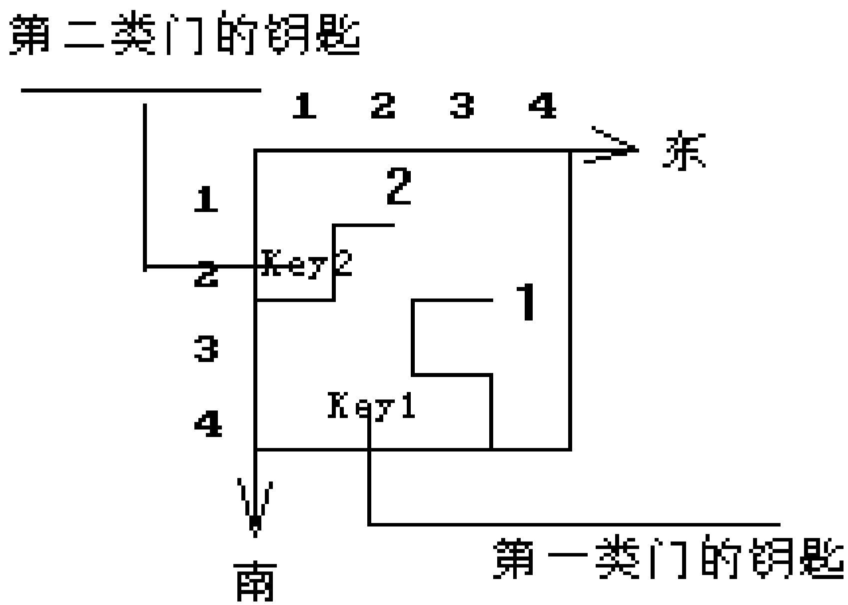
- 问题背景

　　1944年，特种兵麦克接到国防部的命令，要求立即赶赴太平洋上的一个孤岛，营救被敌军俘虏的大兵瑞恩。瑞恩被关押在一个迷宫里，迷宫地形复杂，但是幸好麦克得到了迷宫的地形图。

迷宫的外形是一个长方形，其在南北方向被划分为N行，在东西方向被划分为M列，于是整个迷宫被划分为N*M个单元。我们用一个有序数对（单元的行号，单元的列号）来表示单元位置。南北或东西方向相邻的两个单元之间可以互通，或者存在一扇锁着的门，又或者存在一堵不可逾越的墙。迷宫中有一些单元存放着钥匙，并且所有的门被分为P类，打开同一类的门的钥匙相同，打开不同类的门的钥匙不同。

大兵瑞恩被关押在迷宫的东南角，即（N,M）单元里，并已经昏迷。迷宫只有一个入口，在西北角，也就是说，麦克可以直接进入(1,1)单元。另外，麦克从一个单元移动到另一个相邻单元的时间为1，拿取所在单元的钥匙的时间以及用钥匙开门的时间忽略不计。

你的任务是帮助麦克以最快的方式抵达瑞恩所在单元，营救大兵瑞恩。



迷宫示例

输入输出结构

- **输入：**
- 第一行是三个整数，依次表示N,M,P的值；
- 第二行是一个整数K，表示迷宫中门和墙的总个数；
- 第I+2行（1<=I<=K），有5个整数，依次为Xi1,Yi1,Xi2,Yi2,Gi：
- 当Gi>=1时，表示(Xi1,Yi1)单元与(Xi2,Yi2)单元之间有一扇第Gi类的门，当Gi=0时，表示(Xi1,Yi1)单元与(Xi2,Yi2)单元之间有一堵不可逾越的墙；
- （其中，|Xi1-Xi2|+|Yi1-Yi2|=1，0<=Gi<=P）
- 第K+3行是一个整数S，表示迷宫中存放的钥匙总数；
- 第K+3+J行(1<=J<=S)，有3个整数，依次为Xi1,Yi1,Qi：表示第J把钥匙存放在(Xi1,Yi1)单元里，并且第J把钥匙是用来开启第Qi类门的。（其中1<=Qi<=P）
- 注意：输入数据中同一行各相邻整数之间用一个空格分隔。
- **输出：**
- 输出文件只包含一个整数T，表示麦克营救到大兵瑞恩的最短时间的值，若不存在可行的营救方案则输出-1。

- *网络流标号法*
- 1　航天实验
- 2　家园

1 航天实验

NASA正在为航天飞行计划一系列的太空飞行，在每次飞行中必须决定进行何种商业性实验和配载何种仪器设备. S教授正在研究这一问题. 对每次飞行NASA考察一个实验集合$E=\{e_1, \ldots\ldots, e_n\}$, 并且实验$e_j$的商业赞助人已同意为该实验的结果向NASA支付$p_j$美圆. 实验使用的全部仪器为集合$I=\{I_1, \ldots, I_n\}$; 每个实验室所需要用到的仪器为子集$r_j \in I$.运送仪器$I_k$的运费为$c_k$美圆。

S教授的任务就是找一个有效算法以确定在某一次指定的飞行中要进行哪些实验并运送哪些仪器才能使净收益（进行实验所获得的全部收入与运送仪器的全部费用的差额）最大。

- 输入输出结构

- 2 家园

问题背景

由于人类对自然的疯狂破坏，人们意识到在大约2300年之后，地球不能再居住了，于是在月球上建立了新的绿地，以便在需要时移民。令人意想不到的是，2177年冬由于未知的原因，地球环境发生了连锁崩溃，人类必须在最短的时间内迁往月球。

现有n个太空站处于地球与月球之间（编号1..n），m艘公共交通太空船在其中来回穿梭，每个太空站Si可容纳无限的人，每艘太空船pi只可容纳Hpi人。对于每一艘太空船pi，将周期性地停靠一系列的太空站（Si1,Si2…Sir），如：（1，3，4）表示停靠太空站1 3 4 1 3 4 1 3 4 …。任一艘太空船从任一个太空站驶往另一个任意的太空站耗时为1。人只能在太空船停靠太空站（或地球、月球）时上船或下船。初始时的人全在地球上，太空船全在初始站（太空船pi处于Si1），目标是让所有的人尽快地全部转移到月球上。

输入：

文件第一行为三个正整数 n（太空站个数）、 m（太空船个数）、 k（需要运送的地球上的人的个数），其中1<=m<=13, 1<=n<=20, 1<=k<=50。

接下来的n行给出了太空船的信息，第i+1行说明太空船pi，此行第一个数表示pi可容纳的人数Hpi，第二个数表示pi停靠一个周期的太空站个数r，1<=r<=n+2, 随后r个数便是停靠的太空站的编号(Si1,Si2,…,Sir), 地球用0表示，月球用-1表示。0时刻时，所有太空船都在初始站，随后开始运行，在时刻1，2，3…等正点时刻各艘太空船停靠相应的太空站，即人只有在0,1,2…等正点时刻才能上下太空船。

- **输出：**
- 　文件只有一个数，若问题有解，输出完成全部人员安全转移的时刻，否则输出0。