Annales Academiae Paedagogicae Cracoviensis
Folia 16 Studia Mathematica III (2003)

Karel Hrach

Two problems solved with MS-Excel

Abstract. The first problem arising in the theory of graphs is to find the maximum flow in a network. This is equivalent to the problem of finding a minimum cut, according to Ford-Fulkerson algorithm, which is briefly explained. The second, the probabilistic problem solved by Bernoulli, is to calculate the probability that certain sum is obtained when throwing a dice several times. This contribution briefly presents a recurrent solution for a generalized dice with any number of possible values, with a hint how to solve these two problems simply using the commonly known SW MS-Excel.

1. Introduction

Solutions of the two problems discussed in this contribution are known, but without a computer rather inconvenient for the practical use. When using even the common MS-Excel, the solution becomes very fast and simple, as shown below.

2. The first problem – the maximum flow

The first problem is known in the theory of graphs as the maximum flow in the network. A network is an oriented and valued graph with n middle-knots $k_1, \dots, k_n$ and with two special knots — the entry k_0 and the exit k_{n+1}. No arrow (oriented edge) may lead back to the entry and no arrow may lead from the exit. Its positive capacity belongs to each arrow $(k_i; k_j)$. For non-existing arrows their capacity may be defined as a zero. One can imagine this structure as a system of pipelines. There is an example of the network with $n = 4$ middle-knots I, II, III and IV at figure 1.

A cut is any set K containing entry k_0 and some of the knots $k_1, \dots, k_n$ such that for every knot $k_j \in K \setminus \{k_0\}$ there is a knot $k_i \in K$ such that $(k_i; k_j)$ represents an existing arrow. In the network at figure 1, for example, cut $K = \{k_0, k_2, k_3, k_4\}$ exists, but cut $\{k_0, k_1, k_4\}$ does not exist. The value of every existing cut K is defined as a sum of arrows $(k_i; k_j)$, $k_i \in K, k_j \notin K$, i.e. of arrows leading outside. For the concrete cut K mentioned above, the value equals $75 + 28 + 11 + 5 + 30 = 149$. The question is: What is the maximum flow for such a network? Ford-Fulkerson algorithm gives an answer, see e.g.

[1]. Simply said, every network is as strong, as strong is its weakest cut. More precisely, one has to find all the existing cuts, calculate their values, and detect the minimum among them. Easy combinatorics shows that in the case of n middle-knots the maximum number of existing cuts equals 2^n. In our example with $n = 4$ we have to check 16 combinations of knots. How to find out, which cuts really exist?

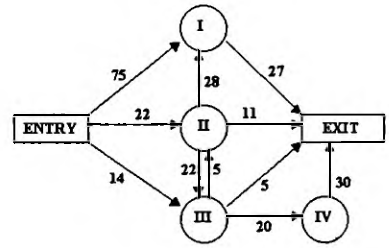

Figure 1. Network — an example

Solution is in the matrix representation of the network. We shall define square matrix (a_{ij}), a_{ij} equals the capacity of $(k_{i-1}; k_j)$, $i, j = 1, \ldots, n+1$. Every cut K_{m+1} consisting of $m+1$ knots might be recurrently defined using all the existing cuts K_m containing m knots. Figure 2 illustrates this approach. For example, the cut $\{k_0, k_2, k_3\}$ exists and the capacity leading from k_3 to k_4 is positive (the logical condition in the cell K3 at figure 2 resulted in 1). That is why cut $K = \{k_0, k_2, k_3, k_4\}$ exists. Its value equals 149, it is the sum of the cells from all the rows of (a_{ij}) except the second one (labeled as I), but excluding the columns labeled II, III and IV. It could be found similarly that among 16 possible cuts in our example eleven do exist. Among them the minimum value is 63 and this is the maximum flow for the network of figure 1.

H2 = =IF(K10>0;SUMA(C2;C4:C6;G2;G4:G6);"NOT")

A	B	C	D	E	F	G	H	I	J	K
	to:	I	II	III	IV	exit	SUM	S,II,III+IV	from S?	NOT
from:	entry	75	22	14	0	0	149		from II?	NOT
	I	0	0	0	0	27			from III?	1
	II	28	0	22	0	11		S,II,IV+III	from S?	NOT
	III	0	5	0	20	5			from II?	NOT
	IV	0	0	0	0	30			from IV?	NOT
								S,III,IV+II	from S?	1
									from III?	1
									from IV?	NOT
								(S...entry)		3

Figure 2. One of the cuts represented in MS-Excel

3. The second problem – summing the thrown dice

The second, the probabilistic problem solved by Bernoulli, is to calculate the probability that a certain sum is obtained when throwing a dice several times. Suppose now we have a generalized "dice" with A possible values, $A \geq 2$. Let the rectangular (uniform) variables X_j are results of the N throws, $N \geq 2$. Thus $P(X_j = a) = 1/A$ for any $a = 1, \ldots, A, j = 1, \ldots, N$. Let variable $Y^{(N)}$ means the sum $X_1 + \ldots + X_N$. There exist $AN-N+1$ values $y_i^{(N)}$ of $Y^{(N)}$, $y_1^{(N)} = N, y_2^{(N)} = N+1, \ldots, y_{AN-N+1}^{(N)} = AN$, namely. As summarized in [3], for the number $n_i^{(N)}$ of adequate results (i.e. of the distinct series $X_1, \ldots, X_N$ resulting in the sum $y_i^{(N)}$)the following recurrent formulas hold:

$$n_i^{(1)} = n_1^{(N)} = 1, \quad i = 1, \ldots, A, \quad N \geq 2,$$
$$n_i^{(N)} = n_{i-1}^{(N)} + n_i^{(N-1)}, \quad i = 2, \ldots, A, \quad N \geq 2,$$
$$n_i^{(N)} = n_{i-1}^{(N)} + n_i^{(N-1)} - n_{i-A}^{(N-1)}, \quad i = A+1, \ldots, (A-1)(N-1)+1, \quad N \geq 3,$$
$$n_i^{(N)} = n_{i-1}^{(N)} - n_{i-A}^{(N-1)}, \quad i = (A-1)(N-1)+2, \ldots, AN-N+1, \quad N \geq 2.$$

Simply said, $n_i^{(N)}$ is the moving sum of the length A (for the "middle" values of $n_i^{(N)}$) or less (otherwise). The calculation of $n_i^{(N)}$ is thus straightforward, as illustrates table 1. Their structure might be called generalized Pascal triangle (ordinary Pascal triangle corresponds with the case $A = 2$).

i	$y_i^{(1)}$	$n_i^{(1)}$	$y_i^{(2)}$	$n_i^{(2)}$	$y_i^{(3)}$	$n_i^{(3)}$
$\Sigma = A^N$	x	6	x	36	x	216
1	1	1	2	1	3	1
2	2	1	3	2	4	3
3	3	1	4	3	5	6
4	4	1	5	4	6	10
5	5	1	6	5	7	15
6	6	1	7	6	8	21
7	x	x	8	5	9	25
8	x	x	9	4	10	27
9	x	x	10	3	11	27
10	x	x	11	2	12	25
11	x	x	12	1	13	21
12	x	x	x	x	14	15
13	x	x	x	x	15	10
14	x	x	x	x	16	6
15	x	x	x	x	17	3
16	x	x	x	x	18	1

Table 1. Recurrent calculation, $A = 6$, $N = 1, 2, 3$

For example,

$n_4^{(3)} = n_1^{(2)} + n_2^{(2)} + n_3^{(2)} + n_4^{(2)} = 1 + 2 + 3 + 4 = 10,$

$n_8^{(3)} = n_3^{(2)} + n_4^{(2)} + n_5^{(2)} + n_6^{(2)} + n_7^{(2)} + n_8^{(2)} = 3 + 4 + 5 + 6 + 5 + 4 = 27,$

$n_{14}^{(3)} = n_9^{(2)} + n_{10}^{(2)} + n_{11}^{(2)} = 3 + 2 + 1 = 6.$

Thus the probabilities that the sum of the results equals 6 (= $y_4^{(3)}$), 10 (= $y_8^{(3)}$) and 16 (= $y_{14}^{(3)}$), respectively when throwing a dice three-times are 10/216, 27/216 and 6/216.

Note, that many similar problems are discussed in [2].

References

[1] *Aplikovaná matematika*, SNTL, Praha, 1978.

[2] J. Anděl, *Matematika náhody*, Matfyzpress, Praha, 2000.

[3] K. Hrach, *Summing discrete rectangular variables*, Acta Economica Pragensia, VŠE, Praha, 2001.

FSE UJEP
Moskevská 54
400 96 Ústi nad Labem
Czech Republic
E-mail: hrach@fse.ujep.cz