IRISH MATHEMATICAL SOCIETY

NEWSLETTER

NUMBER 7 MARCH 1983

PATHS IN A GRAPH

Johannes Siemons

In a connected graph any two vertices can be joined by a sequence of edges. This is the definition of connectedness for graphs. However, how do you find a path joining a given pair of vertices, and how do you decide effectively if a graph is connected? These are the questions I shall discuss in this note. The graphs we consider are finite, undirected and have no loops or multiple edges. A path is a sequence $\{v',v_1\} = e_1$, $\{v_1,v_2\} = e_2$, ..., $\{v_{r-1},v''\} = e_r$ of edges without repetition (of edges: vertices *may* occur repeatedly). The vertices v' and v'' are the end vertices of the path.

A popular version of this problem is to find the exit in a maze. We have to distinguish two cases. In the first instance, imagine that we are actually inside a maze without knowing its overall design. Here the only solution seems to be trial and error. A successful route to the exit is very unlikely to be a path according to our definition. In fact, the probability to reach the exit on a path is less than 2^{-c}, where c is the number of intermediate junctions on a path to the exit (provided that there is only one such path in the maze). In other words, it is almost impossible to avoid walking into a cul-de-sac! However, most commonly, maze puzzles are done with paper and pencil, and the design of the maze is right in front of your eyes. In this situation, can you avoid a cul-de-sac? The answer is yes, there is a *construction* for a path to the exit!

For a set P of edges let V(P) be the set of end vertices of edges in P. For a vertex v in the graph, let $d_P(v)$ be the number of edges in P that end at v. A cycle is a path that ends in its initial vertex. Our construction is based upon the following simple observation:

<u>Lemma</u> Let v' and v" be two vertices in a graph and let P be a set of edges such that $d_P(v')$ and $d_P(v")$ are odd, while $d_P(v)$ is zero or even for all remaining vertices. Then $P = P(v',v") \cup C_1 \cup \ldots \cup C_r$ where P(v',v") is (suitably arranged) a path from v' to v" while each C_i is a cycle that has no vertex in common with P(v',v").

<u>Proof</u> Let G_0, G_1 ... be the connected components of the subgraph with vertices V(P) and edges P. As $d_P(v')$ is at least 1, v' is a vertex in one of the G_i, say in G_0. But then also v" belongs to G_0, for otherwise the total degree sum in G_0 would be odd, which is impossible: In any graph the total degree sum is even. Therefore G_0 is a path from v' to v" and the remaining components are cycles.

How can we effectively determine such a set of edges? And, secondly, how can we ensure that P does not contain cycles? (From a practical point of view, the second problem is less relevant, for if we start our path in v' we will reach v" without entering any of the cycles C_i) We shall say that a set P as in the lemma is *short* if none of its subsets is a cycle. Thus a short path from v' to v" is a path where none of the intermediate vertices is repeated.

We order the vertices of G in some way $v_1, \ldots, v_n$ and also order its edges $e_1, \ldots, e_m$. The graph now can be represented by its incidence matrix I. This is the matrix whose rows are indexed by vertices and whose columns are indexed by edges, such that $(I)_{v,e}$ is 1 if e ends at v and $(I)_{v,e} = 0$ otherwise. A set S of vertices is represented by a 0-1-vector $\underline{S}$ of length n where $(\underline{S})_i = 1$ iff v_i belongs to S. In the same way, an edge set P is represented by a 0-1-vector $\underline{P}$ of length m. The incidence matrix associates a vertex vector to any edge vector: $I \cdot \underline{P}^t$ is a vector of length n and its i th component is easily seen to be $d_P(v_i)$. Now we realise that a set P has the property of the lemma exactly if $\underline{P}$ satisfies a linear congruence modulo 2.

Path Construction: A set P *of edges consists of a path* $P(v',v'')$ *and a number of cycles disjoint from* $P(v',v'')$ *if and only if* $I \cdot \underline{P}^t \equiv \underline{S}$ modulo 2 *where* $S = \{v',v''\}$.

Thus a path from v' to v" can be constructed by solving this linear congruence, for instance by Gauss elimination. This is particularly simple in characteristic 2 where we only need to add rows and possibly permute rows and columns of I. Note also that cycles and unions of cycles correspond to 0-1-vectors in the kernel of I modulo 2. In order that the graph is connected, this congruence has to be solvable for any choice of S. This will be the case if and only if the rank of I is at least n-1 in characteristic 2. However, as each column of I adds up to 2, the rank will be n-1 exactly. Therefore, we obtain a criterion for connectedness in a graph.

The number of connected components in a graph is the number of vertices minus the rank of I *in characteristic* 2.

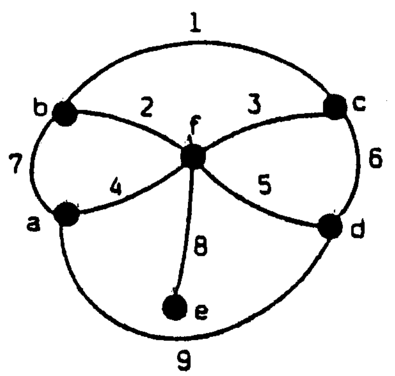
Short paths: Now we shall see that $I \cdot \underline{P}^t \equiv \underline{S}$ can be solved in such a way that a solution automatically will be short, that is, P does not contain a cycle. Using Gauss elimination, the congruence can be transformed into

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & . & . & . & 0 & * & * & * & . & . & . & * \\ 0 & 1 & 0 & 0 & 0 & . & . & . & 0 & * & * & * & . & . & . & * \\ 0 & 0 & 1 & 0 & 0 & . & . & . & 0 & * & * & * & . & . & . & * \\ . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & . & . & . & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{n-1} \\ p_n \\ \vdots \\ p_m \end{pmatrix} \equiv \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ \vdots \\ 0 \end{pmatrix} \bmod 2$$

We now choose $p_n = p_{n+1} = \ldots. = p_m = 0$ and hence have $p_i = s_i$ for $i = 1, \ldots, n-1$. If P is determined in this way, none of its subsets can satisfy the homogeneous congruence and therefore P does not involve any cycle. Thus P is a short path from v' to v". Of course, the above tableau can usually be achieved in a number of distinct ways. This corresponds to the fact that a short path is unique only if the graph contains no cycle.

<u>Maximal and Minimal Short Paths</u>: In the above tableau, the entries s_i are calculated from $\underline{S} = \underline{\{v',v''\}}$ during the subsequent row operations. The number of $s_i \neq 0$ is, as we have seen above, the length $\ell(P)$ of the short path from v' to v". Therefore, the minimum value obtainable for $\ell(P)$ in any tableau is the distance from v' to v". As a short path passes through any vertex at most once, $\ell(P) + 1$ is the number of vertices en route, v' and v" included. Thus $\ell(P)$ is at most n-1, but this may or may not be obtainable in a tableau. For, if $\ell(P) = n-1$, then P passes through all the vertices of the graph. Such a path is called *hamiltonian*. No satisfactory criteria for the existence of such paths exist for graphs in general. In a particular case, however, we notice that hamiltonian paths correspond to tableaux of the above form in which $s_1 = s_2 = \dots = s_{n-1} = 1$.

As an example consider the graph in Fig. 1. It has 6 vertices a,b, ..., f, 9 edges 1,2, ..., 9 and its incidence matrix is the 6x9 matrix I given below. We form the 6x15 matrix (I,Id) where Id is the 6x6 identity matrix.



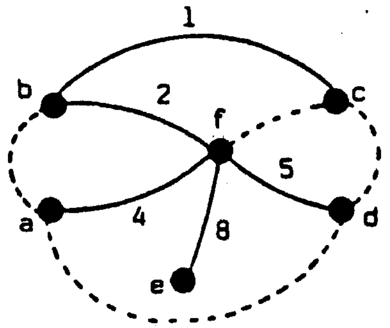
$$
\begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\
\begin{bmatrix}
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0
\end{bmatrix}
\end{array} = I
$$

<u>FIGURE 1</u>

On this matrix subsequently Gauss elimination is performed (allowing permutations of the columns of I) and we obtain

$$
(I^{+},S) = \left[\begin{array}{ccccc c cccc|cccccc}
1 & 2 & 4 & 5 & 8 & & 3 & 6 & 7 & 9 & & & & & & \\
1 & 0 & 0 & 0 & 0 & & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right]
$$

In this process at least 1 column permutation has to take place, e.g. edge 8 has to be included among the first 5 edges. In the event these are the edges 1,2,4,5,8. This means that they are the only edges effectively used in the construction. As no cycle can be formed from them, they automatically build a *spanning tree*.



$P(a,e) = \{4,8\}$

$P(a,c) = \{4,2,1\}$

$P(a,d) = \{4,5\}$

FIGURE 2

The matrix S in a certain way is a generalised inverse of I. For a given pair x,y of vertices the path $P = P(x,y)$ with edges among 1,2,4,5,8 is unique and can immediately be read off from $I^{+}.\underline{P} = S\cdot\{x,y\}$.

Department of Mathematics,
University College Dublin,
Belfield,
Dublin 4.