THE TWO PATHS PROBLEM IS POLYNOMIAL

by

Yossi Shiloach

STAN-CS-78-654
APRIL 1978

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

The Two Paths Problem Is Polynomial

Yossi Shiloach */

Computer Science Department
Stanford University
Stanford, California 94305

September, 1977

Abstract.

Given an undirected graph $G = (V,E)$ and vertices $s_1,t_1;s_2,t_2$, the problem is to determine whether or not G admits two vertex disjoint paths P_1 and P_2, connecting s_1 with t_1 and s_2 with t_2 respectively. This problem is solved by an $O(n\cdot m)$ algorithm $(n = |V|, m = |E|)$, An important by-product of the paper is a theorem that states that if G is 4-connected and non-planar, then such paths P_1 and P_2 exist for any choice of s_1, s_2, t_1, and t_2, (as was conjectured by Watkins in [W]).

Keywords and Phrases: Algorithm, Connectivity, Disjoint paths, Planarity, Two Paths Problem.

*/ This research was supported by a Chaim Weizmann Postdoctoral Fellowship and by National Science Foundation grant MCS75-22870.

1. Basics,

1. A graph in this paper is undirected, without multiple edges or self loops (which are irrelevant to the problem).

2. It is assumed that basic graph theory concepts, such as paths, k-connectivity, planar/complete/bipartite graphs, etc., are familiar to the reader.

3. Disjoint paths means vertex-disjoint paths (excluding their end-points), and k-connectivity means vertex k-connectivity,

4. G has the P2 property if for any $s_1,t_1;s_2,t_2$ there exist two disjoint paths connectiving s_1 with t_1 and s_2 with t_2 ,

A comprehensive treatment of the combinatorial part of the problem (i.e., what conditions imply the P2 prperty) and also more general problems was done by M. Watkins in [W]. Algorithmic partial results were recently obtained by A. Itai [I1] and by Y. Perl and the author [PS]. Another closely related work is that of A. S. LaPaugh, [L].

2. Reductions of the Problem.

R1: We may assume that G is 3-connected.

This reduction relies on a detailed analysis of the problem concerning graphs which are not 3-connected, which was done in [I1]. Itai shows that the problem can be solved in $O(n+m+T)$ time, where T is the time required to solve the problem for a 3-connected graph $G' = (V',E')$ such that $|V'| \leq n$ and $|E'| \leq m$. A brief outline of this work is given in the appendix.

R2: We may assume that G is not planar.

This reduction is a result of the work which was done in [PS]. This work solves the problem for 3-connected planar graphs in $O(n+m)$ time.

By Kuratowski's theorem, G contains a homeomorph to either K_5 (the complete graph on 5 vertices) or to $K_{3,3}$ (the complete bipartite graph with 3 vertices on each side).

R3: We may assume that there are four disjoint paths connecting s_1, t_1, s_2 and t_2 to any other set of four vertices or less,

Proof. Let $S = \{s_1,t_1,s_2,t_2\}$ and let S' be a set of vertices such that there are no four disjoint paths connecting the vertices of S and S', Then, by Menger's theorem, S can be separated from S' by a cut-set C of three vertices. $S \cap C$ and $S' \cap C$ are not necessarily empty, but $G' = G-C$ contains at least one connected component $G_1 = (V_1,E_1)$ such that $V_1 \cap S = \emptyset$ and $V_1 \cap S' \neq \emptyset$. Let $C = \{v_1,v_2,v_3\}$ and let

$\bar{G} = (V-V_1, E-E_1 \cup \{(v_1,v_2),(v_1,v_3),(v_2,v_3)\})$. The following lemma is very easy to prove.

Lemma 2.1. The TPP with $G;s_1,t_1,s_2,t_2$ is equivalent to that with $\bar{G};s_1,t_1,s_2$ and t_2. Moreover, a solution to the first can be easily obtained from a solution to the second.

Since $|V-V_1| < n$, Lemma 2.1 implies that we can reduce the size of the problem by using only a polynomial time computation (required to determine C). If such a reduction is not possible then R3 holds,

Q.E.D.

R4: G does not contain a homeomorph to K_5 .

This reduction is due to Watkins' work. Watkins shows that if G contains a homeomorph to K_5 and R3 holds, then G has the P2 property. Moreover, his proof is completely constructive and can be implemented, step by step, in polynomial time. The exact complexity of his proof will be evaluated in Section 5 where the complexity of the whole algorithm will be determined.

3. Further Reductions.

So far G is a 3-connected graph, containing a subgraph $G_{3,3}$ homeomorphic to $K_{3,3}$. The nine paths of $G_{3,3}$ which consist of the edges of $K_{3,3}$, will be called p-edges (a short form of pseudo edges) and the six vertices of $G_{3,3}$ which represent the vertices of $K_{3,3}$, will be called p-vertices. In the figures to come, three p-vertices will always be drawn as circles while the other three as squares, indicating the two "sides" of $K_{3,3}$. Other vertices of $G_{3,3}$ will be drawn as • (see Figure 3.1). The circled p-vertices will be x_1, x_2 and x_3. The squared ones -- y_1, y_2 and y_3.

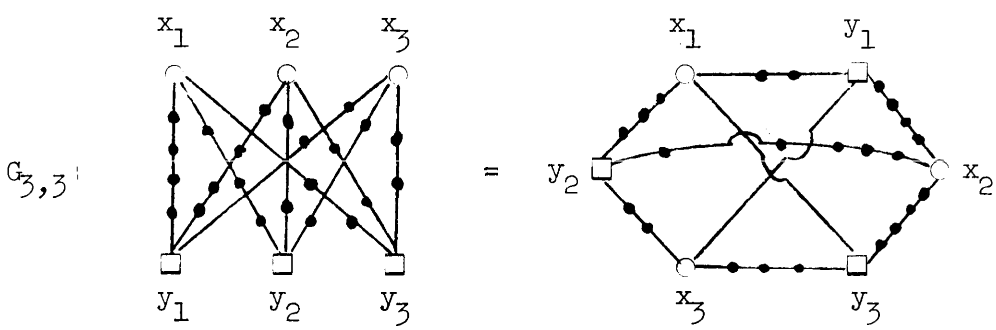


Figure 3.1.

R5: We may assume that s_1 is a p-vertex,

Proof. If s_1 is not a p-vertex, we are going to modify $G_{3,3}$ and make it a p-vertex. We construct three disjoint paths P_1, P_2 and P_3 from s_1 to x_1, x_2 and x_3 respectively. We are interested only in that part of each path from s_1 to the vertex in which it hits $G_{3,3}$ for the first time. These parts of the paths will be denoted by

P'_1 , P'_2 , and P'_3 and the vertices in which they hit $G_{3,3}$ for the first time -- will be f_1 , f_2 , and f_3 respectively, All possible not-symmetric cases are given in Figure 3.2. The cases differ from each other by the location of the f_i's on $G_{3,3}$. The new $G_{3,3}$ in each case is heavily lined.

<u>Case a</u>. The f_i's belong to three different p-edges.

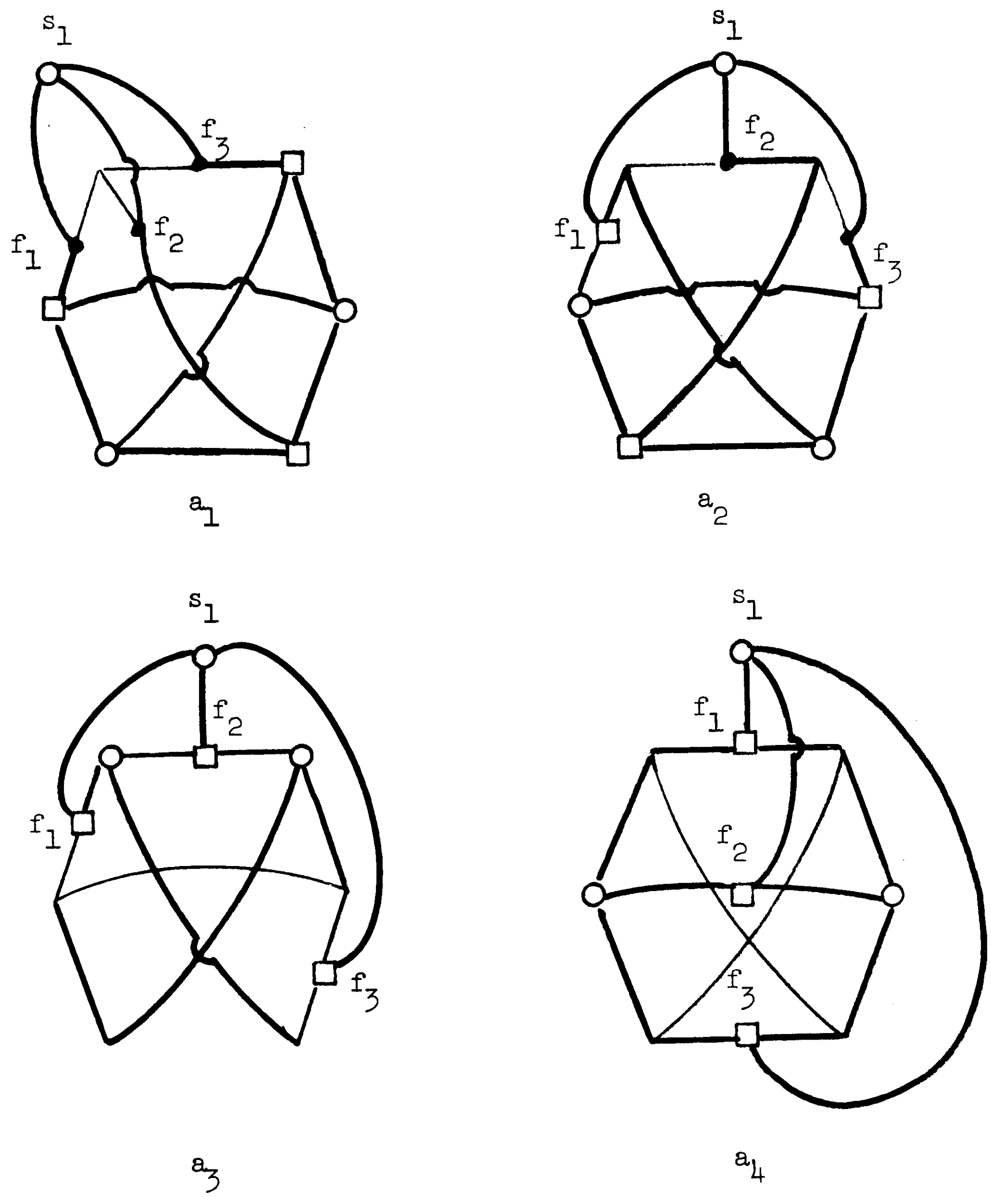
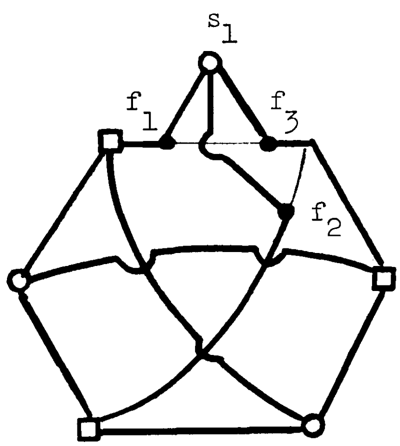


Figure 3.2-a.

Case b. Two f_i's are on the same p-edge.



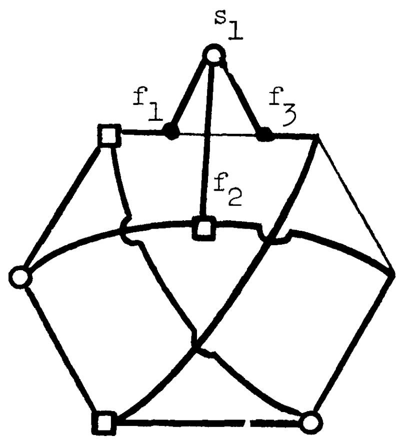
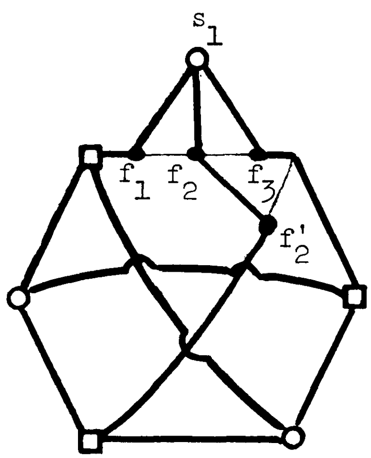
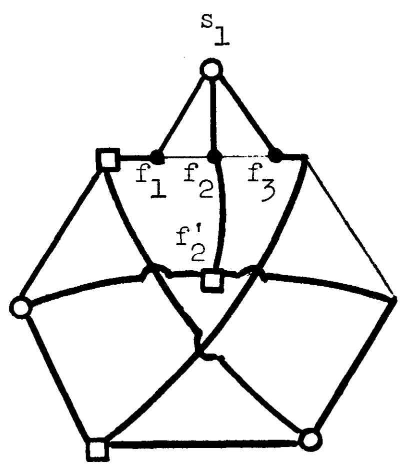


Figure 3.2-b.

Case c. All the f_i's are on the same p-edge. In this case, the path in the middle, say P_2, is continued to its second intersection with $G_{3,3}$ at f_2'.



c1



c2

Figure 3.2-c.

Q.E.D.

Two important remarks:

(1) In Figure 2.3 and in those to come, the f_i's are drawn as vertices of $G_{3,3}$ which are <u>not</u> p-vertices. This is not necessarily true, of course, and the f_i's might be p-vertices as well. However, if one or more of the f_i's are p-vertices, everything is easier. Since it would triple the amount of case analyses involved, we have omitted this case.

(2) In Case c, we have omitted the possibility that f_2' is on the same p-edge as f_2. If that happens, we first modify $G_{3,3}$ as shown in Figure 3.4 and then compress the subpath of P_2 between f_2 and f_2' into one vertex, say f_2'', (see Figure 3.4). This f_2'' is the new first intersection of the modified P_2 with the modified $G_{3,3}$. It is easy to see that all the properties which are relevant to our discussion (such as disjointness of P_1, P_2 end P_3) are preserved by this transformation. Moreover, f_2'' is closer to the end of P_2 then f_2. This implies that these adjustments take place at most O(n) times. The same modification is applied, if necessary, to P_1 and P_3 as well. The same assumption (that the P_i's do not hit the same p-edge twice without hitting another one in between) was also made by Watkins and we shall refer to it as the <u>W-assumption.</u>

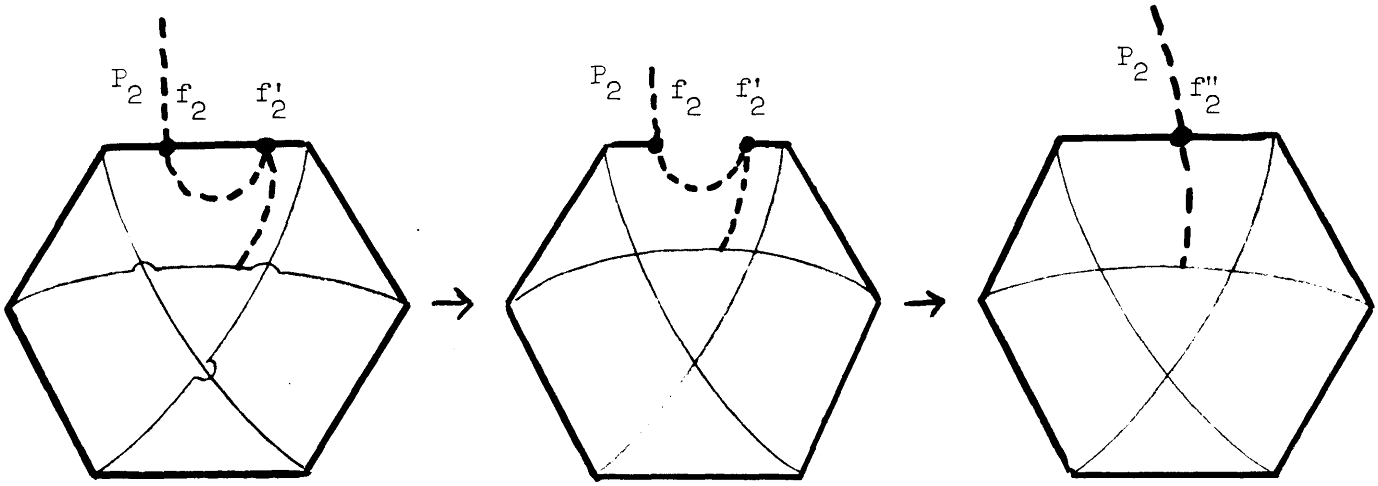


Figure 3.4.

R6: One of the following two cases occurs:

(1) t_1 is also a p-vertex.

(2) G has a subgraph, $G^+_{3,3}$, homeomorphic to that shown in Figure 3.5.

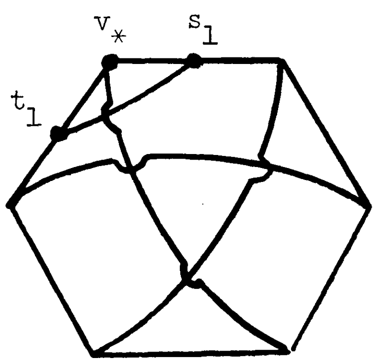


Figure 3.5.

Proof. R5 is now assumed. The three p-edges which are incident with s_1 will be called black p-edges while the others will be white p-edges. The technique of proving R6 is very similar to that of R5. We consider three disjoint paths Q_1, Q_2, and Q_3 connecting t_1 with three distinct p-vertices (no matter which). The first intersection of each Q_i with $G_{3,3}$ will be denoted by g_i, $i = 1,2,3$. We now have more cases to consider since we have black and white p-edges. The main three cases correspond to whether the Q_i's "land" on one, two, or three different p-edges. The subcases take the color of the p-edges into account. Figure 3.6 covers all non-symmetric cases, subject to the remarks made after the proof of R5.

Case 1. The Q_i's land on three different p-edges.

1-a. Three black ones.

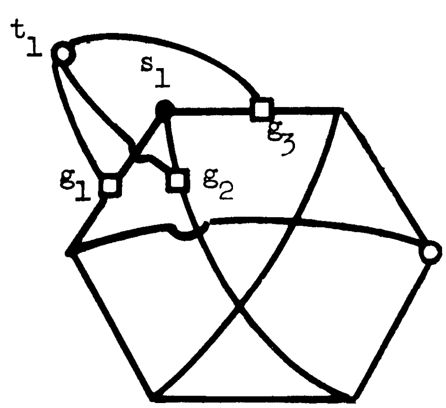


Figure 3.611-a.

l-b. Two blacks and one white.

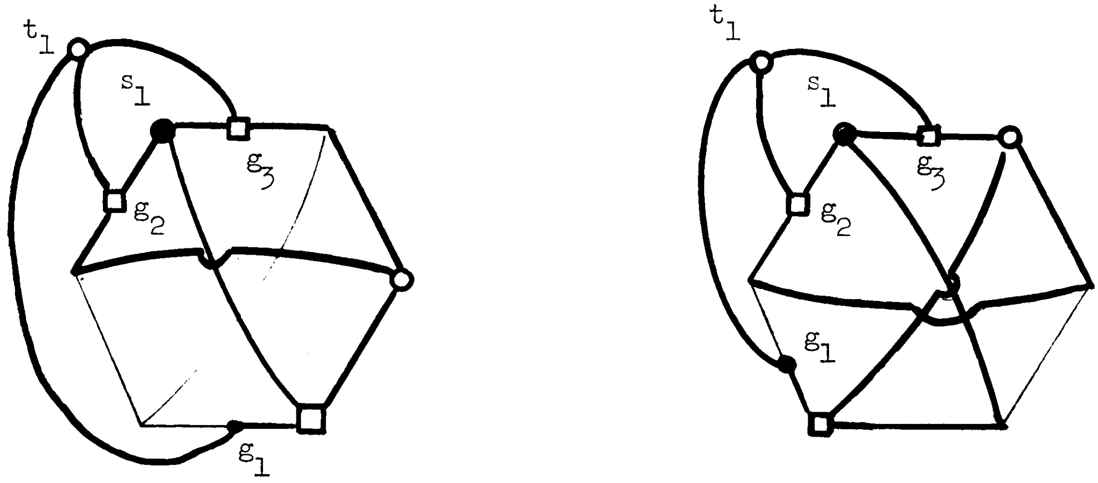


Figure 3.6/1-b.

l-c. Two whites and one black.

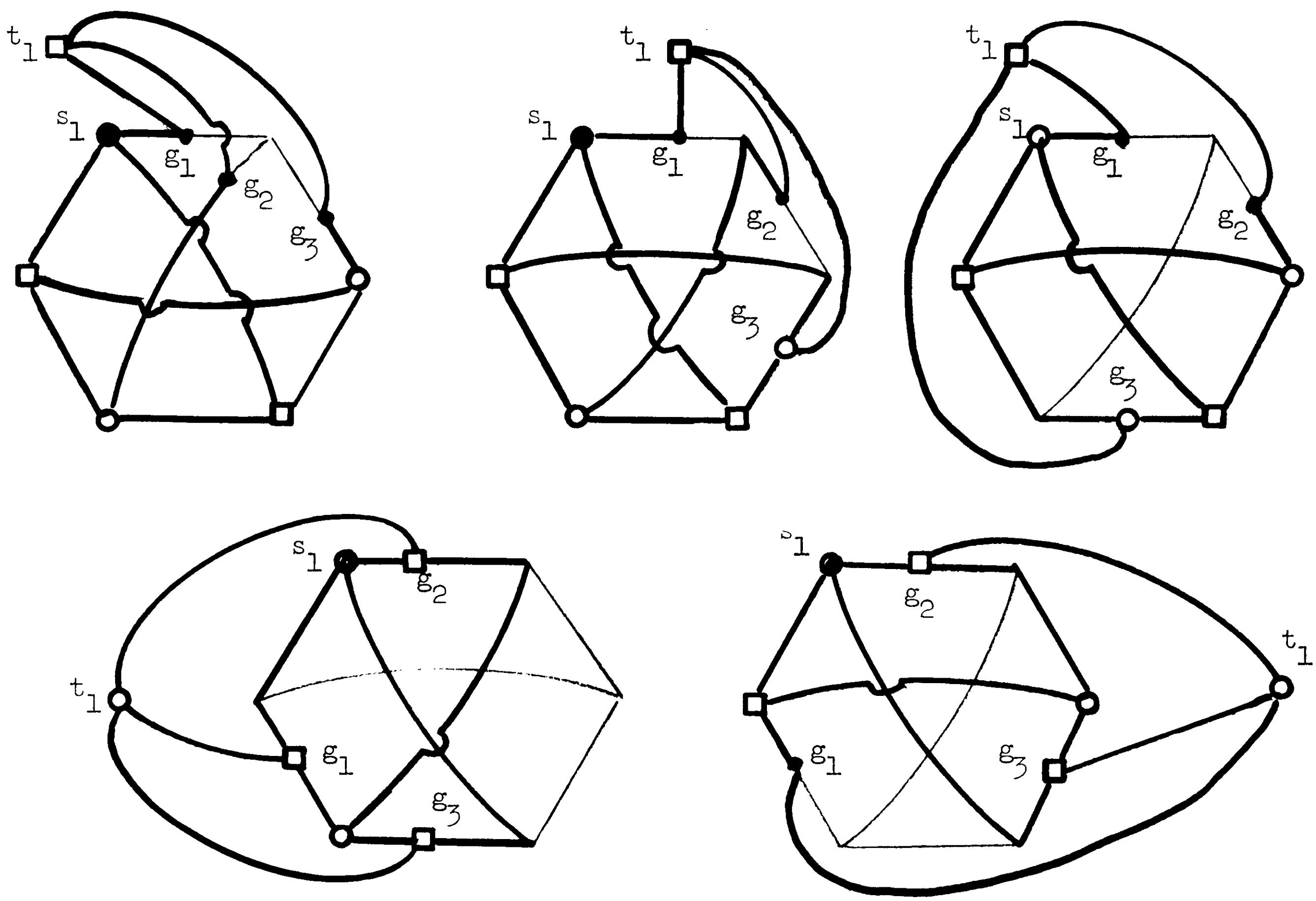


Figure 3.611-c.

1-d. All whites.

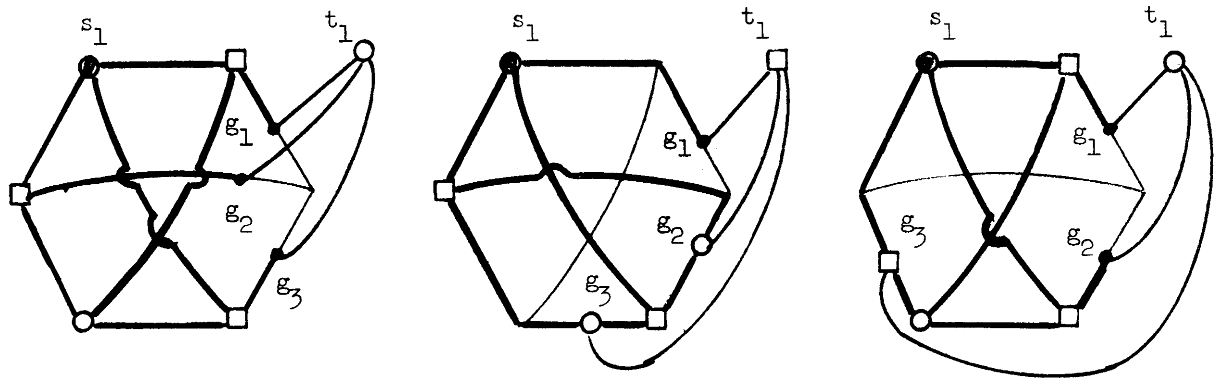
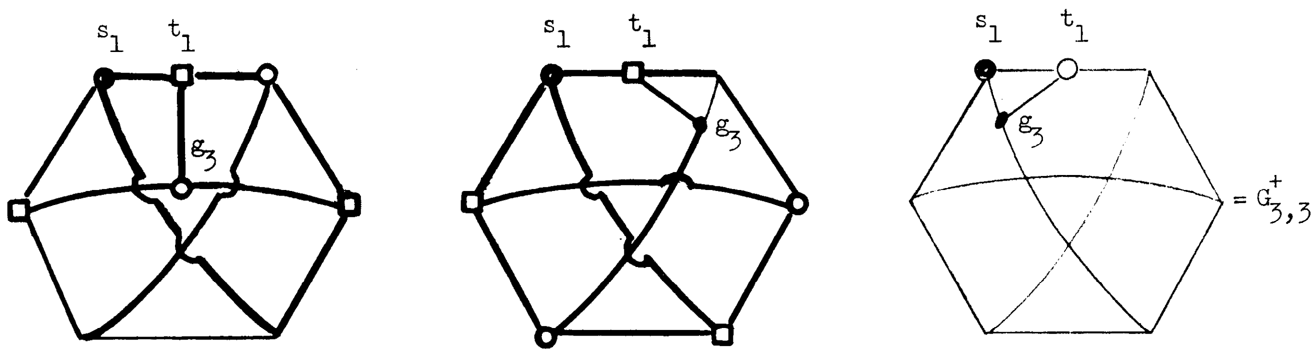


Figure 3.6/1-d.

Case 2, g_1 and g_2 are on the same p-edge and g_3 is on another one. Using Q_1 and Q_2 we can transform $G_{3,3}$, such that t_1 lies on a p-edge and has a connection (disjoint to $G_{3,3}$) to another p-edge.

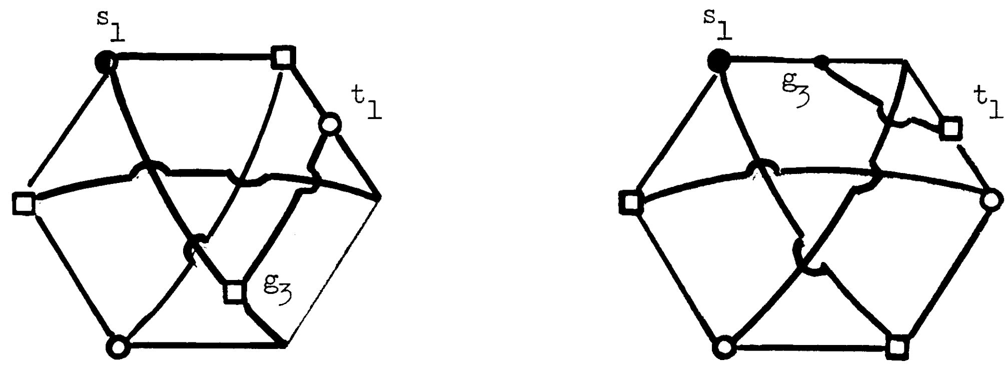
2-a. t_1 lies on a black -p-edge.



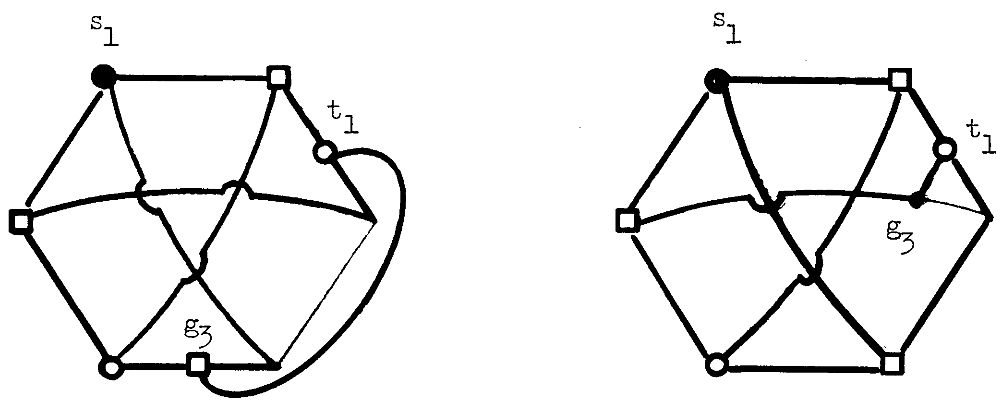
connected to a white p-edge. connected to a black one.

Figure 3.6/2-a.

2-b. t_1 lies on a white p-edge,



and connected to a black one,



and connected to a white one.

Figure 3.612-b.

<u>Case 3</u>. All the g_i's are on the same p-edge. Using the W-assumption, this case is easily reducible to Case 2.

Q.E.D.

4. Cracking the Nut.

This is the time to use R3. We construct four disjoint paths, π_1, π_2, π_3, and π_4. Connecting s_1, t_1, s_2, and t_2 with four p-vertices on $G_{3,3}$ (or $G^+_{3,3}$) which are different from s_1 and t_1. The main idea is to use $G_{3,3}$ ($G^+_{3,3}$) in order to make two disjoint connections, one between π_1 and π_2 and the other between π_3 and π_4, to yield the desired two disjoint paths. The following case is the easiest, and left to the reader.

Case 1. s_1 and t_1 are p-vertices and they are not connected by a p-edge.

Case 2. s_1 and t_1 are p-vertices connected by a p-edge.

Figure 4.1 shows the unique way (ignoring symmetry) in which π_3 and π_4 can block one vertex of the first pair (in this case, s_1). However, s_1 is "saved" by π_1 connecting it to a p-vertex different from itself. Using the W-assumption, π_1 cannot now block 5 or π_4 since it hits $G_{3,3}$ at a vertex which is not on one of the three p-edges incident with s_1. The two non-symmetric cases are illustrated in Figure 4.2. The connections between π_1 and π_2 and between π_3 and π_4 are heavily lined.

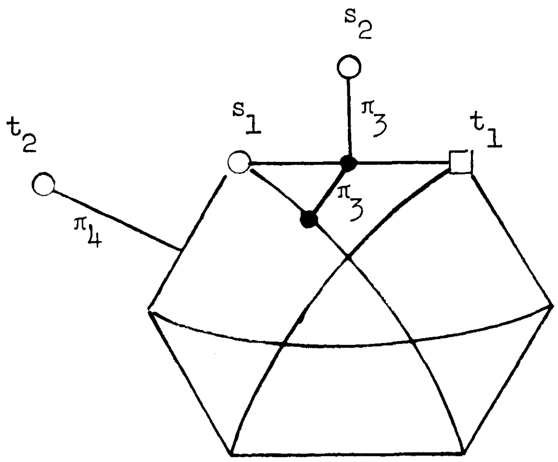


Figure 4.1.

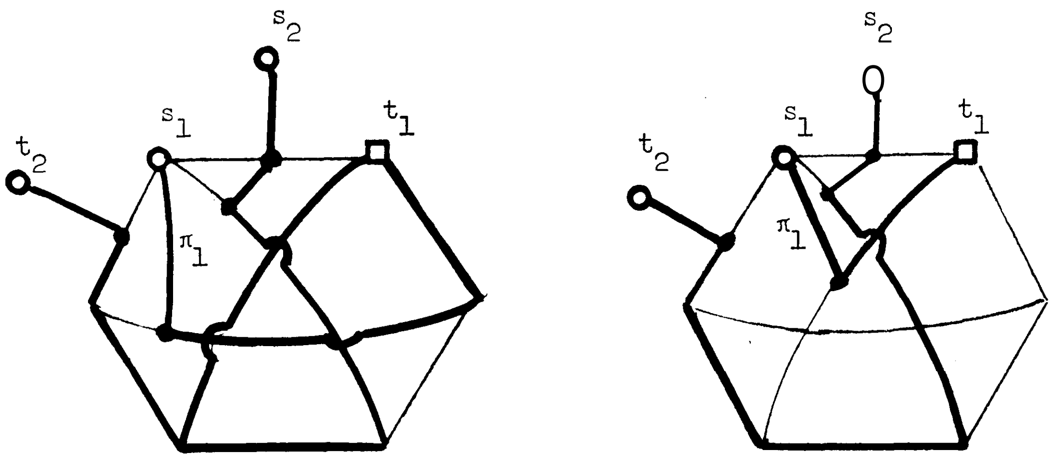


Figure 4.2.

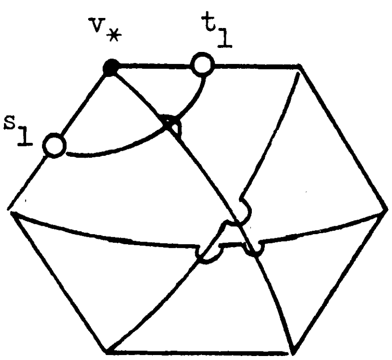
<u>Case</u> 3. $G_{3,3}$ is a subgraph of G .

We proceed in the same way. Four disjoint paths, π_1 , π_2 , π_3 , and π_4 are drawn from s_1 , t_1 , s_2 , and t_2 , respectively, to four distinct p-vertices, different from v_* . (See. Figure 3.5.) Here, we have to consider additions3 cases. However, we have a series of "inevitable moves" which limit the cases analysis. The moves are illustrated in

Figure 4.3. Symmetric cases are ignored.

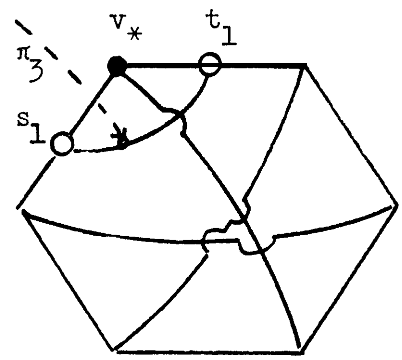
Figure 4.3.

The starting point.



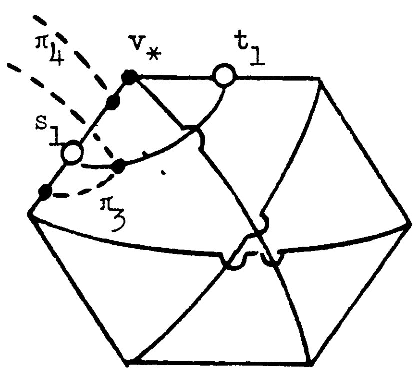
(a)

π_3 (or π_4) must land on the (s_1, t_1) segment, otherwise we are done.



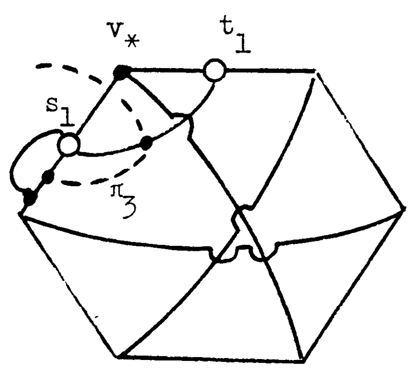
(b)

The only way for π_3 and π_4 to block s_1 (t_1 is symmetric) is shown in Figure 4.3(c).



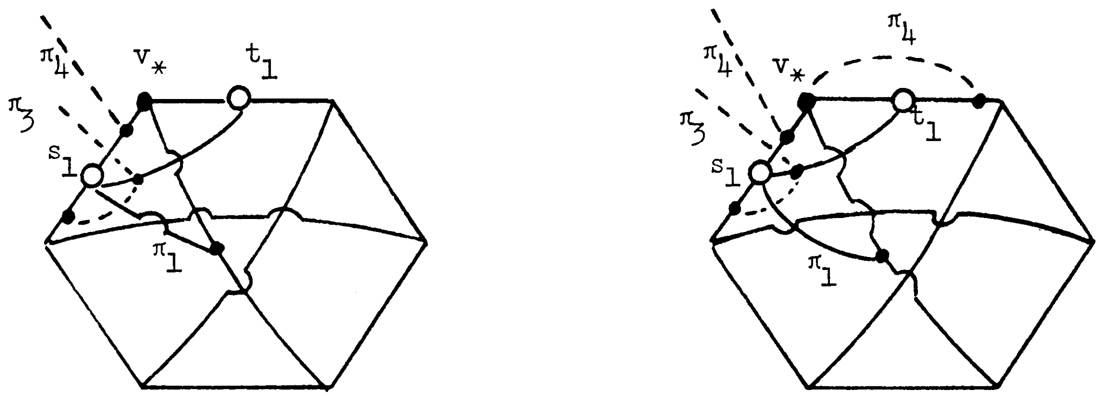
(c)

Forbidden!



(d)

The W-assumption does not allow π_1 to block π_3 in his turn. (See Figure 4.3(d).) However it can still block π_4 . (Figure 4.3(e).) This is its only choice, otherwise we are done.



(e) (f)

π_4 is still alive, since it has to reach a p-vertex different from v_*. The only way to continue the game one step further, is to block t_1, as shown in Figure 4.3(f). Note that we are not allowed to use the W-assumption, here, since if applied, if would throw t_1 out, and we would be left with a graph which is not homeomorphic to that we started with, namely $G^+_{3,3}$.

... t_1 has not yet reached a p-vertex. Thus, π_2 can be used to get it out of the trap. However, it can no longer trap π_3 or π_4.

The nut is finally cracked.

Remark. The words "otherwise we are done" mean that otherwise we can find two disjoint connections between π_1 and π_2 and between π_3 and π_4 as was done in Figure 4.2.

5. -Complexity.

We haven't written the solution as an explicit, long and tedious algorithm. However, its complexity can be easily evaluated if we follow the reductions.

Let linear mean O(n+m) .

R1 is linear. (It involves the linear algorithm [HT1] for decomposing a graph into 3-connected components.)

R2 involves (linear) planarity testing [TH2] and the linear solution for the planar case, [PS].

R3 requires O(n•m) time in worst case. We do not attempt to find four disjoint paths between s_1 , t_1 , s_2 , t_2 and any other four vertices of G_5 (a hameomorph to K_5 , see [W]) in case G_5 is a subgraph of G , or to four p-vertices of $G_{3,3}$ in the other case. If we find such paths -- fine. If not, the size of the graph can be reduced by one vertex at least. Finding these four disjoint paths is linear since at most four augmenting paths are required. (see [ET].)

Following the (constructive) proof of Kuratowski's theorem, we can find a G_5 or a $G_{3,3}$ in G in linear time, provided that G is non-planar.

R4: If a G_5 has been discovered, we follow the lines of [W]. Though quite complicated, Watkins' analysis can be easily implemented in linear time. Assume that four disjoint paths between s_1 , t_1 , s_2 , t_2 and four p-vertices of G_5 are given. Adjustments of G_5 each time the W-assumption is violated have the property of propagating along these paths. This property makes the overall amount of work which is involved in these adjustments to be linear. This is true for $G_{3,3}$ too.

R5 and R6 are obviously linear, and so is the work involved in cracking the nut.

6. Summary.

(a) Not long ago, many people believed that the two paths problem is not polynomial. The two commodity 0-1 flow problem for undirected graphs, which is a close relative of it, is NP-complete, ([I2],[EIS]). We have shown that it is not only polynomial, but "almost" linear. "Almost" means, of course, that there is only one step in the whole algorithm which is not linear. It might be that even this step will be made linear by some sophisticated techniques.

(b) It should be pointed out that this work relies heavily on previous results of Itai, Perl, Shiloach and Watkins.

(c) Generalization of this solution to the case of k (> 2) disjoint paths connecting $s_1,\ldots,s_k$ with $t_1,\ldots,t_k$ respectively, seems to be impossible.

The directed two paths problem also seems to be much more difficult. However, significant results were recently obtained by S. Even, M. Garey, and R. E. Tarjan, [EGT].

(d) The following combinatorially interesting theorem follows from Watkins' work and the results of this paper.

Theorem. If G is an undirected b-connected non-planar graph, then it has the P2 property.

Corollary. Every 6-connected graph has the P2 property.

Proof. A 6-connected graph cannot be planar.

There are 5-connected (planar) graphs that do not have the P2 property, see [W] and [EGT].

Appendix.

We present here a general scheme of the proof of the following theorem.

Theorem A. Let G be an undirected graph with n vertices and m edges. If the TPP (Two Paths Problem) can be solved for any 3-connected graph G' 'having $n' \leq n$ vertices and $m' < m$ edges, in time of T , then it can be solved for G in $O(n+m+T)$ time.

Proof (a general scheme). We present a sequence of polynomial reductions, reducing the TPP from general into 3-connected graphs. Thus, we prove that Theorem A is true if $O(n+m+T)$ is replaced by $O(p(n,m) + T)$ where $p(n,m)$ is a polynomial in n and m. The proof that $p(n,m)$ is actually n+m is not given in full. Most of the reductions have an obvious linear behavior. When linearity is not clear, we support it by more detailed arguments.

The Reductions.

Each of the following reductions assumes that all its predecessors hold. Most of them cannot be proved without this assumption. We may assume that:

Al: G is '&connected.

If not, the problem is reduced (in the worst case) to one of G's 2-connected components. Decomposition of a graph into 2-connected components is linear. Let $ST = \{s_1, t_1, s_2, t_2\}$ be the set of the four vertices of the problem.

A2: If $\{u,v\}$ is a separating set of G , then $ST \cap \{u,v\} = \emptyset$. Otherwise the problem can be reduced to a proper subgraph of G . Some case analysis is involved corresponding to what $ST \cap \{u,v\}$ really is. It is relatively simple (and makes use of A1) and left for the reader. The linearity of this step is not trivial.

Definition. Let $S = \{u,v\}$ be a separating set of G . $G' = (V',E')$ is a weak component mod S if it is a connected component of G-S . $\bar{G}^1 = (V' \cup S, E' \cup E'')$ is a strong component mod S . Here E'' is the set of edges connecting u or v with vertices of V' .

A3: If $G' = (V',E')$ is a weak component mod $\{u,v\}$ then $V' \cap ST \neq \emptyset$. Otherwise we could chop G' off and add the edge (u,v) and obtain an equivalent problem.

Corollary. If $\{u,v\}$ is a separating set of G then G has at most four weak (strong) components mod $\{u,v\}$.

A4: There is no separating set $\{u,v\}$ which separates s_1 and t_1 from s_2 and t_2, otherwise (assuming A1) we have two disjoint paths connecting s_1 with t_1 and s_2 with t_2 .

A5: No set $\{u,v\}$ separates s_1 and s_2 from t_1 and t_2 . Assume to the contrary that such u and v exist. Let G_S and G_T be the strong components mod $\{u,v\}$ containing s_1, s_2 and t_1, t_2 respectively. We first construct two disjoint paths P_1 , P_2 connecting s_1 and s_2 with t_1 and t_2 , (using an O(m+n) flow algorithm such as [ET]). If P_1 connects s_1 with t_1 and P_2 connects s_2 with t_2 , we are done. So let us assume that

P_1 connects s_1 with t_2 and P_2 connects s_2 with t_1 ". Since $\{u,v\}$ separates s_1 and s_2 from t_1 and t_2, we may also assume that P_1 goes through u and P_2 goes through v. It is now easy to see that the original TPP has an affirmative solution iff at least one of the following has.

TPP(S): $G' = G_S$, $s_1' = s_1$, $s_2' = s_2$, $t_1' = v$, $t_2' = u$.

TPP(T): $G'' = G_T$, $s_1'' = u$, $s_2'' = v$, $t_1'' = t_1$, $t_2'' = t_2$.

Note that P_1 and P_2 induce two pairs of disjoint paths, one in G_S and'one in G_T between the sources and sinks in both reduced problems. Thus they are constructed only once and can be used in further reductions of the same type. Theorem A follows now by induction.

If Al through A5 hold and G is not 3-connected, we may assume that s_1 is separated from s_2, t_1, and t_2 by a separating set $\{u,v\}$. If the strong component mod $\{u,v\}$ which contains s_2, t_1, and t_2 is 3-connected, then we are done. The original problem is reduced into two smaller problems, restricted to this 3-connected component, by substituting $s_1 = u$ and $s_1 = v$, one at a time. If this component is not 3-connected, a further decomposition of this component takes place and the worst case is illustrated in Figure A-l. It involves 16 subproblems restricted to the central 3-connected component.

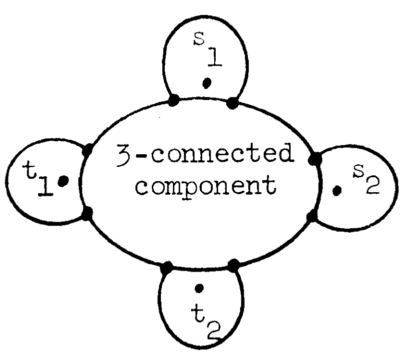


Figure A-l.

References

[EIS] S. Even, A. Itai, and A. Shamir, "On the complexity of time table and multicommodity flow problems," SIAM J. on Computing 5(1976), 691-703.

[EGT] S. Even, M. Carey, and R. E. Tarjan, "The edge-disjoint circuit problem," to appear.

[ET] S. Even and R. E. Tarjan, "Network flow and testing graph connectivity," SIAM J. on Computing 4 (1975), 507-518.

[HT1] J. E. Hopcroft and R. E. Tarjan, "Dividing a graph into triconnected componenets," SIAM J. on Computing 2 (1973), 135-158.

[HT2] J. E. Hopcroft and R. E. Tarjan, "Efficient planarity testing," Journal ACM 21 (1974), 549-568.

[I1] A. Itai, private communication.

[I2] A. Itai, Ph.D. thesis, Weizmann Institute, Rehovot, Israel (1976).

[L] A. LaPaugh, "The subgraph homeomorph problem," Master's thesis, Dept. of Electrical Engineering and Computer Science, Mass. Inst. of Technology.

[PS] Y. Perl and Y. Shiloach, "Finding two disjoint paths between two pairs of vertices in a graph," Journal ACM 25, 1, January 1978.

[W] M. Watkins, "On the existence of certain disjoint arcs in graphs,'* Duke Math. J. 35 (1968), 321-346.