GRAPH 2-ISOMORPHISM IS NP-COMPLETE

by

F. Francis Yao

STAN-CS-79-706
January 19 7 9

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

Graph 2-Isomorphism Is NP-Complete

F. Frances Yao

Computer Science Department
Stanford University
Stanford, California 94305

Abstract.

Two graphs G and G' are said to be k-isomorphic if their edge sets can be partitioned into $E(G) = E_1 \cup E_2 \cup \ldots \cup E_k$ and $E(G') = E'_1 \cup E'_2 \cup \ldots \cup E'_k$ such that as graphs, E_i and E'_i are isomorphic for $1 \leq i \leq k$. In this note we show that it is NP-complete to decide whether two graphs are 2-isomorphic.

This research was supported in part by National Science Foundation grant MCS77-05313.

Given two graphs */ $G = (V, E)$ and $G'(V', E')$ with the same number of edges, by a <u>k-isomorphism of G and G'</u> we mean a partition of $E = E_1 \cup E_2 \cup \ldots \cup E_k$ and a partition of $E' = E'_1 \cup E'_2 \cup \ldots \cup E'_k$ such that as graphs, E_i and E'_i are isomorphic for $1 \le i \le k$. Let $U(G,G')$ be the minimum value of k for which a k-isomorphism of G and G' exists. (See [1] for a study of k-isomorphism of graphs.)

In general, the determination of whether $U(G,G') \le k$ for two graphs G , G' , and positive integer k is an NP-complete problem. For, it clearly belongs to NP; and if we take G' to be a star graph (with the same number of edges as G), then $U(G,G')$ is simply the minimum size of a vertex cover for G , a well-known NP-complete -problem [4]. The question "Is $U(G,G') = 1$?" is the familiar graph isomorphism problem, which is not known to be NP-complete or not [2],[4]. In this note we show that <u>graph-2-isomorphism</u> (G2I), i.e., to decide whether $U(G,G') \le 2$, is an NP-complete -problem.

We will use a transformation from the following problem, which is known to be NP-complete [2].

<u>Exact Cover by 3-Sets (X3C).</u>

Instance: Set $X = \{1,2,\ldots,n\}$ and a family $\mathcal{A} = \{A_i\}$ of 3-element subsets of X .

Question: Does $\mathcal{A}$ contain an exact cover for X , i.e., a subfamily $\mathcal{A}' \subseteq \mathcal{A}$ such that every element of X occurs in exactly one member of $\mathcal{A}'$?

<u>Theorem.</u> X3C is polynomially transformable to G2I. Therefore, the graph 2-isomorphism problem is NP-complete.

*/ We follow [3] for the terminology on graphs.

Proof. Given an instance of X3C, we may assume without loss of generality that $n = 3m \geq 6$, $|\mathcal{S}| = m+\ell \geq m$, and $X \subseteq \bigcup_{i=1}^{m+\ell} A_i$. We shall construct a pair of graphs G and H corresponding to $(X,\mathcal{S})$, as shown in Figure 1.

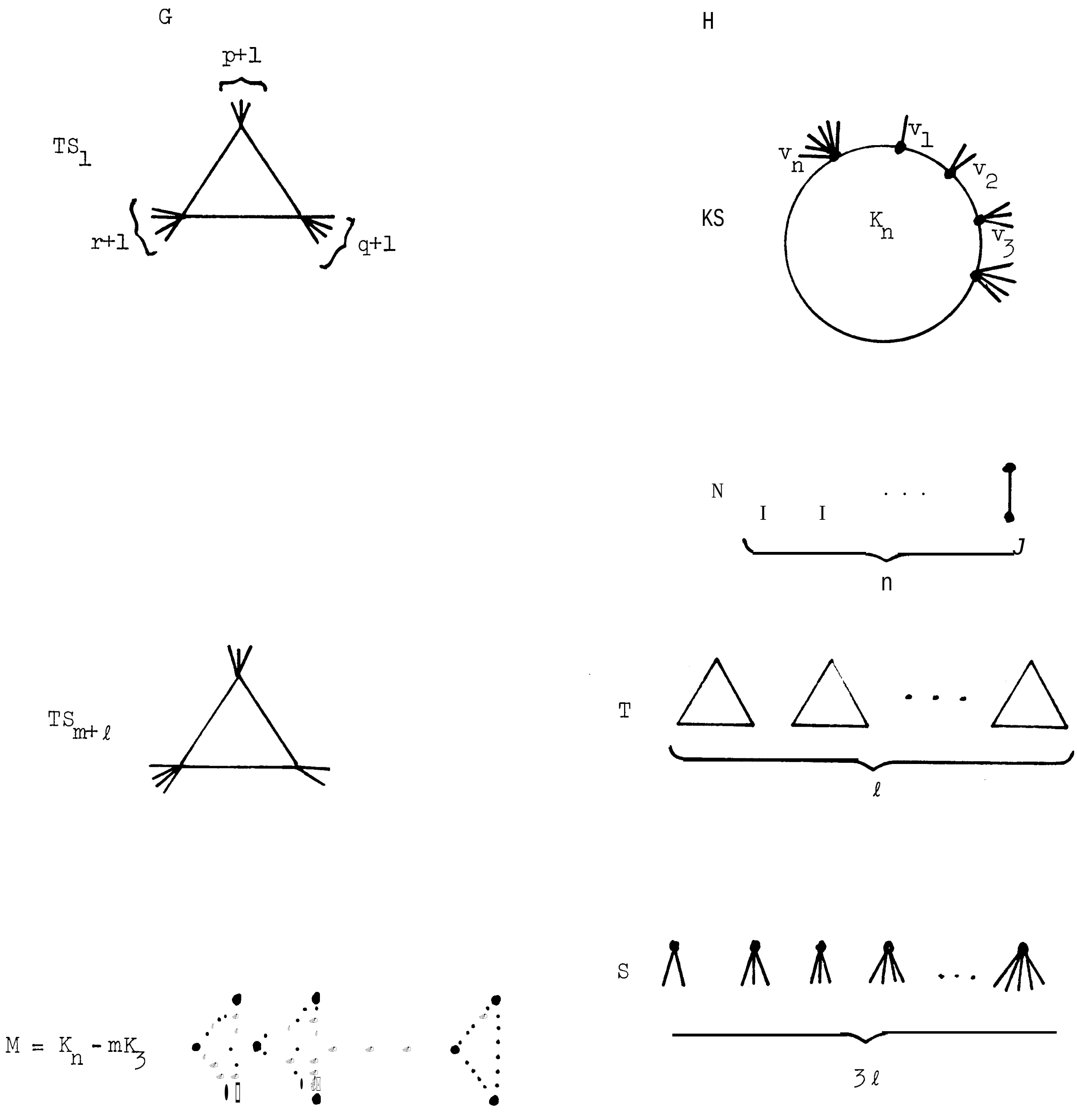
Graph G contains a connected component TS_i corresponding to each A_i in $\mathcal{S}$. If $A_i = \{p,q,r\}$, then TS_i is a triangle, with, additionally, three stars of size $p+1$, $q+1$ and $r+1$ attached to the vertices of the triangle. We will denote $\bigcup_{i=1}^{m+\ell} TS_i$ by TS. In addition to TS, graph G contains a connected component M, which is a complete graph on n vertices with m disjoint triangles removed.

Graph H is the disjoint union of four subgraphs KS, N, T, and S. In KS, we have a complete graph on n vertices $\{v_1,v_2,\ldots,v_n\}$, together with an i-star attached to each v_i. The complete graph of KS will be referred to as K_n henceforth. Subgraph N consists of n disjoint edges, and T consists of ℓ disjoint triangles. Finally, S consists of 3ℓ disjoint stars, one of size $p+1$ for each p that occurs in the multiset $\left(\bigcup_{i=1}^{m+\ell} A_i\right) - X$.

Clearly, G and H can be constructed from $(X,\mathcal{S})$ in polynomial time. Since G and H are not isomorphic, $U(G,H)$ is at least 2. We now show that $U(G,H) < 2$ if and only if $\mathcal{S}$ contains an exact cover for X.

Lemma 1. $U(G,H) \leq 2$ if $\mathcal{S}$ contains an exact cover for X.

Proof of Lemma 1. Without loss of generality, assume that $\{A_1,A_2,\ldots,A_m\}$ forms an exact cover for X. We decompose G and H in two steps as follows.



(dotted lines are the missing edges)

Figure 1. The graphs G and H ,

Step 1. Suppose $A_i = \{p,q,r\}$, where $1 \leq i \leq m$. In the corresponding TS_i, we take a subgraph consisting of the triangle together with stars of size p, q, r, and map it into the triangle in KS incident at $\{v_p,v_q,v_r\}$ with the matching stars. At the same time map the triangles of TS_j, for $m+1 \leq j \leq m+\ell$, onto the ℓ triangles of T .

Step 2. The subgraph that is left in G consists of $n = 3m$ isolated edges from $\bigcup_{i=1}^{m} TS_i$, which are isomorphic to N ; 3ℓ stars from $\bigcup_{j=m+1}^{m+\ell} TS_j$, isomorphic to S ; and subgraph M , which is isomorphic to the remainder of KS . 3

We mention in passing that actually $U(G,H) \leq 3$ for the graphs (G,H) constructed from any $(X,\mathcal{S})$. For we can first map all $m+\ell$ triangles of TS into K_n and T ; next map the $3(m+\ell)$ stars of TS into KS and S ; what is left then in both graphs is isomorphic to $M \cup N$. The rest of this note is devoted to proving the converse of Lemma 1.

Lemma 2. $U(G,H) \leq 2$ only if $\mathcal{S}$ contains an exact cover for X .

We first introduce some notations. Under the assumption $U(G,H) \leq 2$, let $E(G) = G^{(1)} \cup G^{(2)}$, $E(H) = H^{(1)} \cup H^{(2)}$ be fixed partitions of the edge sets, with isomorphism mappings $\varphi_1: G^{(1)} \to H^{(1)}$ and $\varphi_2: G^{(2)} \to H^{(2)}$. For any subgraph F of G (or H), we use $F^{(i)}$ to denote $F \cap G^{(i)}$ (or $F \cap H^{(i)}$, respectively); also, let $\langle F^{(i)} \rangle$ be the isomorphic image of $F^{(i)}$ under $\varphi^{(i)}$ (or $(\varphi^{(i)})^{-1}$, respectively). For a graph $F = (V,E)$, we use $e(F)$ to denote $|E|$. Define vertexcover to be the minimum size of a subset $V' \subseteq V$ such that for every edge $(u,v) \in E$, at least one of u and v belongs to V' . The following facts will be useful.

<u>Fact A</u>. If $F \subseteq K_n$ and vertexcover $\le a$, then K_n-F contains an (n-a) -clique.

<u>Fact B</u>. Let F be a connected component in G . Any edge of H that is incident with a vertex of $\langle F^{(1)} \rangle$ but not contained in $\langle F^{(1)} \rangle$ must belong to $H^{(2)}$.

<u>Proof of Lemma 2</u>. First, we show that any 2-isomorphism of G and H must decompose KS into M and a collection of triangles with stars. Indeed, since KS has more edges than M , we must have either $\langle KS^{(1)} \rangle \cap TS \neq \emptyset$ or $\langle KS^{(2)} \rangle \cap TS \neq \emptyset$. Assume it is the former.

<u>Proposition</u>. Under the assumption that U(G,H) = 2 and $\langle KS^{(1)} \rangle \cap TS \neq \emptyset$, we must have $\langle KS^{(2)} \rangle = \langle K_n^{(2)} \rangle = M$.

<u>Proof of Proposition</u>. Let TS_i be such that $\langle KS^{(1)} \rangle \cap TS_i \neq \emptyset$. Consider the image of $TS_i^{(1)}$ in H . Let $\{v_{i_1}, v_{i_2}, \ldots, v_{i_h}\}$ be the vertices of K_n that are incident with $\langle TS_i^{(1)} \rangle$.

<u>Fact C</u>. (i) $\langle TS_i^{(1)} \rangle \cap K_n$ contains at most h edges.

(ii) $\langle TS^{(1)} \rangle \cap K_n$ contains $\le$ n edges; equality holds only if $\langle M^{(1)} \rangle \cap K_n = \emptyset$.

<u>Proof</u>. (i) is true since TS_i with one edge removed is a tree.

(ii) follows from (i) immediately. □

Fact D. $2 \le h \le n-2$.

Proof. (a) Suppose $h \ge n-1$. Then since K_n has no edges disjoint from $\{v_{i_1}, v_{i_2}, \ldots v_{i_h}\}$, we must have $K_n^{(1)} \subseteq \langle TS_i^{(1)} \rangle$. This implies that $\text{vertexcover}(K_n^{(1)}) \le \text{vertexcover}(TS_i) = 3$. By Fact A, $K_n^{(2)}$ must contain a (n-3)-clique. Since G does not contain a (n-3) -clique when $n \ge 6$, this is impossible.

(b) Next suppose $h = 1$. Then by Fact B, an (n-1) -star R must be contained in $K_n^{(2)}$. Since the maximum degree of a vertex in M is n-3 , we must have $R \subseteq \langle TS_j^{(2)} \rangle$ for some j . But then $\langle TS_j^{(2)} \rangle$ is incident with n vertices of K_n , and the same argument as given in (a), with step 1 and step 2 interchanged, shows that this is impossible, This proves Fact D. □

Fact E. $\langle KS^{(2)} \rangle = \langle K_n^{(2)} \rangle \subseteq M$.

Proof. Given $2 < h < n-2$, and that an $h \times (n-h)$ bipartite graph Y must be contained in $K_n^{(2)}$ because of Fact B, it is easy to see that Y must lie in $\langle M^{(2)} \rangle$, thus $\langle M^{(2)} \rangle$ is incident with all n vertices $\{v_1, v_2, \ldots, v_n\}$. It follows that $\langle KS^{(2)} \rangle = \langle K_n^{(2)} \rangle \subset M$. □

To finish the proof of the Proposition, note that by Fact E, the edges of K_n are divided into those in $\langle TS^{(1)}\rangle \cap K_n$ and those in $(\langle M^{(1)}\rangle \cup \langle M^{(2)}\rangle) \cap K_n$. This is possible only if the latter contains $e(M) = \binom{n}{2} - n$ edges and the former contains n edges, because of Fact C (ii). But then, $\langle M^{(1)}\rangle \cap K_n = \emptyset$ by Fact C, which implies that $e(\langle M^{(2)}\rangle \cap K_n) = \binom{n}{2} - n$, and hence $\langle K_n^{(2)}\rangle = M$. This proves the Proposition. □

We can now complete the proof of Lemma 2. It follows from the Proposition that $KS^{(2)}$ is the isomorphic image of M, while $KS^{(1)}$ consists of m disjoint triangles, each attached with three stars. Without loss of generality, write $\langle KS^{(1)}\rangle = TS_1^{(1)} \cup TS_2^{(1)} \cup \ldots \cup TS_m^{(1)}$ where for $1 \leq i \leq m$, $TS_i^{(1)}$ is a subgraph of TS_i and moreover, they are triangles with stars of size $\{p',q',r'\}$ and $\{p+1, q+1, r+1\}$ respectively, with $p' \leq p$, $q' \leq q$ and $r' \leq r$.

If $\mathcal{S}$ does not contain an exact cover for X, then we will not have $p' = p$, $q' = q$, $r' = r$ in $TS_i^{(1)}$ and TX_i for all $1 \leq i \leq m$. Hence $TS_1^{(2)} \cup TS_2^{(2)} \cup \ldots \cup TS_m^{(2)}$ will contain fewer than n isolated edges. This makes it necessary, because of the subgraph N in H, for $TS_{m+1} \cup TS_{m+2} \ldots \cup TS_{m+\ell}$ to yield $\delta \geq 1$ isolated edges in either step 1 or 2. Assume without loss of generality that TS_{m+1} contributes an isolated edge (u,v) in step 1. We examine two cases.

<u>Case 1</u>. Suppose (u,v) is in the triangle of TS_{m+1}. Then $TS_{m+1}^{(2)}$ contains a path of length 4, which does not exist in $N \cup T \cup S$ of H (Figure 2(a), 2(b)).

<u>Case 2</u>. Suppose (u,v) is in one of the stars of TS_{m+1}. Then in $TS_{m+1}^{(2)}$, u is a vertex of degree ≥ 3, and hence must be

mapped by φ_2 into a star of S . This implies that $TS^{(1)}_{m+1}$ contains a path of length ≥ 3 , which again does not exist in $N \cup T \cup S$ (Figure 2(c), 2(d)).

Thus we can have U(G,H) = 2 only if $\mathcal{S}$ contains an exact cover for X , and this completes the proof of Lemma 2 and the Theorem. □

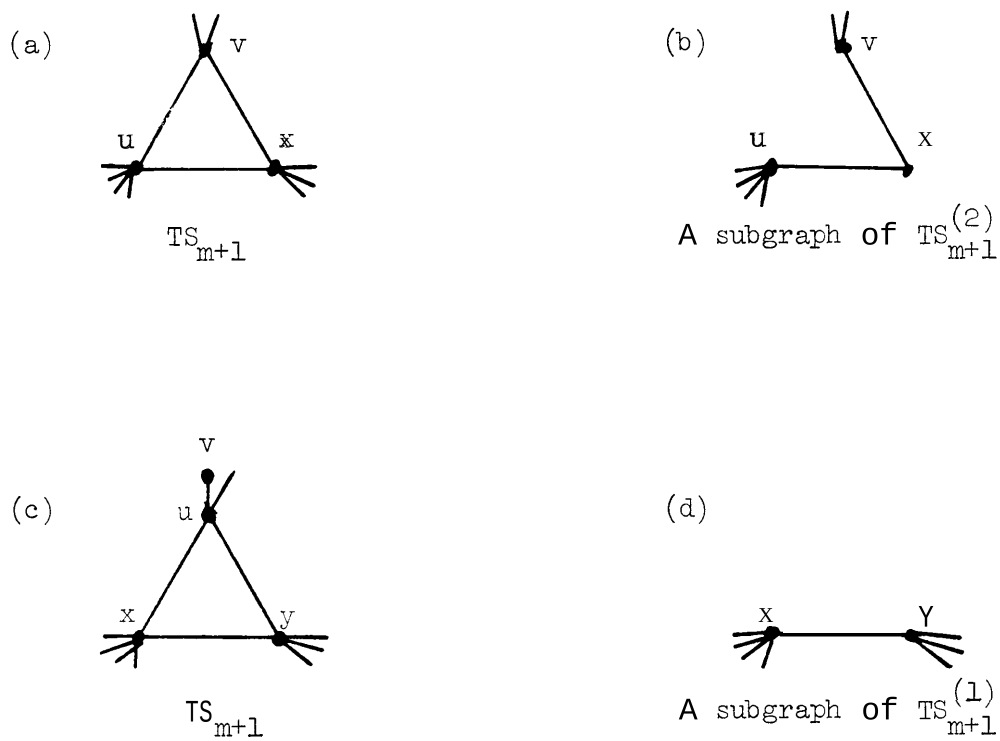


Figure2

We wish to point out that in our construction, it is necessary to employ a different representation for elements of X in TS than in KS (such as using (p+1) -stars versus p-stars for $p \in X$). The following example shows that, for instance, if just p-stars were used in both G' and H' , then one could have U(G',H') = 2 even though $\mathcal{S}$ does not contain an exact cover for X .

Example. Let $X = \{1,2,\ldots,6\}$ and $\mathcal{S} = \{A_1 = \{1,2,5\}, A_2 = \{4,5,6\}, A_3 = \{2,3,4\}\}$. (See Figure 3. We use R_p to denote a p-star.) One can first map two of the edges of R_5 in TS_1 into the R_5 of s ; the triangle of TX_3 into T ; and M into K_n . The remaining subgraphs of G' and H' are then isomorphic. Such unwanted phenomena cannot be remedied simply by choosing other representations, say, using p^2 -stars for $p \in X$, in both G' and H' .

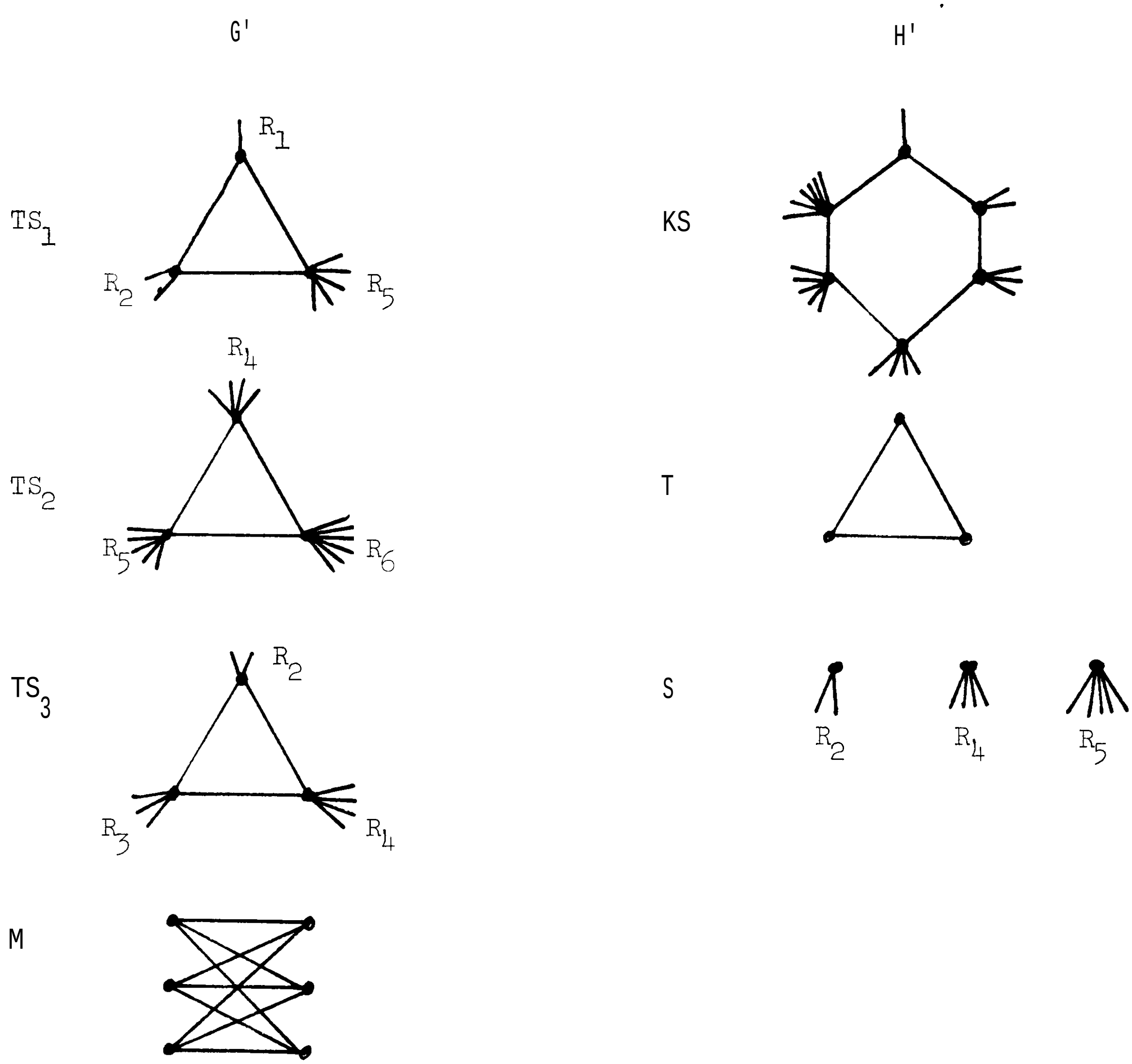


Figure 3. An example with $U(G',H') = 2$.

References

[1] F. Chung, P. Erdős, R. Graham, S. Ulam, F. Yao, "Minimal decompositions of two graphs into pairwise isomorphic subgraphs," to appear.

[2] M. Garey, D. Johnson, Computers and Intractability, Freeman and Company, San Francisco, CA, 1979.

[3] F. Harary, Graph Theory, Addison-Wesley, Reading, Mass,, 1969.

[4] R. Karp, "Reducibility among combinatorial problems," in Complexity of Computer Computations, edited by R. Miller and J. Thatcher, Plenum Press, 1972, 85-104.