Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Discrete Algorithms 4 (2006) 201–208
JOURNAL OF DISCRETE ALGORITHMS
www.elsevier.com/locate/jda

Polynomial recognition of equal unions in hypergraphs with few vertices of large degree

David P. Jacobs [a,*], Robert E. Jamison [b]

[a] *Clemson University, Department of Computer Science, Clemson, SC 29634 USA*
[b] *Clemson University, Department of Mathematical Sciences, Clemson, SC 29634 USA*

Available online 13 April 2005

Abstract

A family of nonempty sets has the *equal union property* if there exist two nonempty disjoint subfamilies having equal unions. If every point belongs to the unions, then we say the family has the *full equal union property*. Recognition of both properties is NP-complete even when restricted to families for which the degree of every point is at most three. In this paper we show that both recognition problems can be solved in polynomial time for families in which there is a bound on the number of points whose degree exceeds two.

Keywords: Hypergraphs; Equal union property; L-matrices; Polynomial time; Turing reduction

1. Introduction

Let $F = (S_1, \ldots, S_m)$ be a family of nonempty subsets of a finite set V. We say that F has the *equal union property*, denoted EUP, if and only if there exist nonempty disjoint sets $\Gamma, \Delta \subseteq \{1, \ldots, m\}$, for which

$$\bigcup_{\gamma\in\Gamma} S_\gamma = \bigcup_{\delta\in\Delta} S_\delta. \tag{1}$$

* Corresponding author.
E-mail addresses: dpj@clemson.edu (D.P. Jacobs), rejam@clemson.edu (R.E. Jamison).

1570-8667/$ – see front matter
doi:10.1016/j.jda.2005.03.001

The set in (1) is called the *support* of the equal union. A stronger requirement is that the support also equal V, in which case we say F has the *full equal union property*, denoted FEUP. Early papers on the EUP were by Lindström in [8] and Tverberg in [12]. From [8] it follows that

Theorem 1 *(Lindström). Any family of at least $k+1$ nonempty sets whose union has k elements has the equal union property.*

Let $F = (S_1, \ldots, S_m)$ where each $S_i \subseteq V = \{x_1, \ldots, x_n\}$. The *incidence matrix* of F is the $n \times m$ matrix $M = (m_{ij})$ of zeroes and ones in which $m_{ij} = 1$ if and only if $x_i \in S_j$. The *degree* of x_i, written $\deg(x_i)$, is the number of sets containing it. Note that the number of 1's in column j is the cardinality of S_j, and the number of 1's in row i is the $\deg(x_i)$.

The equal union property is related to L-matrices. A real $n \times m$ matrix is an *L-matrix* if and only if the columns of every real $n \times m$ matrix having the same sign-pattern are linearly independent. In [3] it was shown that a family $F = (S_1, \ldots, S_m)$ of nonempty subsets of $V = \{x_1, \ldots, x_n\}$ has the EUP if and only if its incidence matrix M is not an L-matrix. In recent papers, Robertson, Seymour, and Thomas [11], and McCuaig [10], independently showed that many related problems, including the recognition of square L-matrices (called sign-nonsingular matrices [9]), can be done in polynomial time. It follows that for families of n sets on n points, the EUP can be recognized in polynomial time. In [5] it was shown that recognizing the FEUP in n sets on n points is NP-complete. Thus, the complexity of the $n \times n$ case is settled.

Our paper is concerned with the complexity of recognizing the EUP and FEUP in the general case of n points and m sets. In [7], Klee, Ladner, and Manber showed that recognizing non-L-matrices, and hence the EUP, is NP-complete. In fact, the recognition problems for the EUP and the FEUP are NP-complete, even when restricted to the very specialized families having only sets of cardinality at most three, and points of degree at most three [6].

In [4] it was shown that if we bound the number of sets whose *cardinality* exceeds two, both problems can be solved in polynomial time. The purpose of this paper is to show that both problems can also be solved in polynomial time by bounding the number of points whose *degree* exceeds two. In Section 2 we obtain the result for the EUP, and in Section 3 we obtain the result for the FEUP.

It will sometimes be more descriptive to speak of a *hypergraph* $H = (V, E)$ rather than a family of sets. Here V is the set of points (or vertices) and E is the family of subsets (or edges).[1] We use *graph* only when all edges have cardinality two.

2. A bounded number of points of large degree

The purpose of this section is to prove:

[1] We will assume that $\emptyset \notin E$, although we allow V or E to be empty. Note that a necessary condition for an equal union is that $|E| \geqslant 3$ and $|V| \geqslant 2$.

Theorem 2. *For each fixed integer b recognition of EUP in the class of hypergraphs with at most b vertices of degree $\geqslant 3$ is polynomial.*

Note that given a hypergraph H, we may always remove vertices of degree zero or one: Clearly any vertices of degree zero cannot contribute to an EUP. The unique edge containing any degree one vertex cannot be used in an equal union, and hence both the edge and the vertex may both be removed without altering EUP. This last operation could introduce other degree zero or degree one vertices, but we may repeat these operations until we obtain a hypergraph H' in which all points have degree at least two. Evidently H has the EUP if and only if H' does.

The main idea in the proof of Theorem 2 is to replace certain sets of degree two points with a single special degree two point, consolidate certain sets of edges, and remove other sets of edges.

We do this by considering a special intersection graph, denoted $I(H)$, whose vertices are the edges of $H = (V, E)$. Two members of E, A and B, are adjacent in $I(H)$ if and only if they contain a point of V of degree two. That is, A and B are adjacent iff there is a point $v \in A \cap B$ but belonging to no other edge. By a *component* we mean a connected component in $I(H)$, but we think of it as an edge set in H. The components form a partition of E. When a component C contains only one edge A, we call C *trivial*. Note this occurs exactly because A does not contain any vertices having degree two.

Lemma 1. *Let $H = (V, E)$ be a hypergraph, and let $C \subseteq E$ be a component in $I(H)$. If some edge in C is used in an equal union, then*

(i) *all edges in C must be used*;
(ii) *C is a* (*possibly trivial*) *bipartite graph in $I(H)$*; *and*
(iii) *the division of the hyperedges in C, between the two parts of the equal union, yields the unique 2-coloring of C.*

Proof. Suppose S and T are edges of H that are adjacent as vertices of C. Since they contain a common degree 2 point, if one is used in an equal union then the other must be used but in the other half. This also precludes an odd cycle in C. Since C is connected, the lemma follows. □

The following observation implies a polynomial time EUP recognition algorithm in the simple case when all vertices have degree two.

Lemma 2. *If $H = (V, E)$ has only vertices of degree 2, then H has EUP if and only if $I(H)$ has a nontrivial bipartite component.*

Proof. Assume H has the EUP. Since all vertices have degree two, all components of $I(H)$ are nontrivial. Hence the "only if" part follows from Lemma 1, part (ii). Conversely, assume $I(H)$ has a nontrivial bipartite component C. Then C may be partitioned into color classes C_1 and C_2. It is easy to see that these two edge sets form an equal union since the only vertices contained in these edges are the degree two points. □

We wish to describe a transformation that simplifies a hypergraph while preserving the status of the EUP property. Let us say that a hypergraph H is *reduced* if

(i) all vertices have degree at least two,
(ii) all components of $I(H)$ are bipartite, and
(iii) every nontrivial component of $I(H)$ has size two.

We describe some operations which may be applied to obtain a reduced hypergraph. Each operation reduces the size of the graph. We have already seen how to remove vertices of degree zero and one, and so property (i) is easy to obtain.

To obtain property (ii), suppose C is a nonbipartite component of $I(H)$. Then by Lemma 1, no edge in C can participate in an equal union, so we may remove all edges in C. Note this operation can injure property (i), so further elimination of degree zero and one vertices may be necessary.

To obtain property (iii), let C be any nontrivial bipartite component of $I(H)$ of size greater than two. Then there is a unique bipartition of C into color classes, say C_1 and C_2. We introduce a new point ∞_C. For $i = 1, 2$, let X_i denote ∞_C together with all vertices v of H with degree $\geqslant 3$, where v lies in a hyperedge of color class C_i. We then *replace* all edges in C with $\{X_1, X_2\}$, and *replace* the corresponding degree-two vertices with ∞_C. It is easy to check that this operation preserves the status of the EUP.

In the previous operation, there is one extreme case that can occur. It is possible that the two unions X_1 and X_2 are equal. But this implies the EUP, so we transform H to the special hypergraph

$$H_2 = \big(\{1, 2\}, \big\{\{1\}, \{2\}, \{1, 2\}\big\}\big) \tag{2}$$

which has the EUP, and which we'll also define as reduced. Let us say a point x has *large degree* if $\deg(x) \geqslant 3$. From the above discussion, it follows that

Lemma 3. *In polynomial time, we can transform any hypergraph H to a reduced hypergraph H' that has the EUP if and only if H does, and that has no more vertices of large degree than does H.*

The next result shows that if a reduced hypergraph has too many nontrivial components, then H is forced to have the EUP.

Lemma 4. *Let H be a reduced hypergraph with b vertices of large degree. If $I(H)$ has $k > b$ nontrivial components, then H has the EUP.*

Proof. If $H = H_2$, we are done. Otherwise, since H is reduced, it has exactly one degree two vertex belonging to each nontrivial component. So there are exactly $k + b$ points in H. Since each nontrivial component has two edges, the number of edges is at least $2k > k + b$. By Theorem 1, H has an equal union. $\square$

Lemma 5. *Let H be a reduced hypergraph with b vertices of large degree. If $I(H)$ has $k \leqslant b$ nontrivial components, then H has at most $2b$ vertices.*

Proof. Edges belonging to trivial components have no degree two vertices, and nontrivial components have one degree two vertex, so H has $k \leqslant b$ degree two vertices. There are also b vertices of large degree, and so H has at most $2b$ vertices. □

Proof of Theorem 2. Let H be a hypergraph having at most b vertices of large degree. By Lemma 3, in polynomial time we may transform this to a reduced hypergraph H' having $b' \leqslant b$ vertices of large degree. If the number of nontrivial components in $I(H')$ exceeds b' then, by Lemma 4, H' (and hence H) has the EUP. Otherwise, by Lemma 5, H' has at most $2b' \leqslant 2b$ vertices, and in constant time we may decide if it has the EUP. □

3. A polynomial time Turing reduction

Recall that if Π_1 and Π_2 are problems, then Π_1 is *polynomial time Turing reducible* to Π_2, written $\Pi_1 \leqslant_T \Pi_2$, provided Π_1 can be solved in polynomial time using an oracle machine which allows queries to solve instances of Π_2 (see [1,2]). Like the more familiar and stronger notion known as a polynomial transformation, if $\Pi_1 \leqslant_T \Pi_2$, then a polynomial time algorithm for Π_2 implies one for Π_1.

The purpose of this section is to obtain a polynomial time Turing reduction from the decision problem **FEUP** to the decision problem **EUP**, when restricted to certain classes of hypergraphs.

We temporarily relax the definition of a hypergraph as follows. A *multi-hypergraph* is a pair (V, F) in which F is a multiset of nonempty subsets of V. Every hypergraph is a multi-hypergraph. The properties EUP and FEUP are defined for multi-hypergraphs in the obvious way. Note that if a multi-hypergraph has a repeated edge, then it has the EUP trivially.

Let $H = (V, F)$ be multi-hypergraph, and let $S \subseteq V$. Let F_S denote the *multiset* $\{A - S \mid A \in F, A - S \neq \emptyset\}$. We form the multi-hypergraph $H_S = (V - S, F_S)$, and say that H_S is a *trace* of H. Even if H is a hypergraph, F_S might have repeated edges.

Lemma 6. *Let $H = (V, F)$ be a multi-hypergraph, and let S be the support of some equal union. If $S \neq V$ then H_S has the FEUP if and only if H has the FEUP.*

Proof. First note that if H has the FEUP with

$$V = \bigcup_{i \in I} A_i = \bigcup_{j \in J} A_j$$

then for *any* $S \neq V$

$$\emptyset \neq V - S = \bigcup_{i \in I} (A_i - S) = \bigcup_{j \in J} (A_j - S).$$

Letting $I' = \{i \in I \mid A_i \nsubseteq S\}$ and $J' = \{j \in J \mid A_j \nsubseteq S\}$ we still have

$$\emptyset \neq V - S = \bigcup_{i \in I'} (A_i - S) = \bigcup_{j \in J'} (A_j - S),$$

and so H_S has the FEUP as well.

Now assume $S \neq V$ is the support of some equal union and for disjoint index sets I and J we have

$$S = \bigcup_{i \in I} A_i = \bigcup_{j \in J} A_j \neq \emptyset, \tag{3}$$

and assume further that H_S has the FEUP. Then there exist disjoint index sets K and L such that

$$V - S = \bigcup_{k \in K} (A_k - S) = \bigcup_{l \in L} (A_l - S).$$

By definition of H_S, the sets A_k and A_l are not contained in S. From (3) we have

$$\begin{aligned} V &= \bigcup_{k \in K} (A_k - S) \cup \bigcup_{i \in I} A_i \\ &= \bigcup_{l \in L} (A_l - S) \cup \bigcup_{j \in J} A_j. \end{aligned}$$

The index sets K, I, L, and J are pairwise disjoint because the A_i and A_j are contained in S while the A_k and A_l are not. Since each $A_k - S \subseteq A_k \subseteq V$ and $A_l - S \subseteq A_l \subseteq V$ we must have

$$V = \bigcup_{x \in K \cup I} A_x = \bigcup_{y \in L \cup J} A_y$$

and so H has the FEUP. $\square$

Assume that $\mathcal{C}$ is any class of multi-hypergraphs. In what follows, **EUP**($\mathcal{C}$) and **FEUP**($\mathcal{C}$) refer to the decision problems when restricted to multi-hypergraphs in $\mathcal{C}$. It will be convenient to introduce a new problem we call **EUI**. Given a multi-hypergraph, this problem asks to *identify* the support of an equal union if it has the equal union property. Strictly speaking, the problem asks to evaluate the function which returns the empty set if no equal union exists, or otherwise returns the support of some equal union. We use **EUI**($\mathcal{C}$) to refer to its restriction to $\mathcal{C}$.

Lemma 7. *For any class $\mathcal{C}$ closed under the trace,* **FEUP**($\mathcal{C}$) $\leqslant_T$ **EUI**($\mathcal{C}$).

Proof. Assume that we have an oracle to evaluate **EUI** for any $H \in \mathcal{C}$. We are given a multi-hypergraph $H = (V, F)$ and wish to know if it has the FEUP. We ask the oracle to identify a support S. If the oracle returns $\emptyset$, then H has no equal union, and hence no full equal union. If $S = V$, we are done. Otherwise, $\emptyset \neq S \subset V$, and we form H_S. By Lemma 6, we have reduced H to an equivalent multi-hypergraph H_S. By assumption, $H_S \in \mathcal{C}$, so we may continue. This must stop in polynomial time because each trace operation reduces the size of the multi-hypergraph. $\square$

Now suppose we have an oracle for solving **EUP**. We wish to use this to solve **EUI**. Given a multi-hypergraph $H = (V, F)$, we first ask if it has any EUP. If not, we return the empty set. If $H = (V, F)$ has the EUP, let us say that an edge $A \in F$ is *necessary*, if the multi-hypergraph $(V, F - \{A\})$ does *not* have the EUP. Given our oracle, we can identify edges which have this property. If every edge is necessary, then H has the FEUP, and V is the required support of an equal union. Otherwise, if some edge A is not necessary, we form the multi-hypergraph $(V, F - \{A\})$ and continue. This process terminates in polynomial time and justifies Lemma 8.

Lemma 8. *If $\mathcal{C}$ is a class of multi-hypergraphs that is closed under the operations of removing edges and removing degree-zero vertices then* $\mathbf{EUI}(\mathcal{C}) \leqslant_T \mathbf{EUP}(\mathcal{C})$.

Lemma 9. *If $\mathcal{C}$ is a class of multi-hypergraphs that is closed under taking traces, removing edges, and removing degree-zero vertices, then* $\mathbf{FEUP}(\mathcal{C}) \leqslant_T \mathbf{EUP}(\mathcal{C})$.

Proof. This follows from Lemmas 7 and 8. □

Now let $\mathcal{C}$ be a class of hypergraphs, and let $\mathcal{C}^*$ be the smallest class of multi-hypergraphs that contains $\mathcal{C}$ and is closed under the three operations of Lemma 9. Let us call $\mathcal{C}$ *closed* if every hypergraph of $\mathcal{C}^*$ is a member of $\mathcal{C}$.

Theorem 3. *For any closed class of hypergraphs $\mathcal{C}$,* $\mathbf{FEUP}(\mathcal{C}) \leqslant_T \mathbf{EUP}(\mathcal{C})$.

Proof. First note that since $\mathcal{C} \subseteq \mathcal{C}^*$, we have

$$\mathbf{FEUP}(\mathcal{C}) \leqslant_T \mathbf{FEUP}(\mathcal{C}^*). \tag{4}$$

By Lemma 9, we have

$$\mathbf{FEUP}(\mathcal{C}^*) \leqslant_T \mathbf{EUP}(\mathcal{C}^*). \tag{5}$$

And finally we claim

$$\mathbf{EUP}(\mathcal{C}^*) \leqslant_T \mathbf{EUP}(\mathcal{C}). \tag{6}$$

An algorithm for $\mathbf{EUP}(\mathcal{C}^*)$ based on an oracle for $\mathbf{EUP}(\mathcal{C})$ is as follows. For H any multi-hypergraph in $\mathcal{C}^*$, if H has a repeated edge, then H has the EUP, so return "TRUE". If H has does not have a repeated edge, then H is a simple hypergraph and hence lies in $\mathcal{C}$, since $\mathcal{C}$ is closed. Hence $\mathbf{EUP}(\mathcal{C})$ may be applied to H and its result returned. By (4)–(6) and the transitivity of $\leqslant_T$, the lemma follows. □

Theorem 4. *For each fixed integer b recognition of FEUP in the class of hypergraphs with at most b vertices of degree $\geqslant 3$ is polynomial.*

Proof. Let $\mathcal{C}$ be a class of hypergraphs having at most b points of degree $\geqslant 3$. We observe that the operations in Lemma 9 never increase the degree of any points, and so $\mathcal{C}$ is a closed class of hypergraphs. By Theorem 3, it follows that $\mathbf{FEUP}(\mathcal{C}) \leqslant_T \mathbf{EUP}(\mathcal{C})$. Finally, the result follows from Theorem 2. □

If we let $\mathcal{C}$ denote the class of all hypergraphs then it follows that $\mathbf{FEUP}(\mathcal{C}) \leqslant_T \mathbf{EUP}(\mathcal{C})$. But it is interesting to note that this is a polynomial time Turing reduction, not a polynomial transformation. Thus, we cannot use the NP-completeness of **FEUP** to directly infer the NP-completeness of **EUP**.

References

[1] J.L. Balcázar, J. Diaz, J. Gabarró, Structural Complexity, vol. I, Springer-Verlag, Berlin, 1988.

[2] M.R. Garey, D.S. Johnson, Computers and Intractability, A Guide to the Theory of NP-Completeness, Freeman, New York, 1979.

[3] D.P. Jacobs, R.E. Jamison, A note on equal unions in families of sets, in: B. Lindström, J. Zaks, D.G. Rogers (Eds.), Discrete Math. Selected Papers in honor of Helge Tverberg 241 (2001) 387–393.

[4] D.P. Jacobs, R.E. Jamison, Complexity of recognizing equal unions in families of sets, J. Algorithms 37 (2000) 495–504.

[5] D.P. Jacobs, R.E. Jamison, A.A. McRae, On the complexity of sign-nonsingularity and equal unions of sets, in: Proceedings of the 38th ACM SE Conference, Clemson, 2000, pp. 232–234.

[6] D.P. Jacobs, R.E. Jamison, NP-completeness for equal unions in families of sets, in: Proceedings of the 39th ACM SE Conference, Athens, 2001, pp. 73–75.

[7] V. Klee, R. Ladner, R. Manber, Sign solvability revisited, Linear Algebra Appl. 59 (1984) 131–157.

[8] B. Lindström, A theorem on families of sets, J. Combin. Theory (A) 13 (1972) 274–277.

[9] T.J. Lundy, J. Maybee, J. Van Buskirk, On maximal sign-nonsingular matrices, Linear Algebra Appl. 247 (1996) 55–81.

[10] W. McCuaig, Pólya's permanent problem, manuscript, June 1997 (93 pages).

[11] N. Robertson, P.D. Seymour, R. Thomas, Permanents, Pfaffian orientations, and even directed circuits, Annals of Math. (2) 150 (3) (1999) 929–975.

[12] H. Tverberg, On equal unions of sets, in: L. Mirsky (Ed.), Studies in Pure Mathematics, Academic Press, London, 1971, pp. 249–250.