

APPLIED MATHEMATICS

UDC 519.1, 512

V. G. Minasyan

On Set of All Maximum Independent Sets of Bipartite Graph

(Submitted by academician Yu. H. Shoukourian 14/XII 2012)

Keywords: *bipartite graph, maximum independent set, generation problem, distributive lattice, algorithm, complexity.*

In this paper it is shown that the set of all maximum independent sets of bipartite graph is a distributive lattice, which allows to view the problem of generating the maximum independent sets of bipartite graph in a new aspect, namely, to find not all, but only the join-irreducible maximum independent sets. Also an algorithm providing these sets is presented, the complexity of which doesn't exceed the complexity of the best algorithm providing just one maximum independent set.

1. Introduction. In this section we present the required preliminaries that can be found e.g. at [1] and [2]. An *independent set* of a graph G is a set of its vertices no two of which are adjacent. Sets with the maximum cardinality are *maximum independent sets* of G , and their cardinality is denoted by $\alpha(G)$. The *Maximum independent set problem* is to find a maximum independent set of the given graph, while the *Maximum independent sets generation problem* is to report all maximum independent sets of the given graph. A *vertex cover* of G is a set of its vertices incident to all its edges. Note that the complement of a vertex cover is an independent set and vice-versa, so the complements of the maximum independent sets are the vertex covers of the minimal cardinality. These vertex covers are the *minimum vertex covers* of G , and their cardinality is denoted by $\tau(G)$. The *Minimum vertex cover problem* and the *Minimum vertex covers generation problem* are defined like the ones for maximum independent set, and obviously they are equivalent respectively. A *matching* of G is a set of its edges no two of which are

incident. Sets of maximal cardinality are the *maximum matchings* of G , and their cardinality is denoted by $\nu(G)$. It can be observed that if M is a matching of G and C is a vertex cover of G , then each edge of M is covered by a separate vertex of C , so it holds $\nu(G) \leq \tau(G)$. The *Maximum matching problem* is to find a maximum matching of the given graph, while the *Maximum matchings generation problem* is to report all maximum matchings of the given graph. In general case the Maximum independent set problem and the Minimum vertex cover problem are NP-complete, while the Maximum matching problem is solvable in polynomial time. Within the study of these problems, the study of the special case of bipartite graphs is of crucial importance. König's theorem states, that for bipartite graphs it holds $\nu(G) = \tau(G)$. It can be checked, that this formulation is equivalent to a notion that no maximum matching has an edge connecting vertices of the same minimum vertex cover. This yields to the fact that for bipartite graphs given a maximum matching, one can construct a vertex cover in linear time, and vice-versa, which, in its turn, makes the Maximum matching problem, the Minimum vertex cover problem and the Maximum independent set problem equivalent to each other for bipartite graphs. The known algorithms solving the last two problems find a maximum matching first, and then obtain a minimum vertex cover or a maximum independent set. In domain of that algorithms the concept of an *augmenting path* is a key concept. For a matching M of a (not necessarily bipartite) graph G path P is called *M-alternating*, if its edges are alternatingly out of and inside the M ; P is called *M-augmenting*, if it is M -alternating and it starts and ends at vertices unmatched by M . It can be observed, that if P is an M -augmenting path, then $M \Delta P$ (here Δ denotes the symmetric difference of two sets) is a matching of cardinality $|M| + 1$. Berge's theorem states that a matching is maximum if and only if there are no augmenting paths with respect to that matching. Thence many known algorithms construct a maximum matching through the search of augmenting paths. Here is a survey on algorithms solving the Maximum matching problem for the given bipartite graph with n vertices and m edges.

1	$O(nm)$	Hungarian method	1955
2	$O(\sqrt{nm})$	Hopcroft-Karp algorithm	1971
3	$O(n^{2.376} \log n)$	Matrices based algorithm	1981
4	$O(n^{1.5} \sqrt{m/\log n})$	Push-reliable flow based algorithm	1991
5	$O((2 - \log_n m) \sqrt{n} m)$	Graphs compressed representation based algorithm	1991

1.1. Generation problems. As it is mentioned above, the Maximum matching problem is equivalent to the Maximum independent set problem for bipartite graphs, since there is a linear time constructive correspondence between the maximum matchings and the maximum independent sets of a bipartite graph. However, various maximum matchings correspond to the same minimum vertex cover, and the generation problems of matching and vertex cover are not known to be equivalent. It makes sense to consider the notion of complexity of an algorithm solving a generation problem carefully, as in these problems the size of output is usually in order higher than the size of the input. Obviously, each such algorithm requires at least as much time as it is needed to obtain one output, plus the time needed to report the whole output. There is an algorithm [2], that given a maximum matching, enumerates all maximum matchings spending just linear time for each of them, so obvious is an algorithm solving the Maximum matchings generation problem for a bipartite graph with n vertices and m edges in $O(t(m,n) + \text{output size})$ time, where $t(m,n)$ is the time needed to obtain a maximum matching at the beginning. For the Maximum independent sets generation problem there is an $O(t(m,n) + \text{output size})$ algorithm [4] (note that the output sizes are not the same in last two bounds), which obviously is asymptotically the best, if it is required to provide all the maximum independent sets one by one. However, it might not be the case if there are better ways to provide the entire set of all maximum independent sets rather than separately providing each element of it. Next we show that actually there is such way, and present an algorithm providing the set of all maximum independent sets in that way in time $O(t(m,n))$.

2. The lattice of maximum independent sets. Here we refer to some basic properties of lattices which can be found e.g. at [3]. A lattice is a partially ordered set in which any two elements have unique supremum, which is called *join*, and unique infimum, which is called *meet*. A lattice can also be defined as an abstract algebra with join and meet operations which are commutative, associative and absorptive. If these operations are also distributive, then the lattice is called distributive. A distributive lattice can also be defined as a topological space, where the union and intersection of the sets correspond to the join and meet operations of the lattice elements. A join-irreducible element of a lattice is one that is not equal to a join of elements all other than itself. As it follows from Birkhoff's representation theorem, a finite distributive lattice is determined by its join-irreducible elements, in sense, that its each element is uniquely represented as an irreducible join of some of its join-irreducible elements. In this section we show that the set of all maximum

independent sets of bipartite graph forms a distributive lattice with respect to some simple set theoretic operations, and show how to construct the join-irreducible elements of that lattice.

2.1. The lattice. Let $G = (U, V, E)$ be a bipartite graph. For a set of vertices $W \subseteq U \cup V$ we will denote $W_U := W \cap U$ and $W_V := W \cap V$. Let X and Y be two maximum independent sets of G , and let \mathcal{X} denote the set of all maximum independent sets of G , so $X, Y \in \mathcal{X}$. We define $X \vee Y$ and $X \wedge Y$ as follows:

$$X \vee Y := (X_U \cup Y_U) \cup (X_V \cap Y_V), \quad (1)$$

$$X \wedge Y := (X_U \cap Y_U) \cup (X_V \cup Y_V). \quad (2)$$

It is easy to check that $X \vee Y$ and $X \wedge Y$ are independent sets of G . Moreover, they are maximum independent sets. As X and Y are maximum independent sets, we have $|X| = |Y| = \alpha(G)$, and as $X \vee Y$ and $X \wedge Y$ are independent sets, we have $|X \vee Y| \leq \alpha(G)$ and $|X \wedge Y| \leq \alpha(G)$. Further, note, that

$$\begin{aligned} |X \vee Y| + |X \wedge Y| &= |(X \vee Y) \cup (X \wedge Y)| + |(X \vee Y) \cap (X \wedge Y)| = \\ &= |X \cup Y| + |X \cap Y| = |X| + |Y| = 2\alpha(G), \end{aligned}$$

so we get $|X \vee Y| = |X \wedge Y| = \alpha(G)$. Thus (1) and (2) define two binary operations over the set of maximum independent sets of G , that are $\vee : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$ and $\wedge : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{X}$. It holds the following.

Lemma 1. *The triple $(\mathcal{X}, \vee, \wedge)$ is a distributive lattice.*

Proof. Indeed, from (1) and (2) it follows that the set family $\{X_U \mid X \in \mathcal{X}\}$ is closed towards the set union and intersection operations, and thus forms a distributive lattice with respect to them. Further, there is an obvious isomorphism between this lattice and $(\mathcal{X}, \vee, \wedge)$, so the lemma is proved.

Consider the partial order induced in $(\mathcal{X}, \vee, \wedge)$. By definition, for $X, Y \in \mathcal{X}$ we have $X \preceq Y$ if and only if $X = X \wedge Y$ (which in its turn is the case if and only if $Y = X \vee Y$). From (1) and (2) it follows that it is the case if and only if $X_U \subseteq Y_U$ (which in its turn is the case if and only if $X_V \supseteq Y_V$). We denote by \check{X} the top of $(\mathcal{X}, \vee, \wedge)$, that is $\check{X} := \bigvee_{X \in \mathcal{X}} X$. Note that from (1) it follows, that \check{X}_U is the U -part of the union of all maximum independent sets of G , and \check{X}_V is the V -part of the intersection of all maximum independent sets of G . Obviously, also the dual claim holds for the bottom of $(\mathcal{X}, \vee, \wedge)$, which is $\hat{X} := \bigwedge_{X \in \mathcal{X}} X$. As it is mentioned above $(\mathcal{X}, \vee, \wedge)$ is determined by its join-irreducible elements. As \hat{X} is the bottom, then it is one of them. Next we indicate the others. Recall, that $(\mathcal{X}, \vee, \wedge)$ is naturally isomorphic to the finite topology $(\mathcal{X}_U, \cup, \cap)$, where $\mathcal{X}_U := \{X_U \mid X \in \mathcal{X}\}$. The non-bottom join-irreducible elements of a finite topology are the closures of its points (the closure of a point is the

least set containing it). So the closure of $u \in \check{X}_U \setminus \hat{X}_U$ is the least set of $(\mathcal{X}, \cup, \cap)$ containing u , and obviously the isomorphic image of that set in $(\mathcal{X}, \vee, \wedge)$ is the least maximum independent set of G containing u . We call that independent set the *closure maximum independent set* of vertex u . Thus the non-bottom join-irreducible elements of $(\mathcal{X}, \vee, \wedge)$ are the closure maximum independent sets of vertices $u \in \check{X}_U \setminus \hat{X}_U$. The next two lemmas form this claim, and their proof basically explains the proof of the corresponding claim for finite topologies to the terms of lattice $(\mathcal{X}, \vee, \wedge)$.

Lemma 2. *In G , the closure maximum independent set of each vertex $u \in \check{X}_U \setminus \hat{X}_U$ is join-irreducible.*

Proof. Let X be the closure maximum independent set of u , and let $X = Y \vee Z$ for some $Y, Z \in \mathcal{X}$. Since $u \in X$, then from (1) it follows, that $u \in Y$ or $u \in Z$, and as X is the least maximum independent set containing u , then $X \leq Y$ or $X \leq Z$. On the other hand, since $X = Y \vee Z$, we have $X \geq Y$ and $X \geq Z$. Thus we obtain that $X = Y$ or $X = Z$, which means that X is join-irreducible.

Lemma 3. *In G , if X is a join-irreducible maximum independent set, then the closure maximum independent set of each vertex $u \in X_U$ is the X .*

Proof. Let for vertices $u \in X_U$, X_u denote the closure maximum independent set of u . Observe, that from (1) it follows, that $\bigvee_{u \in X_U} X_u \geq X$. On the other hand, for each $u \in X_U$, since X_u is the least maximum independent set containing u , then it holds $X_u \leq X$ for all $u \in X_U$, which means that $\bigvee_{u \in X_U} X_u \leq X$. Thus we obtain $\bigvee_{u \in X_U} X_u = X$. Since X is join-irreducible, then for some $u \in X_U$ it holds $X = X_u$.

Lemma 4. *The join-irreducible maximum independent sets of G are \hat{X} and the closure maximum independent sets of vertices $u \in \check{X}_U \setminus \hat{X}_U$.*

Proof. Immediately follows from Lemma 2 and

Lemma 3.

In this section we have described the set of maximum independent sets of a bipartite graph, and in the next two sections we will show how to maintain that set.

2.2. The least maximum independent set algorithm. As it follows from

Lemma 4, in order to construct a join-irreducible element of $(\mathcal{X}, \vee, \wedge)$, one has to construct the least maximum independent set containing some given vertex of G . Next we discuss a wider problem of finding the least maximum independent set containing the given set of vertices $R \subseteq U \cup V$, if there is one. To do it we first prove some properties.

Lemma 5. *In G , if X is a maximum independent set and M is a maximum matching, then each edge of M is incident to exactly one vertex out of X , and there are no other vertices out of X .*

Proof. Immediately follows from König's theorem. Recall, that the complement of X , which we denote by \bar{X} , is a minimum vertex cover of G , so each edge of M is incident to at least one vertex \bar{X} . But it is incident to exactly one vertex of \bar{X} , and there are no other vertices of \bar{X} , as König's theorem states that $|\bar{X}| = |M|$.

Corollary. *In G , the union of all sets of vertices unmatched by some maximum matching is contained in the intersection of all maximum independent sets.*

Proof. Indeed. If X is some maximum independent set, and $u \in U$ is some vertex not incident to some maximum matching M , then it is also not out of X , since, as it follows from the lemma, all vertices out of X are incident to an edge in M . Thus $u \in X$.

Lemma 6. *In G , if a vertex $u \in U$ is contained in some maximum independent set X , then for any maximum matching M the endpoint of any M -alternating path of even length starting at u is also contained in X .*

Proof. Let M be some maximum matching. Note that it is sufficient to prove the lemma only for an M -alternating path of length two, that is for a path uv_1u_1 , where $(u, v_1) \in E$ and $(u_1, v_1) \in M$, and this case it obvious. Indeed, as X is an independent set, then $v_1 \notin X$, and as $(u_1, v_1) \in M$, then from

Lemma 5 it follows that $u_1 \in X$.

For a maximum matching M of G , we denote by \vec{G}_M the directed graph obtained from G by directing the edges of M from V to U , and the other edges from U to V . We use \vec{G}_M to operate with the M -alternating paths of G . Note that the last are just the directed paths of \vec{G}_M . Now we present an algorithm providing the least maximum independent set containing the given set of vertices $R \subseteq U \cup V$, if there is one, and otherwise reporting that no maximum independent set contains R , given a maximum matching.

Algorithm A

Input: a bipartite graph $G = (U, V, E)$, a maximum matching M of G , a set of vertices $R \subseteq U \cup V$.

Output: if R is contained in some maximum independent set of G , then the least one of them, otherwise, a message indicating that no maximum independent set of G contains R .

Step 1. Denote by W the set of vertices unmatched by M .

Step 2. Construct \vec{G}_M and find all directed paths starting at $W_U \cup R_U$; denote by F the set of their vertices.

Step 3. Denote by \bar{F} the complement of F .

Step 4. If $W_V \cup R_V \not\subseteq \bar{F}_V$, then report that no maximum independent set of G contains R and exit.

Step 5. Provide $F_U \cup \bar{F}_V$ as the least maximum independent set of G containing R .

Now we show that this algorithm is correct.

Lemma 7. *The Algorithm Ais correct, and it works in time $O(n + m)$, where n is the number of vertices and m is the number of edges of the input bipartite graph.*

Proof. The claim regarding complexity is obvious, as the Step 2 can be implemented via a traversal over \vec{G}_M , which takes $O(n + m)$ time, and the other steps can be implemented by a lookup over the vertices of G . Now we prove that the algorithm is correct. As it follows from the corollary of

Lemma 5, after Step 1 it holds that W is contained in all maximum independent sets of G , so if there is one containing R_U , then it also contains W_U . Obviously, after Step 2 it holds that F_U and F_V are respectively the even and the odd vertices of the M -alternating paths starting at $W_U \cup R_U$. Observe that after Step 3, by denotation of \bar{F} , we have that (i) no edge connects a vertex of F_U with a vertex of \bar{F}_V , so $F_U \cup \bar{F}_V$ is an independent set, and $\bar{F}_U \cup F_V$ is a vertex cover; (ii) no edge of M connects a vertex of F_V with a vertex of F_U ; (iii) $W_U \subseteq F_U$, so each vertex of \bar{F}_U is matched by some edge of M . If $W_V \not\subseteq \bar{F}_V$, then there is an edge between F_U and W_V , so, as it follows from the corollary of

Lemma 5, no maximum independent set contains F_U , and as it follows from

Lemma 6, no maximum independent set contains R_U , therefore no maximum independent set contains R , as the algorithm reports at Step 4 in this case. Otherwise, if $W_V \subseteq \bar{F}_V$, then each vertex of F_V is matched by some edge of M , so referring to (ii) and (iii), we get that $|M| = |\bar{F}_U \cup F_V|$, further referring to (i) and König's theorem, we get that $F_U \cup \bar{F}_V$ is a maximum independent set, and moreover, referring to

Lemma 6 and the construction of F , we get that $F_U \cup \bar{F}_V$ is the least maximum independent set containing R_U . The last means that F_V is the union of the V -parts of all maximum independent sets containing R_U , so if $R_V \not\subseteq \bar{F}_V$, then no maximum independent set contains $R = R_U \cup R_V$, as the algorithm reports at Step 4 in this case. Otherwise, if $R_V \subseteq \bar{F}_V$, then obviously $F_U \cup F_V$ is the least maximum independent set containing R , which the algorithm provides at Step 5.

Corollary. *In G , the union of all sets of vertices unmatched by some maximum matching coincides with the intersection of all maximum independent sets.*

2.3. The join-irreducible maximum independent sets algorithm.As it follows from

Lemma 4, the join-irreducible maximum independent sets of G are $\hat{X} \cup \{X_u \mid u \in \check{X}_U \setminus \hat{X}_U\}$, where X_u is the least maximum independent set containing u . For any maximum matching M of G it holds that \hat{X} is the least maximum independent set containing a vertex unmatched by M , so one can obtain a maximum matching, execute Algorithm A for each vertex $u \in U$, thus obtaining the join-irreducible maximum independent sets. Note that such approach may obtain the same join-irreducible maximum independent set twice, and that it doesn't indicate the partial order between them. Next we provide an algorithm which doesn't have these disadvantages. For simplicity we assume that G is connected.

Let M be a maximum matching of G , and let W denote the set of vertices unmatched by M . Denote by \mathcal{H} the set of strong components of \vec{G}_M . Obviously some $L \in \mathcal{H}$ contains W_U and some $R \in \mathcal{H}$ contains W_V . Let \preceq denote the partial order over \mathcal{H} induced by the arcs of \vec{G}_M between its strong components, so for a pair of strong components \preceq indicates whether there is a path of \vec{G}_M from second to the first. It can be checked that L and R are correspondingly the bottom and the top of (\mathcal{H}, \preceq) . It also can be checked that \mathcal{H} , as well as L and R , doesn't depend on M , and that (\mathcal{H}, \preceq) except L and R , corresponds to the Dulmage-Mendelsohn decomposition [2] of G . Now let $H \in \mathcal{H} \setminus \{L, R\}$ be a strong component, and let H^* be the union of all strong components preceding or equal to H .

Lemma 8. For a vertex $u \in H_U$ it holds that $H_U^* \cup H_V^*$ is the least maximum independent set containing u .

Proof. From the definition of H^* it follows that in $\vec{G}_M u$ is connected with all vertices in H^* , and that it is not connected with any vertex in $\overline{H^*}$, so the proof immediately follows from Algorithm A and Lemma 7. Finally we provide the algorithm obtaining the partially ordered set of the join-irreducible maximum independent sets of the given bipartite graph.

Algorithm B

Input: a bipartite graph $G = (U, V, E)$.

Output: the partially ordered set of the join-irreducible maximum independent sets of G .

Step 1. Obtain a maximum matching M of G .

Step 2. Construct \vec{G}_M .

Step 3. Construct (\mathcal{H}, \preceq) , and indicate L and R .

Step 4. Denote $X_L := L_U \cup \overline{L_V}$, and provide it as the bottom join-irreducible maximum independent set.

Step 5. Iterate $\mathcal{H} \setminus \{L, R\}$ in a non-decreasing order, and for each H in it do Step 6.

Step 6. Denote $X_H := H_U^* \cup \overline{H_V^*}$ and provide it as the join-irreducible maximum independent set succeeding all already provided $X_{H'}$ -s, such that $H' \preceq H$.

Now we show that this algorithm is correct.

Lemma 9. The Algorithm B is correct, and it works in time $O(t(n, m))$, where n is the number of vertices and m is the number of edges of the input bipartite graph, and $t(n, m)$ is the complexity of an algorithm finding a maximum matching of the given bipartite graph.

Proof. The claim regarding complexity is obvious, as the strong components can be obtained in linear time [2]. Next, from

Lemma 4 and

Lemma 8 it follows that Algorithm B provides all the join-irreducible maximum independent set of G , and from the definitions of $(\mathcal{X}, \vee, \wedge)$ and (\mathcal{H}, \preceq) it follows that at Step 6, if $H' \preceq H$, then it holds $X_{H'} \preceq X_H$. Thus the lemma is proved.

3. Conclusion. We have studied the underlying structure of maximum independent sets of a bipartite graph and we have shown that it is a distributive lattice with respect to join and meet operations defined by (1) and (2). The join-irreducible maximum independent sets describe the entire set of the maximum independent sets, in sense, that each maximum independent set is uniquely represented by an irreducible join of some of the join-irreducible maximum independent sets. This faced allows to view the Maximum independent sets generation problem in the aspect of obtaining the join-irreducible maximum independent sets rather than reporting all the maximum independent sets. Algorithm B obtains them in time needed to obtain just one maximum independent set.

Institute for Informatics and Automation Problems of NAS RA

V. G. Minasyan

On Set of All Maximum Independent Sets of Bipartite Graph

It is shown that the set of all maximum independent sets of bipartite graph is a distributive lattice, which allows to view the problem of generating the maximum independent sets of bipartite graph in a new aspect, namely, to find not all, but only the join-irreducible maximum independent sets. Also an algorithm providing these sets is presented, the complexity of which doesn't exceed the complexity of the best algorithm providing just one maximum independent set.

Վ. Գ. Մինասյան

Երկկողմանի գրաֆի բոլոր մաքսիմալ անկախ բազմությունների բազմության մասին

Ցույց է տրվում, որ երկկողմանի գրաֆի մաքսիմալ անկախ բազմությունների բազմությունը բաշխական կավար է, ինչը թույլ է տալիս երկկողմանի գրաֆի մաքսիմալ անկախ բազմությունների գեներացման խնդիրը դիտել որոշակիորեն նոր ասպեկտում, այն է՝ գտնել ոչ թե բոլոր, այլ միայն միավորմամբ անբաղադրելի մաքսիմալ անկախ բազմությունները: Նաև ներկայացվում է այդ բազմությունները տրամադրող ալգորիթմ, որի բարդությունը ավելին չէ, քան միայն մեկ մաքսիմալ անկախ բազմություն տրամադրող լավագույն ալգորիթմի բարդությունը:

В. Минасян

О множестве всех максимальных независимых множеств двудольного графа

Показано, что множество всех максимальных независимых множеств двудольного графа есть дистрибутивная решетка, что позволяет рассматривать задачу генерации максимальных независимых множеств двудольного графа в некотором новом аспекте, а именно, найти не все, а только неразложимые в объединение максимальные независимые множества. Также приведен алгоритм, предоставляющий эти множества, сложность которого не больше сложности наилучшего алгоритма, предоставляющего только одно максимальное независимое множество.

References

1. *Schrijver A.*– Combinatorial optimization: polyhedra and efficiency, volume A. Springer. 2003.
2. *Ловас Л., Пламмер М.* – Прикладные задачи теории графов: теория паросочетаний в математике, физике, химии. М. Мир. 1998.
3. *Буркгоф Г.* – Теория структур. М. ИЛ. 1952.
4. *Kashiwabara T., Masuda S., Nakajima K., Fujisawa T.*–Journal of Algorithms, V.13.N 1. 1992. P. 161-174.
5. *Karp R., Vazirani U., Vazirani V.*– STOC '90 Proceedings of the twenty-second annual ACM symposium on Theory of computing. 1990. P. 352-358.
6. *Lovász L.* –Journal of Combinatorial Theory, Series B. V. 43. N. 2. 1987. P. 187–222.
7. *Alt H., Blum N., Mehlhorn K., Paul M.*–Information Processing Letters. V. 37.N 4. 1991. P. 237-240.

8. *Feder T., Motwani R.*—STOC '91 Proceedings of the twenty-third annual ACM symposium on Theory of computing. 1991. P. 123-133.
9. *Ibarra O., Moran S.* —Information Processing Letters, V. 13. Issue 1. 1981. P. 12-15.