ELSEVIER

Discrete Applied Mathematics 59 (1995) 285–292

DISCRETE APPLIED MATHEMATICS

Note

How do I marry thee? Let me count the ways

Arthur T. Benjamin[a,*], Cherlyn Converse[b], Henry A. Krieger[a]

[a] Mathematics Department, Harvey Mudd College, Claremont, CA 91711, USA
[b] Mathematics Department, California State University, Fullerton, CA 92634, USA

Received 4 October 1993; revised 3 August 1994

Abstract

A stable marriage problem of size $2n$ is constructed which contains $\Theta(2^n\sqrt{n})$ stable matchings. This construction provides a new lower bound on the maximum number of stable matchings for problems of even size and is comparable to a known lower bound when the size is a power of 2. The method of construction makes use of special properties of the latin marriage problem, which we develop.

1. Introduction

The stable marriage problem consists of n men and n women who are to be matched up into married couples. Each man ranks the women from most desirable to least desirable, and each woman does the same for the men. A matching is said to be *unstable* if there exists a man and woman who prefer each other to the partners they have been assigned. If no such people exist, the matching is said to be *stable*. In [2], Gale and Shapley proved that a stable matching always exists, but it need not be unique. The problem of determing the maximum number of possible stable matchings among all stable marriage problems of size n was posed by Knuth [5] and remains an open question. As reported in [3], Knuth established that this maximum number exceeds $2^{n/2}$ for $n > 1$. When n is a power of 2, Gusfield and Irving established that this maximum number is at least 2^{n-1}, which can be improved to $(2.28)^n/(1+\sqrt{3})$ based on a construction by Irving and Leather [4]. In this article we construct, for all even values of n, a stable marriage problem such that the number of stable matchings lies between these two lower bounds.

*Corresponding author.

SSDI 0166-218X(94)00129-4

$$\begin{bmatrix} 0 & 2 & \boxed{4} & 1 & 3 & 5 \\ \boxed{2} & 4 & 0 & 5 & 1 & 3 \\ 4 & 0 & 2 & \boxed{3} & 5 & 1 \\ 1 & 3 & 5 & 0 & \boxed{2} & 4 \\ 5 & \boxed{1} & 3 & 2 & 4 & 0 \\ 3 & 5 & 1 & 4 & 0 & \boxed{2} \end{bmatrix}$$

Fig. 1. An unstable matching for DS_3.

2. Latin marriages

We call the stable marriage problems constructed here *latin marriages* since they can be described by a latin square, an $n \times n$ matrix where every row and column is a permutation of the numbers $0, 1, \ldots, n-1$. For a given latin square A with (i,j) entry a_{ij}, $0 \leqslant i, j \leqslant n-1$, we shall interpret a_{ij} to be man i's rank of woman j, and $n-1-a_{ij}$ to be woman j's rank of man i, where 0 is the best rank and $n-1$ is the worst rank. A matching on A can be described by a sequence $X_0, X_1, \ldots, X_{n-1}$, where X_j denotes the number selected in column j. To avoid polygamy, no two selected numbers may lie in the same row. If $X_j = a_{ij}$, we say a_{ij} is the *selected cell* of column j (as well as the selected cell of row i). Alternatively, a matching can also be described by a sequence $X^0, \ldots, X^{n-1}$, the numbers selected for each row. For example, the matching presented in Fig. 1 (described by column sequence 2, 1, 4, 3, 2, 2 and row sequence 4, 2, 3, 2, 1, 2) is unstable since man 0 prefers woman 1 over his assigned partner (woman 2), and woman 1 prefers man 0 over her current partner (man 4).

Lemma 1. *A matching on a latin square A with column selections $X_0, \ldots, X_{n-1}$ and row selections $X^0, \ldots, X^{n-1}$ is unstable if and only if $X^i > a_{ij} > X_j$ for some $i, j \in \{0, \ldots, n-1\}$.*

Proof. By our interpretation of A, $X^i > a_{ij} > X_j$ means that man i prefers woman j to his assigned partner, and woman j prefers man i to her assigned partner. □

Lemma 2. *A matching on A is unstable if and only if $X^i < a_{ij} < X_j$ for some $i, j \in \{0, \ldots, n-1\}$.*

Proof. By the previous lemma, it suffices to show that $X^i < a_{ij} < X_j$ if and only if there exist $k, l \in \{0, \ldots, n-1\}$ satisfying $X^k > a_{kl} > X_l$. Suppose that $X^i < a_{ij} < X_j$, where a_{ij} is equal to the number x. Now create a bipartite multigraph G with nodes $r_0, \ldots, r_{n-1}$ and $c_0, \ldots, c_{n-1}$ as follows. Label each node r_s and c_t with X^s and X_t, respectively. For each selected cell a_{st}, draw a red edge between r_s and c_t. For each $a_{st} = x$, draw a blue edge between r_s and c_t. In the resulting graph, each node has exactly one red edge and one blue edge leaving it, with the red edges connecting nodes with the same label. The connected components of this graph are alternating red–blue cycles. By assumption, there exists a blue edge, and not a red edge, from r_i to c_j. Traverse the cycle beginning with this blue edge until we first encounter a node c_l with label $X_l < x$. This node was entered by a blue edge from a node r_k with label $X^k > x > X_l$. Thus $X^k > a_{kl} > X_l$. The converse can be proven in exactly the same way. □

Combining our two lemmas we have the following *stability condition*:

Theorem 1. *A matching on A is stable if and only if there do not exist* $i, j \in \{0, \dots, n-1\}$ *such that* $X^i > a_{ij} > X_j$ *or* $X^i < a_{ij} < X_j$.

Consequently, whereas instances of the general stable marriage problem can have as few as one stable matching, size n latin marriage problems must possess at least n stable matchings, namely the *constant* matchings where every man receives his mth choice for some m.

In fact, the matrix S_n defined to have (i,j) entry $s_{ij} = (i+j) \bmod n$ achieves this minimum.

Lemma 3. S_n *has only n stable matchings, the constant ones.*

Proof. Suppose to the contrary that $X_0, \dots, X_{n-1}$ is a non-constant stable matching on S_n. Let $m = \min\{X_0, \dots, X_{n-1}\} < n-1$. Then there exists j such that $X_j = s_{ij} = m$ and $X_{j+1} \neq m$ (where X_n is defined to be X_0). Now $s_{i,j+1} = m+1$ is in the same row as selected cell s_{ij} and therefore cannot be selected in column $j+1$. Thus, $X_{j+1} > m+1$. But this produces a stability violation since $X^i = m < m+1 = s_{i,j+1} < X_{j+1}$. □

3. DS_n and valid sequences

We now introduce DS_n, a $2n \times 2n$ latin square possessing many stable matchings. DS_n is built up from S_n as follows. For $0 \leqslant i, j \leqslant n-1$, we define $a_{ij} = a_{i+n,j+n} = 2s_{ij}$. The other quadrants are defined by the *vertical reflection* relation $a_{ij} + a_{i,2n-1-j} = 2n-1$. Fig. 1 contains an example of DS_3. Algebraically, the (i,j) entry of DS_n is equal to

$$a_{ij} = \begin{cases} 2(i+j) \bmod 2n & \text{if } 0 \leqslant i,j \leqslant n-1 \text{ or } n \leqslant i,j \leqslant 2n-1, \\ 2(j-i)+1 \bmod 2n & \text{otherwise.} \end{cases}$$

It is easy to see that DS_n is indeed a latin square, where opposite quadrants are identical and have elements of the same parity. Furthermore, for $j \neq 0$ or n, $a_{ij} = (a_{i,j-1} + 2) \bmod 2n$. Also $a_{i0} = (a_{i,n-1} + 2) \bmod 2n$ and $a_{in} = (a_{i,2n-1} + 2) \bmod 2n$.

The remaining crucial property of DS_n that we exploit follows immediately from our algebraic description, and is called the *rectangle property*:

Lemma 4. *If* (i, j) *and* (k, l) *are in opposite quadrants of* DS_n *and* $a_{ij} = a_{kl}$ *then* $a_{il} = a_{kj}$.

It may seem unusual to base a stable marriage problem construction of size $2n$ on a latin marriage problem of size n with the minimum number of stable matchings. Nevertheless, Converse [1] empirically demonstrates that this construction outperforms other seemingly more natural constructions.

To characterize and count the stable matchings of DS_n, we define a sequence of numbers $X_0, \ldots, X_k$ to be a *valid sequence* if for $j = 0, \ldots, k-1$, $0 \leqslant X_{j+1} \leqslant X_j + 1$.

Lemma 5. *If* $X_0, \ldots, X_{2n-1}$ *is a stable matching for* DS_n, *then the sequences* $X_0, X_1, \ldots, X_{n-1}, X_0$ *and* $X_n, X_{n+1}, \ldots, X_{2n-1}, X_n$ *are valid sequences.*

Proof. For $j \neq n-1, 2n-1$, if $X_j = a_{ij} = X^i$, then $X_{j+1} \neq a_{i,j+1} = X_j + 2$, provided that $X_j \leqslant 2n-3$. (If $X_j \geqslant 2n-2$, then $X_{j+1} \leqslant X_j + 1$ is automatic.) Further, if $X_{j+1} > X_j + 2$, then $X^i = a_{ij} < a_{i,j+1} < X_{j+1}$ violates the stability condition. Hence, $X_{j+1} \leqslant X_j + 1$. In the same way, it can be shown that $X_0 \leqslant X_{n-1} + 1$ and $X_n \leqslant X_{2n-1} + 1$. □

The next theorem states that if we can select cells from the first n columns in a valid way, then there is exactly one way to select cells from the remaining n columns to produce a stable matching.

Theorem 2. *For every valid sequence* $X_0, X_1, \ldots, X_{n-1}, X_0$, *where* $0 \leqslant X_j \leqslant 2n-1$ *for all* j, *there exists exactly one stable matching* $X_0, \ldots, X_{2n-1}$ *in* DS_n.

Proof (By induction on $\sum_{j=0}^{n-1} X_j$). First we prove the theorem for our base case where $X_j = 0$ for $j = 0, \ldots, n-1$. Thus $X_0 = 0 = X^i$ for some $i \leqslant n-1$. Hence $a_{il} = 1$ for some $l \geqslant n$. Hence, $X_l \neq 1$. If $X_l > 1$ then we have an instability $X^i = 0 < 1 = a_{il} < X_l$. Therefore X_l must be 0. Each subsequent $X_j = 0$, $j \leqslant n-1$ forces a different $X_l = 0$, $l \geqslant n$. Hence the only stable matching whose first n terms are 0 is the constant matching.

Next we observe that for any valid sequence, not all zero, there exists at least one element in the sequence (any maximum one, for instance) that we can diminish by one and still have a valid sequence. In this way, we can construct a "continuous" path of valid sequences from any given valid sequence to the zero sequence. Inductively, suppose there exists a unique stable matching M associated with the valid sequence $X_0, X_1, \ldots, X_{n-1}, X_0$. Suppose further that changing X_j from m to $m+1$ yields a new valid sequence. We shall prove that there exists a unique stable matching M' associated with the new valid sequence.

Since this is a legal change, it must be true that $X_{j-1} \geqslant m$, for $j \geqslant 1$. (If $j = 0$, $X_{n-1} \geqslant m$.) Now suppose that $a_{ij} = m$, and $a_{kj} = m + 1$. Now in our stable matching M, we must have $X^k \leqslant m$ since $X^k \neq m + 1$ and if $X^k > m + 1$ we would have an instability $X_j = m < m + 1 = a_{kj} < X^k$. In fact we must have $X^k = m$. For $X^k \neq m - 1 = a_{k,j-1}$ since $X_{j-1} \geqslant m$. Further, if $X^k < m - 1$, then we would have an instability $X^k < a_{k,j-1} < X_{j-1}$. Thus, $X^k = m = a_{k,l}$ for some column $l \geqslant n$. Since $a_{ij} = m = a_{kl}$, we have by our rectangle property that $a_{il} = a_{kj} = m + 1$. Thus by selecting $X_j = m + 1 = X^k$ and $X_l = m + 1 = X^i$ we have a matching M' on DS_n that is consistent with the new valid sequence.

We claim that this new matching must also be stable. For suppose, to the contrary, that an instability was caused by a_{xy}. Then x must equal i or k or y must equal j or l (otherwise this instability would exist in the previous matching). Since a_{xy} has $m + 1$ selected in its row or column, $a_{xy} \neq m + 1$. Also, $a_{xy} \neq m$, since by our construction a_{xy} would have $m + 1$ selected in its row and column and thus not cause an instability. But since a_{xy} is not equal to m or $m + 1$, any instability involving a_{xy} caused by its row or column selection being equal to $m + 1$ would have existed in the previous matching when that selection was m, contradicting the stability of the previous matching. Hence the new matching is also stable.

As for uniqueness, if two stable matchings are consistent with our new sequence, then arguing as before, legally changing X_j from $m + 1$ to m necessarily implies that $X_l = m + 1$ in both matchings, and by the induction hypothesis, when X_j and X_l are reassigned m, both matchings are the same. Hence the two matchings must have been the same to begin with. □

Corollary 1. *The number of stable matchings in DS_n is equal to the number of valid sequence $X_0, \ldots, X_n$ where $X_n = X_0$ and $0 \leqslant X_j \leqslant 2n - 1$ for all j.*

4. Counting valid sequences

Lemma 6. *For any integer $0 \leqslant k \leqslant 2n - 1$, the number of valid sequences $X_0, \ldots, X_n$ where $X_n = X_0 = k$ and $0 \leqslant X_j \leqslant 2n - 1$ for all j is equal to the number of valid sequences $Y_0, \ldots, Y_n$ where $Y_n = Y_0 = 2n - 1 - k$ and $0 \leqslant Y_j \leqslant 2n - 1$ for all j.*

Proof. Defining $Y_j = 2n - 1 - X_{(2n-1-j)}$ for all j gives us a valid sequence for the desired one to one correspondence. □

Note that in order for a sequence to be valid we must have $X_j \leqslant X_0 + j$ for all j. Hence if $X_0 \leqslant n - 1$ the condition $0 \leqslant X_j \leqslant 2n - 1$ is automatically satisfied since our sequences have length n.

Thus if we define $I(2n)$ to be the number of stable matchings of DS_n, and for $0 \leqslant k \leqslant n - 1$, $h(n, k)$ to be the number of valid sequences $X_0, \ldots, X_n$ where $X_n = X_0 = k$, we have

Corollary 2.

$$I(2n) = 2\sum_{k=0}^{n-1} h(n,k).$$

For any valid sequence $X_0, \dots, X_n$ with $X_n = X_0 = k$, if we define $1 - u_j = X_j - X_{j-1}$, $j = 1, \dots, n$, we obtain the following equivalent counting problem.

Lemma 7. *$h(n,k)$ also counts the number of non-negative integer solutions to $\sum_{j=1}^{n} u_j = n$, with the restriction that for $j = 1, \dots, n$, $\sum_{i=1}^{j} u_i \leqslant k + j$.*

We can interpret $u_1, \dots, u_n$ as unique instructions to a random walk from $(0,0)$ to $(2n-1, 1)$ (i.e., a sequence $(0, Y_0), (1, Y_1), \dots, (2n-1, Y_{2n-1})$ with $Y_0 = 0$, $Y_{2n-1} = 1$, and $|Y_j - Y_{j-1}| = 1$ for all j) which takes n steps up and $n-1$ steps down, where for $j = 1, \dots, n-1$, u_j denotes the number of consecutive up steps preceding the jth down step. u_n is the number of up steps following the last down step. The inequality constraint, rewritten as $\sum_{i=1}^{j} u_i - (j-1) \leqslant k+1$, eliminates precisely those random walks with $Y_j \geqslant k+2$ for some j.

We can enumerate the walks which satisfy $Y_j \geqslant k+2$ for some j by reflecting along the line $Y = k+2$ after the first point of intersection (that is, if $Y_i < k+2$ for $i < j$, and $Y_j = k+2$, then we replace Y_l with $2(k+2) - Y_l$ for $l > j$). This reduces the problem to enumerating random walks from $(0,0)$ to $(2n-1, 2k+3)$ with $n+1+k$ steps up and $n-2-k$ steps down. In summary, we have

Lemma 8.

$$h(n,k) = \binom{2n-1}{n-1} - \binom{2n-1}{n-2-k}.$$

We note that when $k = 0$, $h(n,k)$ simplifies to $(1/(n+1))\binom{2n}{n}$, the nth Catalan number. In fact it was the presence of these numbers appearing in constructions of DS_n that encouraged us to believe that a closed form for $I(2n)$ existed. In fact, we originally used generating functions and determined $h(n,k)$ to be the x^n coefficient of $(1 - x^{k+2}(C(x))^{2k+4})/(1 - x(C(x))^2)$ where $C(x) = (1 - \sqrt{1-4x})/2x$ is the Catalan number generating function. Finally, we prove

Theorem 3. $I(2n) = (n+1)\binom{2n}{n} - 2^{2n-1}$.

Proof.

$$I(2n) = 2\sum_{k=0}^{n-1} h(n,k) = 2n\binom{2n-1}{n-1} - 2\sum_{k=0}^{n-1}\binom{2n-1}{n-2-k}$$
$$= n\binom{2n}{n} - \left[\sum_{i=0}^{2n-1}\binom{2n-1}{i} - \left(\binom{2n-1}{n-1} + \binom{2n-1}{n}\right)\right]$$
$$= (n+1)\binom{2n}{n} - 2^{2n-1}. \qquad \square$$

Using Stirling's approximation, we obtain

$$I(2n) \sim 2^{2n}\left[\frac{\sqrt{n}(1+1/n)}{\sqrt{\pi}} - \frac{1}{2}\right].$$

Equivalently, when n is even,

$$I(n) \sim 2^{n}\left[\frac{\sqrt{n}(1+2/n)}{\sqrt{2\pi}} - \frac{1}{2}\right]$$

and is bounded below by

$$2^{n}\left[\frac{1}{\sqrt{2\pi}}\sqrt{n}\left(1+\frac{2}{n}\right)\left(1-\frac{1}{4n}\left(1+\frac{1}{36n}\right)\right) - \frac{1}{2}\right].$$

5. Extensions

We note that [3] also suggests a procedure for creating stable marriage problems when n is composite. Specifically, if $n = n_1 n_2$ then from a marriage problem of size n_1 with x_1 stable matchings and a marriage problem of size n_2 with x_2 stable matchings, they create two marriage problems of size n, one with at least $n_2 x_1^{n_2}$ and the other with at least $n_1 x_2^{n_1}$ stable matchings. For example, starting with the best marriage problems of size 2 and 3 (which happen to be latin marriage problems S_2 and S_3), then by appropriately "duplicating" S_3 and "triplicating" S_2, this procedure generates two latin marriage problems of size 6 with exactly 18 and 24 stable matchings, respectively. Our construction is different, since DS_3 produces 48 stable matchings.

In [1], Converse proposes a more sophisticated latin square construction when the size of the problem is a multiple of 4, but not a multiple of 8. The matrix D^2S_n has $4n$ rows and columns, divided into four quadrants. As before, if we denote the (i,j) entry of DS_n and D^2S_n by a_{ij} and $a_{ij}^{(2)}$, respectively, we have for $0 \leqslant i,\ j < 2n$, $a_{ij}^{(2)} = a_{i+2n,\,j+2n}^{(2)} = 2a_{ij}$, and when i or j exceeds $2n-1$, $a_{ij}^{(2)} = (4n-1) - a_{i,4n-1-j}^{(2)}$. An analogous process is used to create marriage problems divisible by higher powers of two. Numerically (for matrices of even size $n \leqslant 26$) the more complicated construction possesses more stable matchings than simpler constructions of the same size (e.g., if n is a multiple of 8, $D^3S_{n/8}$ has more stable matchings than $D^2S_{n/4}$ which has more stable matchings than $DS_{n/2}$). In fact, when n is a power of two, this procedure produces the same stable marriage problem as the one in [4] reported in [3] to possess at least $(2.28)^n/(1+\sqrt{3})$ stable matchings. We refrain from conjecturing that our constructions produce the maximum number of stable matchings since our size 26 construction DS_{13} has fewer stable matchings than D^3S_3, our size 24 construction which can be extended to a size 26 construction with the same number of stable matchings.

Acknowledgements

We would like to thank our referees for many valuable suggestions.

References

[1] C. Converse, Lower bounds for the maximum number of stable pairings for the general marriage problem based on the latin marriage problem, Ph. D. Thesis, Claremont Graduate School, Claremont, CA (1992).

[2] D. Gale and L.S. Shapley, College admissions and the stability of marriage, Amer. Math. Monthly 69 (1962) 9–15.

[3] D. Gusfield and R.W. Irving, The Stable Marriage Problem: Structure and Algorithms, Foundations of Computing Series (MIT Press, Cambridge, 1989).

[4] R.W. Irving and P. Leather, The complexity of counting stable marriages, SIAM J. Comput. 15 (1986) 655–667.

[5] D.E. Knuth, Mariages Stables (Les Presses de l'Université de Montréal, Montreal, 1976).