HAPPINESS LEVELS IN RANDOM MATCHINGS

By

Hamilton Scott

April 2008

An Honors Thesis presented

to the faculty of the Department of Mathematics

EAST TENNESSEE STATE UNIVERSITY

In Partial Fulfillment of the requirements

for the Honors–In–Discipline Program for a

BACHELOR OF SCIENCE

in MATHEMATICS

Approved:

____

Anant Godbole, Ph.D.
Thesis Adviser

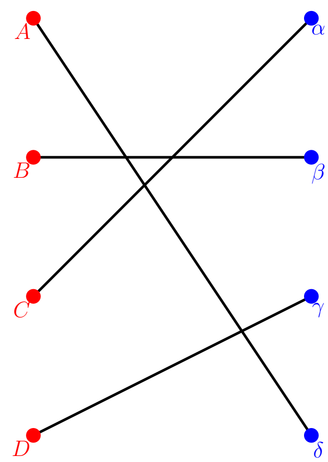
CONTENTS

Abstract

Consider a group consisting of an equal number of men and women. Every man makes a preference list ranking each woman according to whom he likes best and each woman makes a similar preference list of the men. In our analysis, we randomly match men and women and gauge their happiness based on their preference lists. A woman's happiness level is based on her partner's place in her preference list, as is a man's happiness level. We limit the women's lists and the men's lists to a Latin square structure. In other words, no two women can prefer the same man in the same place of their lists. Likewise, no two men can prefer the same woman in the same place of their lists. We also show the minimum and maximum happiness levels and the feasibility of values between the minimum and maximum. We also look at the distributions of happiness levels for 3×3, 4×4, and 5×5 preference matrices.

1. Terminology

Let A be a set of n women and B be a set of n men, a *matching* is a bijection from A to B. In other words, a matching is n monogamous relationships between n men and n women.[4] The figure below is a graphical representation of a matching:



Every man makes a list ranking each woman according to whom he likes best and every woman makes a similar list of the men. These are called *preference lists*. The preference lists are grouped into two $n \times n$ matrices, one for the men and one for the women, which can be called *preference matrices*. We are only allowing the preference matrices to be structured as Latin squares. A *Latin square* is an $n \times n$ matrix that consists of a set of entries $S = \{1, 2 \ldots n\}$ such that each element of S occurs exactly once in each row and column. The set of all Latin squares of order n is Λ_n. We can say that the preference matrices of the men and women are elements of Λ_n. A permutation chosen from S_n matches the men and women.

The following figure is an example of the preference matrices for a set of men $\{\alpha, \beta, \gamma, \delta\}$ and a set of women $\{A, B, C, D\}$.

	1	2	3	4
A	α	β	γ	δ
B	β	γ	δ	α
C	γ	δ	α	β
D	δ	α	β	γ

Women's Lists

	1	2	3	4
α	C	D	A	B
β	D	A	B	A
γ	A	B	C	D
δ	B	C	D	A

Men's Lists

We can also represent preference matrices as illustrated below.

	α	β	γ	δ
A	1	2	3	4
B	2	3	4	1
C	3	4	1	2
D	4	1	2	3

Women's Lists

	A	B	C	D
α	3	4	1	2
β	4	1	2	3
γ	1	2	3	4
δ	2	3	4	1

Men's Lists

If we choose a permutation $\Pi_{A,B,C,D}$ from S_4, say $(\gamma, \beta, \delta, \alpha)$, the resulting matching is displayed below.

	α	β	γ	δ
A	1	2	**3**	4
B	2	**3**	4	1
C	3	4	1	**2**
D	**4**	1	2	3

Women's Lists

	A	B	C	D
α	3	4	1	**2**
β	4	**1**	2	3
γ	**1**	2	3	4
δ	2	3	**4**	1

Men's Lists

A man is happier if he is paired with a woman high on his preference list. Likewise, a woman is happier if she is paired with a man high on her preference list. The *happiness level* for a person is the numerical value of the placement in the preference list of the person with which they are matched. If a man is paired with the second woman on his preference list, then his happiness ranking is 2. If he is paired with the fourth woman on his preference list, then his happiness level is 4. A happiness level of 2 is better than a happiness level of 4. A lower happiness level is better than a higher happiness level.

Both the men's and women's preference matrices will have a happiness level. The happiness level for the men's preference matrix is the sum of every man's happiness level. In the above example, the happiness level for the men's preference matrix is $2 + 1 + 1 + 4 = 8$. Similarly, the happiness level for the women's preference matrix is $3 + 3 + 2 + 4 = 12$.

The *total happiness level* is the sum of every person's happiness level. In the above example the total happiness level is $3+3+2+4+2+1+1+4 = 20$. The happiest matching is the one whose sum is $2n$. The most miserable matching is the one whose sum is $2n^2$.

2. Objective

Our objective is to analyze the matchings of every combination of men's preference lists, women's preference lists, and permutations. The total number of matchings for a given n is $|\Lambda_n|^2 n!$. As can be seen from the following table, the number of matchings grows very quickly as n increases.

n	# of L.S.[7]	# of matchings
3	12	$12^2 3!$
4	576	$576^2 4!$
5	161280	$161280^2 5!$
6	812851200	$812851200^2 6!$
7	61479419904000	$61479419904000^2 7!$
8	108776032459082956800	$108776032459082956800^2 8!$
$\vdots$	$\vdots$	$\vdots$

We will look at the matchings for $n = 3, 4$ and try to generalize our observations for all n. When we have looked at all possible combinations we will analyze the happiness levels of each permutation looking for a patterns. In addition, we will examine the feasible values for a preference matrix. We will also examine the distribution of the happiness values for an individual preference matrix. Using computer programs, we will be able to determine the distributions for $n = 3, 4, 5$.

3. Baseline Results

Proposition 1. *The minimum total happiness level is equal to $2n$.*

Proof. The most optimal matching is when every person gets their first choice. Therefore, if every person gets their first choice, the happiness level will be 1 for every person. Since there are $2n$ people, the total happiness level will be $2n \cdot 1 = 2n$. $\square$

Proposition 2. *The maximum total happiness level is equal to $2n^2$.*

Proof. The least optimal matching is when every person gets their last choice. Therefore, if every person gets their last choice, the happiness level will be n for every person. Since there are $2n$ people, the total happiness level will be $2n \cdot n = 2n^2$. $\square$

Proposition 3. *The distribution of happiness levels for a given n is symmetric.*

Proof. Note that each total happiness level must fall in one of the three classes: below the median, the median, above the median, where the median is $(2n^2 + 2n)/2$. The symmetry is defined by a bijective mapping from the set of happiness levels below the median to the set of happiness levels above the median. We can define this mapping with the following. For each preference list that leads to a total happiness level less than $(2n^2 + 2n)/2$, we replace each preference value p by the value $n - p$. $\square$

4. Feasible Values

Recall that we are only allowing individual preference matrices to be of the form of a Latin square. For an individual preference matrix, the minimum happiness level is n and the maximum is n^2. Our goal is to determine the feasible values for happiness levels between n and n^2 for an individual preference matrix. We first look at the feasible values for 3×3 and 4×4 Latin squares and then look at the general case.

4.1 Identity Permutation

The *identity permutation* is defined by $\Pi_{A,B,C,D,\ldots} = (\alpha, \beta, \gamma, \ldots)$. On a preference matrix that has the identity permutation, the match of each person is on the diagonal. In the following figure, the dots along the diagonal represent the matching; the person in the row and column of a dot are matched.

$$\begin{array}{c|cccc} & \alpha & \beta & \gamma & \cdots \\ \hline A & \bullet & & & \\ B & & \bullet & & \\ C & & & \bullet & \\ \vdots & & & & \ddots \end{array}$$

Any matching on an individual preference matrix can be arranged so that the match of each person is on the diagonal. Note that this is because columns and rows of a Latin square can be swapped and the resulting matrix will still be a Latin square. Therefore, we only have to evaluate the values of the sum of the diagonal to determine all of the possible happiness levels of an individual preference matrix. An example of how an individual preference matrix with a random matching can be manipulated to look like a matrix with the identity permutation is displayed below. In this example, A's column is moved to be the third column, D's column is moved to be the first column and C's column is moved to be the forth column.

	A	B	C	D
α	3	4	1	**2**
β	4	**1**	2	3
γ	**1**	2	3	4
δ	2	3	**4**	1

Men's List

$\longrightarrow$

	D	B	A	C
α	**2**	4	3	1
β	3	**1**	4	2
γ	4	2	**1**	3
δ	1	3	2	**4**

Men's List

4.2 3×3 Latin Squares

We want to show the feasible happiness levels for preference matrices of order $n = 3$. For a 3×3 Latin square there are only two possible constructions for the diagonal as shown below.

a	b	c
c	a	b
b	c	a

a	c	b
c	b	a
b	a	c

When constructing a 3×3 Latin square, the only possible diagonals are $a + b + c$ and $3a$. If the diagonal is $a + b + c$, then the only possible value is $1 + 2 + 3 = 6$. If the diagonal is $3a$ and $a \in \{1, 2, 3\}$, then the possible values for the sum of the diagonal are 3, 6, 9. It is impossible to construct a 3×3 Latin square with a diagonal of $2a+b$. As can be seen in following figure, an a must be placed in the same row and column as b. However, this is impossible because if an a was placed in the same row and column as b, then the matrix would contain a row and column with two a's and this would not be a Latin square.

a		
	a	
		b

4.3 4×4 Latin Squares

Next, we look at the 4×4 Latin squares and are interested in the feasible happiness levels from the minimum 4 to the maximum 16. We only consider three structures for 4×4 Latin squares at shown in the figure below. Using these structures, we are able to construct Latin squares that have a happiness level for all values from 4 to 16 except 5 and 15.

a	b	d	c
b	a	c	d
c	d	a	b
d	c	b	a

a	b	c	d
b	a	d	c
c	d	b	a
d	c	a	b

a	b	c	d
c	a	d	b
d	c	b	a
b	d	a	c

The two values 5 and 15 are impossible to attain. In order for an individual preference matrix to have a happiness level of 5, the entries along the diagonal would have to be $1 + 1 + 1 + 2$. However, this is not possible, because it is a similar situation as the 3×3 matrix with a diagonal of $2a + b$. In the 4×4 Latin square, it is impossible to have a diagonal of $3a + b$. The happiness value 15 is also impossible for the same reason. The only way for the sum of the diagonal to equal 15 is to have entries of $4 + 4 + 4 + 3$. Similarly, this is of the form $3a + b$, which is not possible. This is illustrated below.

1			
	1		
		1	
			2

4			
	4		
		4	
			3

4.4 General Case

The minimum and maximum happiness levels for an individual matrix are n and n^2, respectively. Our goal is to prove that all of the happiness levels between n and n^2 are possible or not possible. First, there are 2 values between n and n^2 that we can show are not possible. It is impossible to construct an individual preference matrix such that it has a happiness level of $n+1$ or n^2-1. The only way to construct a Latin square with one of these values along the diagonal is with a diagonal of the form $(n-1)a+b$. This is not possible because an a has to be placed in the same column as b; however, there is already an a in every row where an a could be placed. Similarly, an a has to be placed in the same row as b; however, there is an a in every column where an a could be placed. This is illustrated below. There is no where to put an a in the yellow places.

$$\begin{pmatrix} a & & & & \\ & a & & & \\ & & a & & \\ & & & \ddots & \vdots \\ & & & \cdots & b \end{pmatrix}$$

We want to define three structures of Latin squares with which we can use to construct the remaining values between n and n^2. The diagonals we want are na, $(n-2)a+2b$ and $(n-2)a+b+c$. With these structures we can construct the remaining values from n to n^2. First, we need to show that Latin squares can be constructed with these diagonals. To construct a Latin square with the the diagonal na is trivial. Below are the structures for $n \times n$ Latin squares with diagonals of $(n-2)a+2b$ and $(n-2)a+b+c$.

$$
\left(\begin{array}{cccccccccc}
1 & 2 & 3 & \cdots & & & & & n-1 & n \\
n & 1 & 2 & 3 & \cdots & & & & n-2 & n-1 \\
n-1 & n & 1 & 2 & 3 & \cdots & & & n-3 & n-2 \\
\vdots & & & \ddots & & & & & & \vdots \\
 & & & & & & & & & \\
 & & & & & & & & & \\
5 & 6 & \cdots & & & n & 1 & 2 & 3 & 4 \\
\hline
2 & 4 & 5 & 6 & 7 & \cdots & n-1 & 1 & n & 3 \\
3 & 5 & 4 & 7 & 6 & \cdots & & & 2 & 1 \\
4 & & & & & & & & 1 & 2 \\
\multicolumn{10}{c}{\textbf{Diagonal with structure } (n-2)a+2b}
\end{array}\right)
$$

$$
\left(\begin{array}{cccccccccc}
1 & 2 & 3 & \cdots & & & & & n-1 & n \\
n & 1 & 2 & 3 & \cdots & & & & n-2 & n-1 \\
n-1 & n & 1 & 2 & 3 & \cdots & & & n-3 & n-2 \\
\vdots & & & \ddots & & & & & & \vdots \\
 & & & & & & & & & \\
 & & & & & & & & & \\
5 & 6 & \cdots & & & n & 1 & 2 & 3 & 4 \\
\hline
3 & 4 & 5 & 6 & 7 & \cdots & n-1 & 1 & n & 2 \\
4 & 3 & 6 & 5 & 8 & \cdots & & & 2 & 1 \\
2 & & & & & & & & 1 & 3 \\
\multicolumn{10}{c}{\textbf{Diagonal with structure } (n-2)a+b+c}
\end{array}\right)
$$

We can use a simple algorithm and the above Latin square structures to obtain all the happiness levels that are possible between n and n^2.

Algorithm

Suppose we want to obtain a happiness level of k in an $n \times n$ Latin square. There exist integers t and r such that $k = tn + r$.

1. For $n - 2$ places in the diagonal put a t. We need the last two places of the diagonal to add up to $k - (n-2)t$. Pick two numbers from 1 to n to put in these places such that this is true. If this is not possible, try the next step.

2. For $n - 2$ places in the diagonal put a $t + 1$. We need the last two places of the diagonal to add up to $k - (n-2)(t+1)$. Pick two numbers from 1 to n to put in these places such that this is true.

4.4.1 Examples

The following are two examples that illustrate the use of the algorithm. Suppose we want a happiness value of $k = 1527$ on a 50×50 Latin square. Note that $1527 = 30 \cdot 50 + 27$. Place a 30 on $n - 2$ places of the diagonal. The last two places need to add up to $1527 - 30 \cdot 48 = 1527 - 1440 = 87$. We can use 43 and 44. Our diagonal will look like the following.

$$\begin{pmatrix} 30 & & & & & \\ & 30 & & & & \\ & & \ddots & & & \\ & & & 30 & & \\ & & & & 43 & \\ & & & & & 44 \end{pmatrix}$$

Suppose we want a happiness value of $k = 1549$ on a 50×50 Latin square. Note that $1549 = 30 \cdot 50 + 49$. If we place a 30 on $n-2$ places of the diagonal, then we need the last two places to add up to $1549 - 30 \cdot 48 = 109$. It is impossible for two numbers between 1 and 50 to add up to 109. We can try to place a 31 on $n-2$ places of the diagonal. The last two places need to add up to $1549 - 31 \cdot 48 = 1549 - 1488 = 61$. We can use 32 and 29. Our diagonal will look like the following.

$$\begin{pmatrix} 31 & & & & & \\ & 31 & & & & \\ & & \ddots & & & \\ & & & 31 & & \\ & & & & 32 & \\ & & & & & 29 \end{pmatrix}$$

5. Distribution

Next, we examine the distribution of happiness levels for an individual preference matrix. In this section we will look at the $3 \times 3, 4 \times 4$ and 5×5 matrices.

5.1 3×3 Latin Squares

Recall that we are only looking at individual preference matrices of Latin square structures. In addition, we need only examine the diagonals of these Latin squares to determine all of the happiness levels for a given n.

As we discussed earlier, the only possible diagonals of 3×3 Latin squares are of the form $3a$ and $a+b+c$. The only possible happiness levels are 3, 6 and 9. For the 3×3 Latin squares we counted the happiness levels by hand. Below are all of the 3×3 Latin squares and their happiness levels.

1	2	3
2	3	1
3	1	2

Happiness level = 1+3+2=6

1	2	3
3	1	2
2	3	1

Happiness level = 1+1+1=3

1	3	2
2	1	3
3	2	1

Happiness level = 1+1+1=3

1	3	2
3	2	1
2	1	3

Happiness level = 1+2+3=6

2	1	3
1	3	2
3	2	1

Happiness level = 2+3+1=6

2	1	3
3	2	1
1	3	2

Happiness level = 2+2+2=6

2	3	1
1	2	3
3	1	2

Happiness level = 2+2+2=6

2	3	1
3	1	2
1	2	3

Happiness level = 2+1+3=6

3	1	2
1	2	3
2	3	1

Happiness level = 3+2+1=6

3	1	2
2	3	1
1	2	3

Happiness level = 3+3+3=9

3	2	1
1	3	2
2	1	3

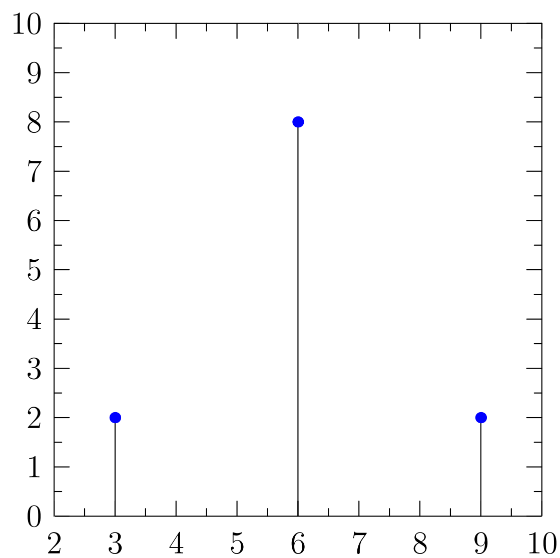
Happiness level = 3+3+3=9

3	2	1
2	1	3
1	3	2

Happiness level = 3+1+2=6

The distribution of the happiness levels for the 3×3 matrices are displayed in the following chart and graph.

Happiness Level	# of Occurrences
3	2
4	0
5	0
6	8
7	0
8	0
9	2



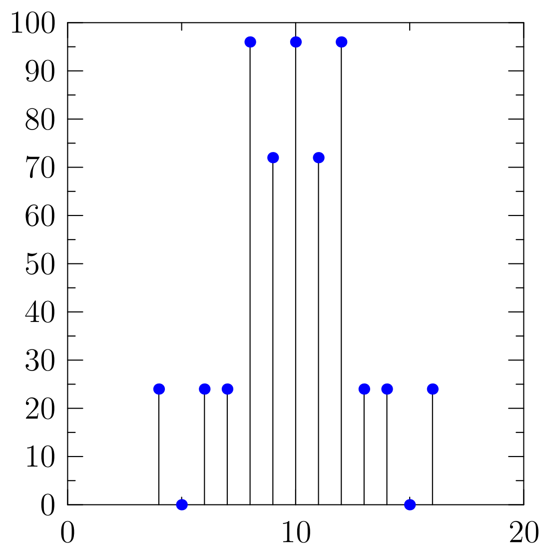
5.2 4×4 and 5×5 Latin Squares

For the 4×4 and 5×5 Latin squares we used computer programs to generate all of the Latin squares and calculate the happiness level for each matrix.

The distribution of the happiness levels for the 4×4 matrices are displayed in the following chart and graph.

Happiness Level	# of Occurrences
4	24
5	0
6	24
7	24
8	96
9	72
10	96

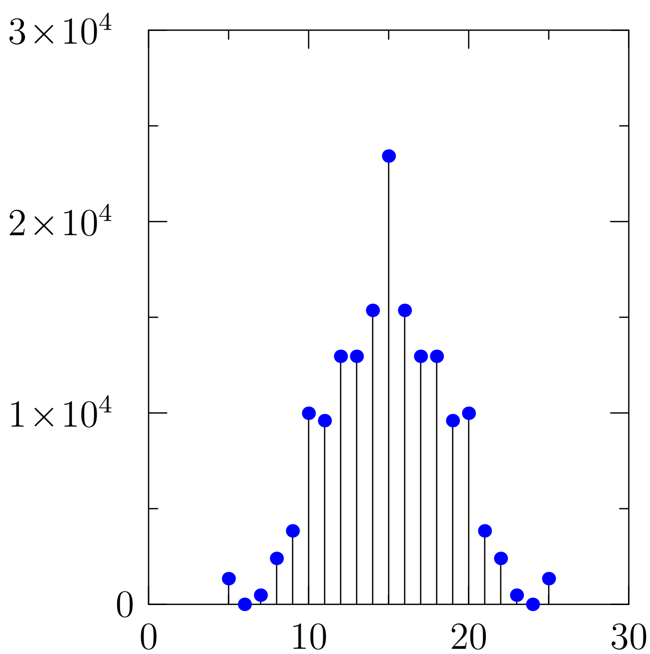
Happiness Level	# of Occurrences
11	72
12	96
13	24
14	24
15	0
16	24



The distribution of the happiness levels for the 5×5 matrices are displayed in the following chart and graph.

Happiness Level	# of Occurrences
5	1344
6	0
7	480
8	2400
9	3840
10	9984
11	9600
12	12960
13	12960
14	15360
15	23424

Happiness Level	# of Occurrences
16	15360
17	12960
18	12960
19	9600
20	9984
21	3840
22	2400
23	480
24	0
25	1344



Note that the distributions for $n = 3, 4, 5$ are not monotone.

6. Open Problems

The shape of the distribution for $n \geqslant 6$ is not known. It would not be easy to determine the exact distribution for $n \geqslant 6$ because of the limitations of the computer programs we used for $n = 4, 5$. In addition, the number of Latin squares in not known for $n \geqslant 12$.[7]

Our research is inspired by stable matchings. A *stable matching* is a matching such that there is not a man-woman pair who both prefer each other more than their current partners.[4] The connection between stability and happiness could be investigated using the limitation of Latin square structures for preference matrices.

Bibliography

[1] Gale, D.; Shapley, L. S. *College Admissions and the Stability of Marriage* American Mathematical Monthly 69, 9–14, 1962.

[2] Golle, Philippe. *A Private Stable Matching Algorithm* In proc. of Financial Cryptography 2006. crypto.stanford.edu/ pgolle/papers/stable.pdf

[3] Gusfield, D.; Irving, R. W. *The Stable Marriage Problem: Structure and Algorithms* MIT Press, p. 54, 1989

[4] Knuth, Donald E. *Stable Marriage and Its Relation to Other Combinatorial Problems* American Mathematical Society. Providence, Rhode Island, 1997.

[5] Roth, Alvin E. *The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory.* University of Chicago Press (1984). *The Journal of Political Economy*, vol 92. April 2007 http://www.jstor.org/view/00223808/di951028/95p0086q/0

[6] Teo, Chung–Piaw; Sethuraman, Jay; Tan, Wee–Peng. *Gale–Shapley Stable Marriage Problem Revisited: Strategic Issues and Applications (2001). March 2007* http://www.columbia.edu/ js1353/pubs/tst–ms01.pdf

[7] *The On–Line Encyclopedia of Integer Sequences.* http://www.research.att.com/˜njas/sequences/A002860