The Deal, the Whole Deal, and Nothing but the Deal

From time to time, one is dealt a "freak" hand of some kind; even several in a duplicate session. There is a view, amongst Bridge players at Club level, that there may well be a causal reason. "This shouldn't happen by mere chance", is sometimes opined. I have heard such hands attributed to new packs of cards, to poor shuffling, or more commonly to computer-dealt hands, such as those used for simultaneous pairs, or championship events. There have been several research articles, over the past 10 years or more, assessing such aspects as the lengths of suits, and the shapes of hands. I am not aware, however, that there has been a detailed consideration of just how extreme, or otherwise, is a particular deal as a whole, nor of an entire session of play. I have been investigating.

Suit lengths

Many players have a good feel for the frequency of long or short suits. The percentage probabilities at random are approximately:

length	0	1	2	3	4	5	6	7+
%	1	8	21	29	24	12	4	1

To interpret these figures more concretely, consider an evening's duplicate session of 25 boards, say. During the course of the evening, each player will hold 100 suit lengths. From the table it becomes clear that each player should expect, at random, about 1 void, 8 singletons, and one suit of at least 7 cards long. More likely, however, there will be about 4 voids, 32 singletons, and 4 suits as long as or longer than 7 cards, distributed around the room in a more uneven fashion.

Suit/Hand distributions

pattern	%
4333	10.54
4432	21.55
4441	2.99
5332	15.52
5422	10.60
5431	12.93
5440	1.24
5521	3.17
5530	0.90
6322	5.64
6331	3.45
6421	4.70
6430	1.33
6511	0.71
6520	0.65
6610	0.07
7***	3.52
8***	0.36
total	99.87

There is a symmetry in a Bridge deal: 4 suits of 13 cards, shared 13 each, among 4 players. This means that, at random, the percentage probabilities of hand patterns and suit distribution patterns will be the same for the same pattern. These are well known. As given in "The Official Encyclopaedia of Bridge" (1984) they are shown on the left below. They are a useful guide to whether one should be surprised or concerned about a particular holding. The table, however, refers only to individual suits, or hands. You may hold what you consider a "freak", but what about others at the table. They may or may not have extreme distributions in their hands. They may or may not share your concern. What is really needed is a way of estimating the extreme nature, or otherwise, of the entire deal of all four hands, all four suits.

We can represent the situation by a 4 x 4 array of numbers such as:

	N	E	S	W	
S	5	5	2	1	13
H	3	2	4	4	13
D	4	3	3	3	13
C	1	3	4	5	13
	13	13	13	13	52

It transpires that there is simple, method for measuring how extreme is any deal. Before describing it, however, I must introduce a new concept.

The Roseman Distance

There are millions of distinct Bridge deals, yet when considering suit lengths, or hand shapes, one can reduce the number of distinct categories to a manageable size, as shown earlier. For suit lengths, 8 categories are effectively enough. For hand shapes, 20 is more than enough. The Roseman Distance gives us a way of reducing to a manageable size the number of categories for complete deals. It also allows us to address what I consider to be the real question to be asked about “freaks”. The question should be not “How rare is the deal itself?”, but “How rare is its extreme nature?”. Players could benefit from having a way of assessing this numerically.

My approach has been to develop a measure of how far away a particular deal is from some baseline deal. The deal I have chosen as a basis does not exist! Nevertheless I claim is a good place to start. It is:

	N	**E**	**S**	**W**	
S	3.25	3.25	3.25	3.25	13
H	3.25	3.25	3.25	3.25	13
D	3.25	3.25	3.25	3.25	13
C	3.25	3.25	3.25	3.25	13
	13	13	13	13	52

One now needs to choose a measure of “distance”, of real deals, from this hypothetical deal. Consider the flattest of all.

	N	**E**	**S**	**W**	
S	4	3	3	3	13
H	3	4	3	3	13
D	3	3	4	3	13
C	3	3	3	4	13
	13	13	13	13	52

We measure the "Roseman Distance" by, in each cell, initially replacing the entry with the square of the difference of that entry from that in the corresponding cell of the hypothetical base deal (in each case 3.25), then add them all. Finally we subtract 1, and divide by 2! For this flat deal, this produces:

	N	E	S	W	(sum-1)/2
S	9/16	1/16	1/16	1/16	12/16
H	1/16	9/16	1/16	1/16	12/16
D	1/16	1/16	9/16	1/16	12/16
C	1/16	1/16	1/16	9/16	12/16
(sum-1)/2	12/16	12/16	12/16	12/16	(3-1)/2

The total "distance" of this deal from the hypothetical base is therefore 1. Interestingly, this is a whole number. In fact, every "Roseman Distance" will be a whole number. For:

	N	E	S	W	
S	5	3	3	2	13
H	2	3	4	4	13
D	3	3	3	4	13
C	3	4	3	3	13
	13	13	13	13	52

The corresponding table is:

	N	E	S	W	(sum-1)/2
S	49/16	1/16	1/16	25/16	76/16
H	25/16	1/16	9/16	9/16	44/16
D	1/16	1/16	1/16	9/16	12/16
C	1/16	9/16	1/16	1/16	12/16
(sum-1)/2	76/16	12/16	12/16	44/16	4

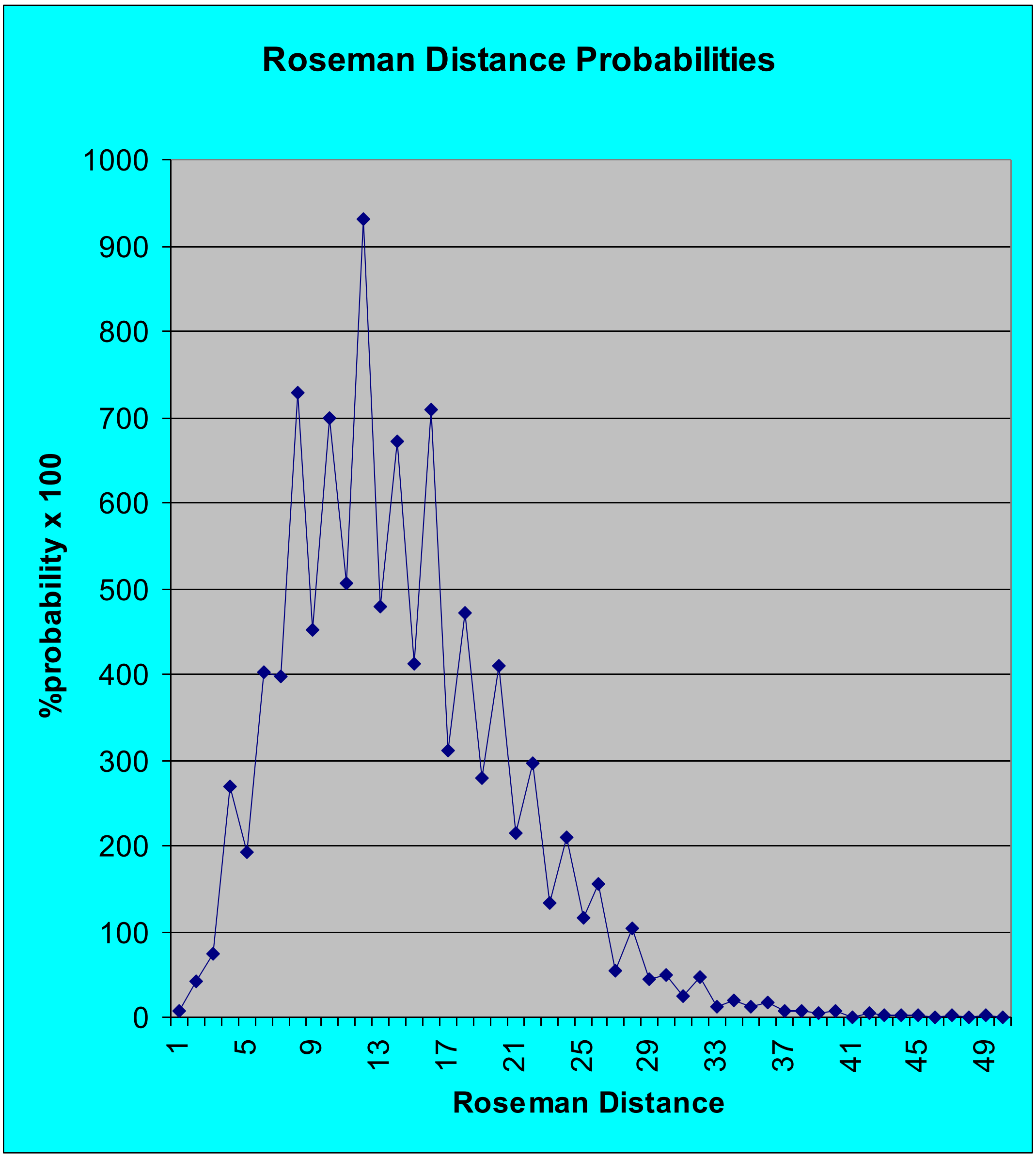
If you replicate this process for these two other deals, however, you will discover that they also produce the “distance” 4.

	N	E	S	W	
S	4	3	3	3	13
H	1	4	4	4	13
D	4	3	3	3	13
C	4	3	3	3	13
	13	13	13	13	52

	N	E	S	W	
S	4	4	3	2	13
H	3	2	4	4	13
D	3	4	2	4	13
C	3	3	4	3	13
	13	13	13	13	52

In this way, although there are millions to consider, the deals all fall into a limited number of categories as measured by their Roseman Distance from our base. The probabilities of the individual 4x4 deal patterns within each “distance” category will differ, but my contention is that to judge whether a complete deal is extreme, it is this “distance” that is best assessed, not the absolute probability of the deal in isolation.

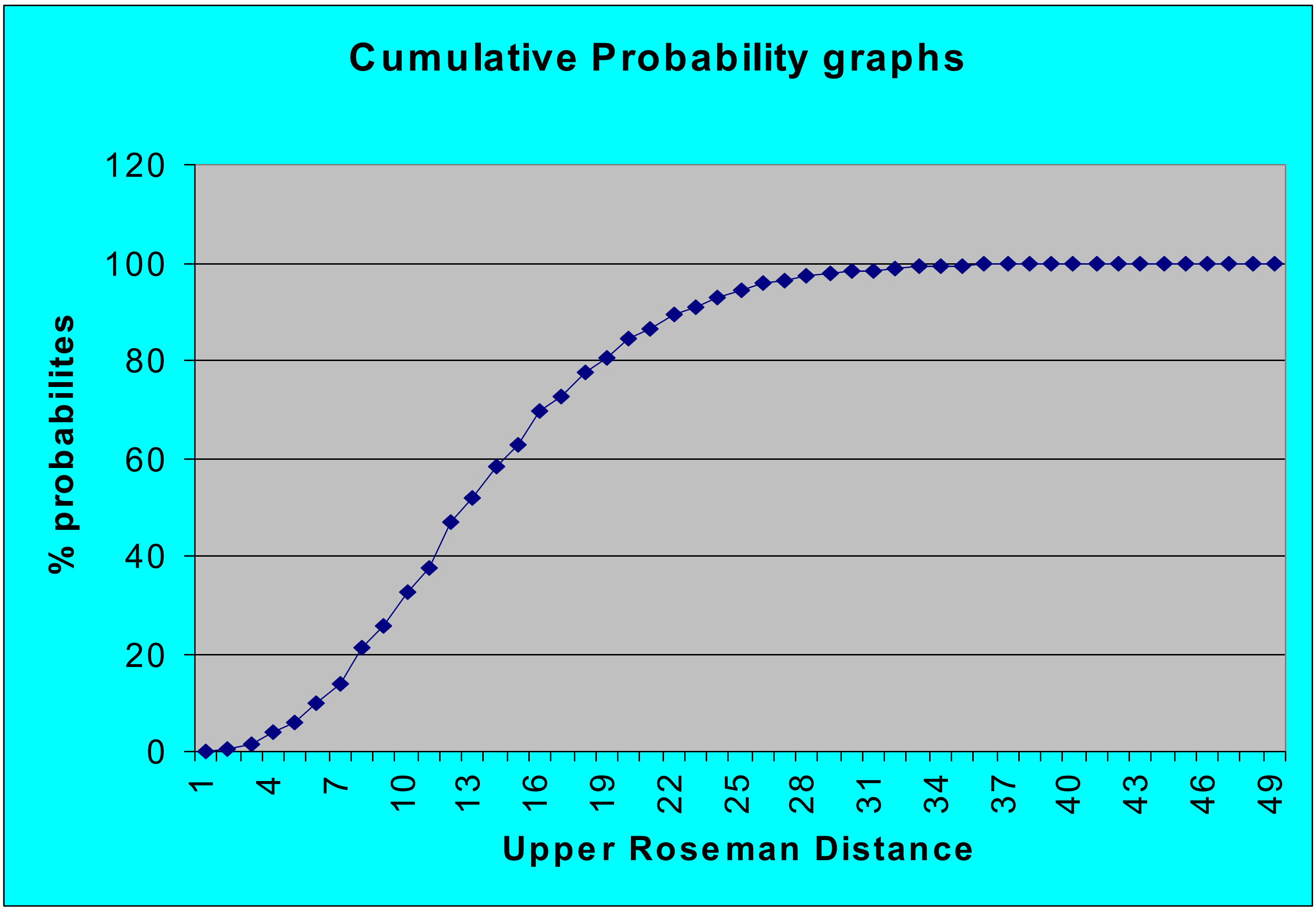
Almost all possible deals fall into 50 such categories. In fact 99.97% of all deals are encompassed by “distances” 1 to 50 in steps of 1. See Appendix 1. Indeed, I have never come across a deal, manual or computer, of “distance” greater than 40. The percentage probability graph of all such “distances” is shown on the following page.

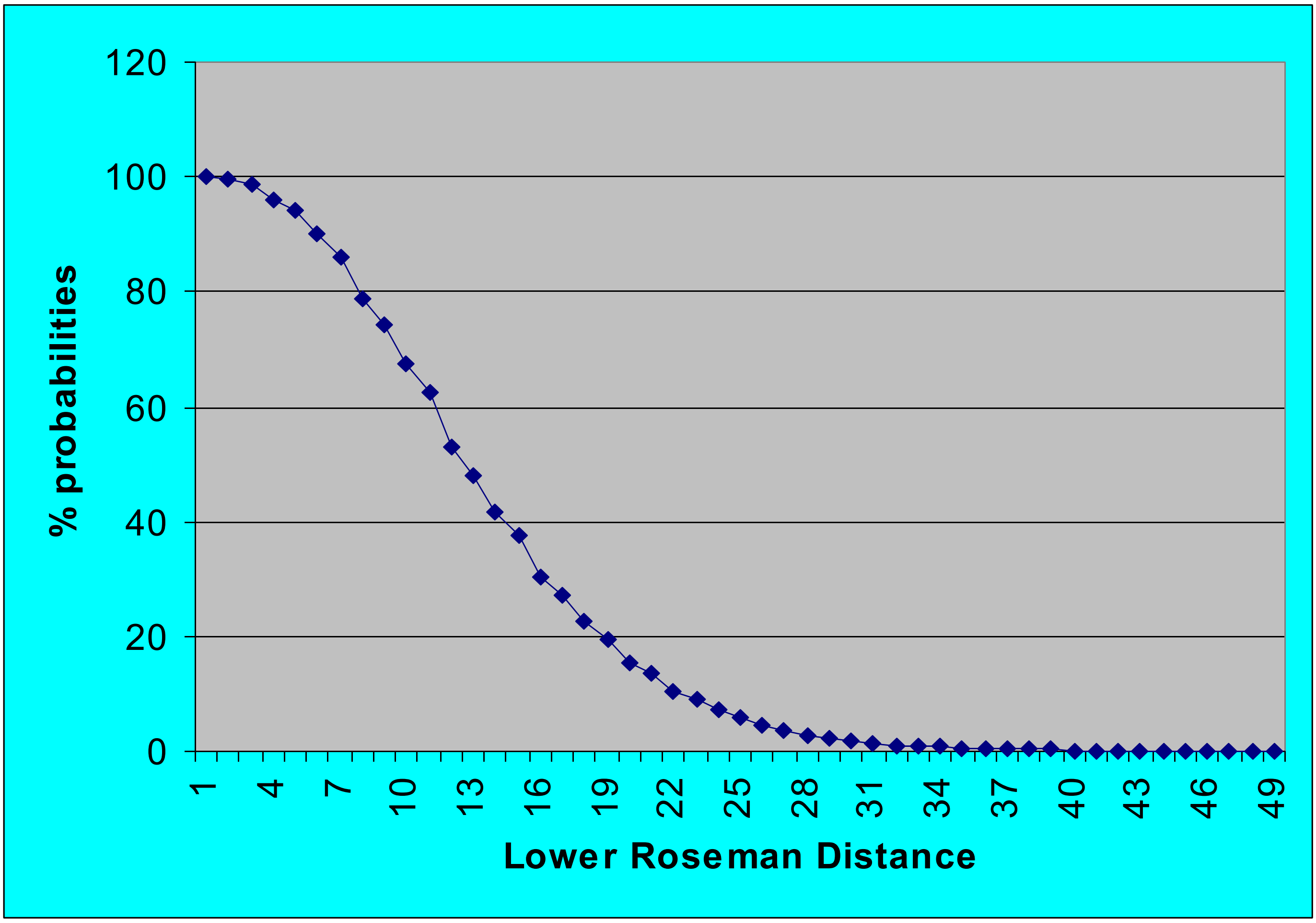


Note the regularity of the jagged envelope. The graph may be notionally split into two, one for the "troughs" at odd "distances" 1, 3, 5 etc. These cover three-eighths of all possibilities, leaving the remaining five-eighths for the "peaks" at even "distances" 2, 4 , 6 etc..

That the "Roseman Distances" fall into two sub-categories, the peaks and the troughs, lies in the unavoidable fact that all bridge hands contain either one odd-length suit, or three odd-length suits. A little arithmetic will produce the result that all Bridge deals (that is, all 4 hands considered together) will contain an even total number of odd-length suits. These totals are confined to 4,6,8,10, or 12. There are no others!

The troughs occur when the this total is 4, 8 or 12. The peaks when the total is 6 or 10. Appendix 2 shows a set of hand/suit patterns all associated with the Roseman Distance of 12. It is easy to check that all the odd-length-suits totals are 6 or 10.


Cumulative Probability graphs
% probabilites
120
100
80
60
40
20
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49
Upper Roseman Distance


% probabilities
120
100
80
60
40
20
0
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49
Lower Roseman Distance

These cumulative charts enable a player to answer directly two questions about any deal encountered.

A. What is the probability of meeting, at random, a deal at least as extreme as this?
B. What is the probability of meeting, at random, a deal no more extreme as this?

One simply calculates the Roseman Distance associated with the 4x4 array representing the deal, then looks up the corresponding probability in the appropriate chart. To help calculate this "distance", Dr Alastair King, of The University of Bath has developed the table below which can be used as a ready-reckoner.

Suit length	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
contribution	21	10	3	0	1	6	15	28	45	66	91	120	153	190

It lists a contribution "N" towards that "distance" for each possible suit length "S", where N = (2S - 7)*(S-3). Add up all the numbers so acquired and divide by 4. That is your final distance. Take a reading from a cumulative chart to draw a conclusion about the probability of meeting a deal more (or less) extreme.

Consider this deal, for example:

	N	**E**	**S**	**W**	
S	7	3	2	1	13
H	2	3	4	4	13
D	2	3	3	5	13
C	2	4	4	3	13
	13	13	13	13	52

Looking up the suit lengths in the list and inserting instead the corresponding figure for the "distance" contribution, we see:

	N	E	S	W	sum
S	28	0	3	10	41
H	3	0	1	1	5
D	3	0	0	6	9
C	3	1	1	0	5
sum	37	1	5	17	60

Now divide this grand total by 4, (60 by 4 = 15). The Roseman Distance for such a deal is therefore 15. Looking up 15 on one cumulative graph we see that we would expect a deal no more extreme than this about 60% of the time. Not surprisingly the other chart says that we would expect a deal at least as extreme as this about 40% of the time.

North may be reflecting that the holding of a 7-card suit is a once-in-a-hundred event for the suit in question. A deal as a whole, however, at least as "extreme" as this one, will occur more than one time in three. Indeed, the following deals (which may not raise any player's eyebrows) also have Roseman Distance 15, hence are just as extreme.

	N	E	S	W	
S	2	6	3	2	13
H	2	3	3	5	13
D	6	1	4	2	13
C	3	3	3	4	13
	13	13	13	13	52

	N	E	S	W	
S	2	4	6	1	13
H	3	3	2	5	13
D	4	3	4	2	13
C	4	3	1	5	13
	13	13	13	13	52

It is my contention that a judicious consideration of these Roseman Distances will put subjective "freak" deals into a better objective perspective.

NB. For those who find the graphical representations difficult to read Appendix 3 lists the cumulative percentages in a look-up form.

COMPLETE SESSIONS

From time to time, at random, one will come across individual deals, during a single session, that seem more extreme than one would like to expect. Yet one should not complain too vigorously without considering all the other deals in the same session. We therefore need a way of assessing whether the collection of deals comprising a complete session is unusual or unexpected in some way. This can be done using the distance concept described earlier.

The “distance” concept allows us to place a separate probability on the “extremity” of each deal of a session. Each deal in a duplicate session is independent of every other deal because different packs of cards are used. [This may not always be true of Rubber Bridge or Chicago where a manual shuffle of the same physical pack intervenes]. The probability of the entire session may then be calculated by multiplying these probabilities together.

This is easier said than done. Each probability is small, so the product of say 24 numbers will be too small to grasp readily. Fortunately there is a standard technique for dealing with the situation. Instead of multiplying the probabilities themselves, one simply adds their logarithms! I did this for 5000 sessions of 24 boards each, using Monte Carlo methods. The log scores for each “distance” are given at Appendix 4. The score totals over all 5000 24-board sessions ranged from -28 to -40 distributed as follows:

Score	-28/-29	-29/-30	-30/-31	-31/-32	-32/-33	-33/-34	-34/-35	-35/-36	-36/-37	-37/-38	-38/-39	-39/-40	<-40
% prob	0.02	0.7	5.6	15.0	24.6	22.2	16.6	9.1	4.0	1.5	0.5	0.2	0.04

The average score for a 24-board session was -33.3, with standard deviation 1.64. Scores more negative than -33.3 will signify that there were relatively more rare deals in the session, but not necessarily more extreme deals. Very flat deals are quite as rare as many of the so-called freaks. Comparable tables for any session length could be readily computed in the same way. Although, if the above experiment were considered general enough, an expectancy of a score of (-33.3/24) = -1.4 per board may be a good approximation for comparison. Hence, if so wished, using the distance concept and scores, an assessment of the oddity, or otherwise, of any pre-dealt session could be comprehensively determined before play begins.

A WAY AHEAD?

What has been achieved by the above discussion has been a way of dividing all of the millions of possible Contract Bridge deals into a mere 40 or so [what a mathematician may call] equivalence classes. Coincidentally, and by a little contrivance, they themselves can be meaningfully numbered 1 to 40 plus, for ease of reference. One might well ask of what use is this? I suggest it has not been just an academic exercise. I could be useful in several ways.

For example, it has already detected human intervention in an apparently randomly dealt duplicate session. Unknown to the players involved, or to myself, the person responsible for the computer generation of an evening's session at a local Club repeatedly produced sets of deals, over and over, until the algorithm produced a set that he felt the players would accept as being "not too extreme". There was no malicious deception intended. The sets were not edited, just censored. This manoeuvre was not apparent from the standard statistics produced by the program algorithm itself. They covered merely rudimentary suit-length and hand-shape counts.

A mere glance at the Roseman distances, however, suggested that something was potentially amiss. Distance "12", the most frequent expected at random, did not appear even once. The perpetrator explained immediately when challenged. I had not been previously suspicious in any way, even by reading the print-out in detail.

I suggest that the Roseman distance calculation be built into every deal-generating algorithm, and its value published beside each deal in any print-out. Further, I suggest that such computer-aided dealing algorithms should be able to use the Roseman distance as an input parameter. It would enable sets of practice deals to be produced rapidly, for such a purpose as system development. Any standard system should be able to deal satisfactorily with deals having Roseman distances less than 20 say, which cover over 80 % of all deals. More expert pairs may also like to feel that they could deal adequately with the more extreme cases with distance 20 or greater.

That is where they may be able to gain a winning edge over their usual rivals. Such deals are rare at the table. A computer, however, could rapidly generate hundreds.

This concept could also be exploited for novelty events, or for any occasion where the extremity, or otherwise, of the distribution of cards in each whole deal needed to be specified beforehand. It would be interesting too, if the logarithmic score for each set of computer-dealt sessions were to be published along with the rest of the usual statistics. It would instantly give an indication as to how unusual, but not necessarily extreme, the deals in the session had been. After all, "flat" deals can be just as "rare" as freaks.

APPENDIX ONE

In a sample set of 10,000 simulated deals these are the observed frequencies of the Roseman Distances from 1 to 50.

"Distance"	frequency	"Distance"	frequency
1	8	26	155
2	42	27	55
3	74	28	103
4	269	29	45
5	193	30	50
6	402	31	24
7	398	32	48
8	729	33	12
9	451	34	20
10	698	35	12
11	506	36	17
12	932	37	7
13	478	38	7
14	671	39	6
15	413	40	7
16	708	41	1
17	311	42	5
18	471	43	2
19	278	44	2
20	410	45	-
21	216	46	2
22	296	47	1
23	133	48	2
24	209	49	1
25	117	50	2

They cover 9999 of the deals.

APPENDIX TWO

The following 32 complete 4-hand (or 4-suit) shapes held in one deal by the four players, or distributed among the four suits, in some allocation, all have the same Roseman Distance from the hypothetical base deal. This distance is "12", the most common, occurring about 9% of the time. The deals are to be read, left to right. Each suit/hand shape is in high to low order for ease of address. I would claim that a player should not regard any such 4-hand combination as more, or less, "extreme" than any other.

6421	5431	4333	4333
6421	5422	4432	4333
6421	5332	5332	4333
6421	5332	4432	4432
6421	4441	4432	4333
6331	5431	4432	4333
6331	5422	4432	4432
6331	5332	5332	4432
6331	5332	4441	4333
6331	4441	4432	4432
6322	5431	5332	4333
6322	5431	4432	4432
6322	5422	5422	4333
6322	5422	5332	4432
6322	5332	4441	4432
5530	5332	4432	4333
5521	5431	4432	4333
5521	5422	5332	4333
5521	5422	4432	4432
5521	5332	5332	4432
5521	4441	4432	4432
5440	5422	4432	4333
5440	5332	5332	4333
5440	5332	4432	4432
5431	5431	5422	4333
5431	5431	5332	4432
5431	5431	4441	4333
5431	5422	5422	4432
5431	5422	5332	5332
5431	5422	4432	4441
5431	5332	5332	4441
5422	5422	5332	4441

APPENDIX THREE

Cumulative percentages of Roseman Distances:

UP TO:

1	0.08
2	0.50
3	1.24
4	3.93
5	5.86
6	9.88
7	13.86
8	21.15
9	25.66
10	32.64
11	37.70
12	47.02
13	51.80
14	58.51
15	62.64
16	69.72
17	72.83
18	77.54
19	80.32
20	84.42
21	86.58
22	89.54
23	90.87
24	92.96
25	94.13
26	95.68
27	96.23
28	97.26
29	97.71
30	98.21
31	98.45
32	98.93
33	99.05
34	99.25
35	99.37
36	99.54
37	99.61
38	99.68
39	99.74
40	99.81
41	99.82
42	99.87
43	99.89
44	99.91
45	99.93
46	99.94
47	99.96
48	99.97

OVER:

1	99.92
2	99.50
3	98.76
4	96.07
5	94.14
6	90.12
7	86.14
8	78.85
9	74.34
10	67.36
11	62.30
12	52.98
13	48.20
14	41.49
15	37.36
16	30.28
17	27.17
18	22.46
19	19.68
20	15.58
21	13.42
22	10.46
23	9.13
24	7.04
25	5.87
26	4.32
27	3.77
28	2.74
29	2.29
30	1.79
31	1.55
32	1.07
33	0.95
34	0.75
35	0.63
36	0.46
37	0.39
38	0.32
39	0.26
40	0.19
41	0.18
42	0.13
43	0.11
44	0.09
45	0.07
46	0.06
47	0.04
48	0.03

APPENDIX FOUR

Table of Logarithmic Probabilities for each individual”distance”

Distance	Log prob
1	-3.09691
2	-2.37675
3	-2.13077
4	-1.57025
5	-1.71444
6	-1.39577
7	-1.40012
8	-1.13727
9	-1.34582
10	-1.15614
11	-1.29585
12	-1.03058
13	-1.32057
14	-1.17328
15	-1.38405
16	-1.14997
17	-1.50724
18	-1.32698
19	-1.55596
20	-1.38722
21	-1.66555
22	-1.52871
23	-1.87615
24	-1.67985
25	-1.93181
26	-1.80967
27	-2.25964
28	-1.98716
29	-2.34679
30	-2.30103
31	-2.61979
32	-2.31876
33	-2.92082
34	-2.69897
35	-2.92082
36	-2.76955
37	-3.1549
38	-3.1549
39	-3.22185
40	-3.1549
41	-4
42	-3.30103
43	-3.69897
44	-3.69897
45	-3.69897
46	-3.69897
47	-4
48	-3.69897
49	-3.39794