Genetic Algorithms

1 Introduction

1.1 Chromosomes, Mutation, and Mating

- Genetic algorithm is the concept of chromosome, and the operation of crossover and mutation.
- Each chromosome consists of a string of bits.
- The chromosome length is affected by desired accuracy.
- The operation of mutation randomly changes the value of a single bit.
- The operation of crossover is a form of mating which combines two chromosomes to produce two new chromosomes.

1.2 Populations, Fitness, and Generations

- A genetic algorithm operates by maintaining a population of chromosomes.
- Chromosome population members are called individuals.
- Each individual is assigned a fitness value based on a problem-specific evaluation function.
- The average fitness of the population increases. Genetic algorithms reward maximum fitness.
- Each step of operations is called a generation. A new population results from these operations.

2 Genetic Algorithms

2.1 Genetic Algorithms Involve Myriad Analogs

- Kookie, a cookie maker, who is trying to optimize the amount of sugar and flour in his cookies.

	9	1	2	3	4	5	4	3	2	1
	8	2	3	4	5	6	5	4	3	2
	7	3	4	5	6	7	6	5	4	3
	6	4	5	6	7	8	7	6	5	4
Sugar	5	5	6	7	8	9	8	7	6	5
	4	4	5	6	7	8	7	6	5	4
	3	3	4	5	6	7	6	5	4	3
	2	2	3	4	5	6	5	4	3	2
	1	1	2	3	4	5	4	3	2	1
		1	2	3	4	5	6	7	8	9

Flour

Bump mountain

Figure 2.1-1 Cookie quality is dependent on the number of kilograms of flour and sugar per batch.

- Kookie wants to find a good combination without trying every combination.
- Kookie decides that each batch of cookies is an "individual."
- A "chromosome" consists of two "genes," each of which is a number from 1 to 9.

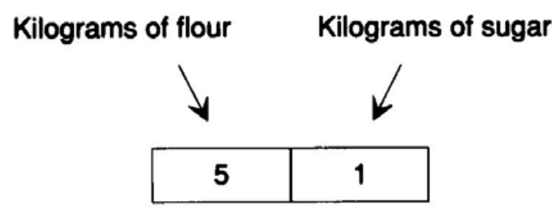


Figure 2.1-2 A chromosome in the cookie world consists of two numbers, which act as gene analogs.

- To mimic chromosome mutation, Kookie selects one of the chromosome's two genes randomly, and alters it randomly by adding or subtracting 1, taking care to stay within the 1-to-9 range.

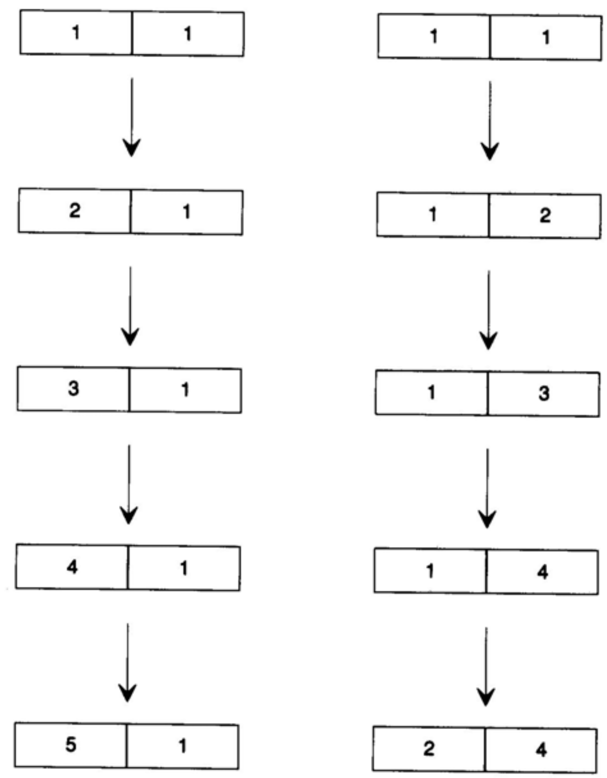


Figure 2.1-3 Two chromosomes undergoing a series of mutations.

- Uniform mutation: constant mutation range
- Non-uniform mutation: varied mutation range (large->small)
- Binary mutation: 0->1 or 1-> 0
- Mutation probability = (number of mutated chromosomes in each generation)x100%/(population size)

- To mimic the crossover involved in mating, Kookie cuts two chromosomes in the middle and rejoins them.

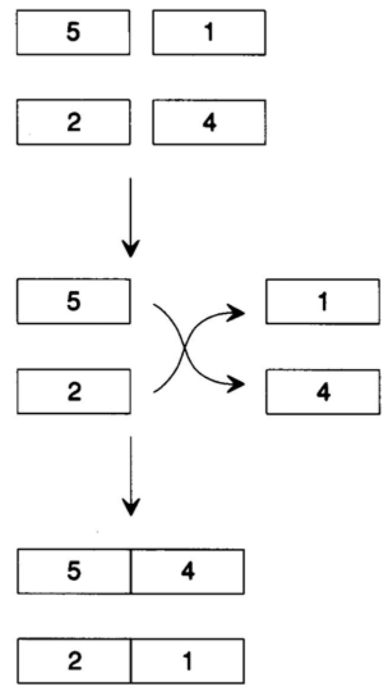


Figure 2.1-4 Two chromosomes undergoing crossover.

- One-point crossover
- Two-point crossover
- Uniform crossover
- Crossover probability = (number of crossed over chromosomes in each generation)x100%/(population size)

2.2 The Standard Method Equates Fitness with Relative Quality

- The fitness of a chromosome is the probability that the chromosome survives to the next generation.

$$f_i = \frac{q_i}{\sum_j q_j} \tag{2.2-1}$$

A population consists of four chromosomes, collectively exhibiting 1-4, 3-1, 1-2, and 1-1 chromosomes.

Chromosomes	Quality	Standard fitness
1 4	4	0.40
3 1	3	0.30
1 2	2	0.20
1 1	1	0.10

To mimic natural selection in general,

- Create an initial "population" of one chromosome.
- Mutate one or more genes in one or more of the current chromosomes.
- Mate one or more pairs of chromosomes.
- Add the mutated and offspring chromosomes to the current population.
- Create a new generation by keeping the best of the current population's chromosomes, along with other chromosomes selected randomly from the current population. Bias the random selection according to assessed fitness.

2.3 Genetic Algorithms Generally Involve Many Choices

- How many chromosomes are to be in the population? If the number is too low, all chromosomes will soon have identical traits and crossover. It will do nothing. If the number is too high, computation time will be unnecessarily excessive.
- What is the mutation rate? If the rate is too low, new traits will appear too slowly in the population; if the rate is too high, each generation will be unrelated to the previous generation.
- Is mating allowed? If so, how are mating pairs selected, and how are crossover points determined?
- Can any chromosome appear more than once in a population?

2.4 It Is Easy to Climb Bump Mountain Without Crossover

- Kookie starts with a single chromosome located at 1-1.
- No chromosome is permitted to appear more than once in each generation.
- A maximum of four chromosomes survive from one generation to the next.
- Each survivor is a candidate for survival to the next generation, along with any new chromosomes produced.
- One gene is selected at random in each of the survivors, and is mutated at random. If the mutant is different from any candidate accumulated so far, that mutant is added to the candidates.
- There is no crossover.
- The chromosome with the highest score survives to the next generation.
- The remaining survivors from one generation to the next are selected at random from the remaining candidates, according to the standard method for fitness computation.

- Among 1000 simulation experiments, on average the best combination is obtained at generation 16.
- The luckiest produced the best combination eight generations

Generation 0:

Chromosome	Quality	
1 1	1	-> 12

Generation 1:

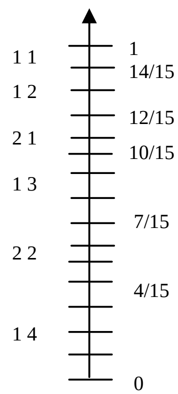
Chromosome	Quality	
1 2	2	-> 1 3
1 1	1	-> 1 2

Generation 2:

Chromosome	Quality	
1 3	3	-> 1 4
1 2	2	-> 2 2
1 1	1	-> 2 1

Combined operations result:

Chromosome	Quality	Fitness
1 4	4	4/15
2 2	3	3/15
1 3	3	3/15
2 1	2	2/15
1 2	2	2/15
1 1	1	1/15



Generation 3:

Chromosome	Quality	
1 4	4	-> 2 4
1 3	3	-> 2 3
1 2	2	-> 3 1
2 1	2	-> 1 3

Operation result:

Chromosome	Quality
2 4	5
2 3	4
3 1	3

Generation 4:

Chromosome	Quality	
2 4	5	-> 2 5
1 4	4	-> 1 5
1 3	3	-> 2 3
2 1	2	-> 2 2

Generation 5:

Chromosome	Quality	
2 5	6	-> 3 5
1 5	5	-> 1 5
2 3	4	-> 1 4
2 2	3	-> 3 2

Generation 6:

Chromosome	Quality	
3 5	7	-> 4 5
1 5	5	-> 1 4
3 2	4	-> 3 1
1 4	4	-> 1 5

Combined operation results:

Chromosome	Quality
4 5	8
3 5	7
1 5	5
3 2	4
1 4	4
3 1	3

Generation 7:

Chromosome	Quality	
4 5	8	-> 5 5
1 5	5	-> 2 5
1 4	4	-> 1 5
3 1	3	-> 2 1

Generation 8:

Chromosome	Quality
5 5	9
4 5	8
2 5	6
2 1	2

2.5 Crossover Enables Genetic Algorithms to Search High-Dimensional Spaces Efficiently

Now suppose Kookie wants to see whether crossover does any good. To decide which chromosomes to cross, Kookie proceeds as follows:

- Kookie considers only the chromosomes that survived from the previous generation.
- For each such chromosome, Kookie selects a mate from among the other survivors. Mate selection is done at random, in keeping with the standard method for computing fitness.
- Each mating pair is crossed in the middle, producing two crossed, offspring chromosomes. If an offspring chromosome is different from any candidate accumulated so far, that offspring chromosome is added to the candidates.

- Using this crossover method, Kookie finds the best combination of ingredients on bump mountain, on average, at generation 14, two generations sooner than without crossover.

2.6 Crossover Enables Genetic Algorithms to Traverse Obstructing Moats

Sugar									
9	1	2	3	4	5	4	3	2	1
8	2	0	0	0	0	0	0	0	2
7	3	0	0	0	0	0	0	0	3
6	4	0	0	7	8	7	0	0	4
5	5	0	0	8	9	8	0	0	5
4	4	0	0	7	8	7	0	0	4
3	3	0	0	0	0	0	0	0	3
2	2	0	0	0	0	0	0	0	2
1	1	2	3	4	5	4	3	2	1
Flour	1	2	3	4	5	6	7	8	9

Moat mountain

Figure 2.6-1 The function relating cookie quality to cookie ingredients exhibits a moat.

- Among 1000 simulation experiments, Kookie finds the best combination on moat mountain problem, on average, only after 155 generations.

2.7 The Rank Method Links Fitness to Quality Rank

- The rank method does not only offer a way of controlling the bias toward the best chromosome, but also eliminates implicit biases, introduced by unfortunate choices of the measurement scale.

- Sort the *n* individuals by quality.
- Let the probability of selecting the *i*th candidate, given that the first *i*-1 candidates have not been selected, be *p*, except for the final candidate, which is selected if no previous candidate has been selected.
- Select a candidate using the computed probabilities.

p = 0.667,

Chromosome	Quality	Rank	Standard fitness	Rank fitness	
1 4	4	1	0.40	0.667	0.667(1-0)
3 1	3	2	0.30	0.222	0.667(1-0.667)
1 2	2	3	0.20	0.074	0.667(1-0.667-0.222)
1 1	1	4	0.10	0.025	0.667(1-0.667-0.222-0.074)
7 5	0	5	0.0	0.012	(1-0.667-0.222-0.074-0.025)

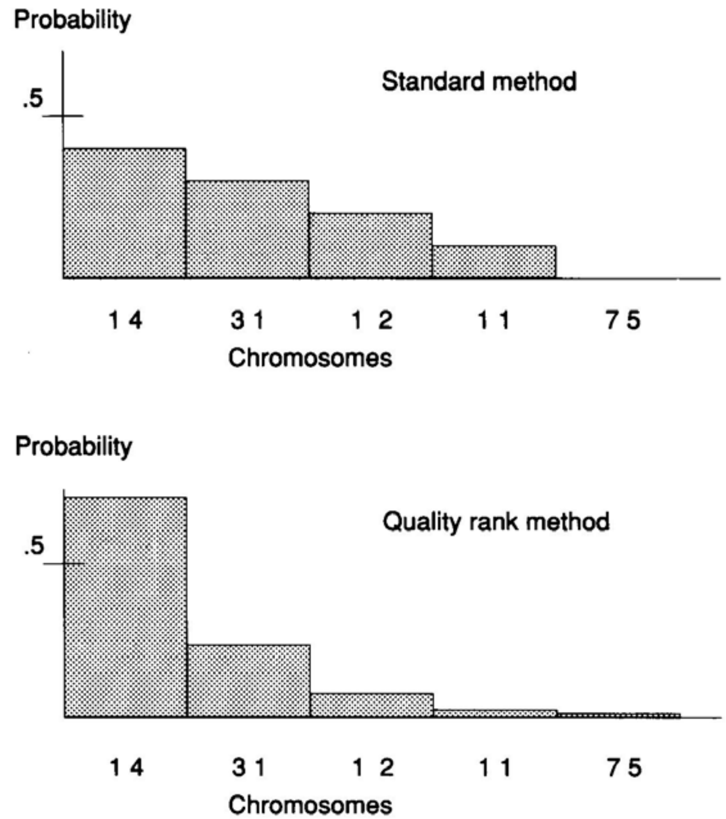


Figure 2.7-1 Fitnesses for five chromosomes, as measured by both the standard method and the rank method, using quality scores as determined by moat mountain.

Assume that chromosome 1-4 is selected, recalculated the probability.

Chromosome	Quality	Rank	Standard fitness	Rank fitness	
3 1	3	1	0.30	0.667	0.667(1-0)
1 2	2	2	0.20	0.222	0.667(1-0.667)
1 1	1	3	0.10	0.074	0.667(1-0.667-0.222)
7 5	0	4	0.0	0.012	(1-0.667-0.222-0.074)

- Kookie finds the best combination, on average, after 75 generations by rank method compared with 155 generations by standard method.

3 Survival of the Most Diverse

3.1 The Rank-Space Method Links Fitness to Both Quality Rank and Diversity Rank

The diversity rank of a chromosome is determined by that inverse squared distance sum:

$$\sum_i \frac{1}{d_i^2} \tag{3.1-1}$$

Consider six candidates that include 5-l, 1-4, 3-1, 1-2, 1-1, and 7-5. The highest-scoring candidate is 5-1 and already selected.

Chromosome	Score	$\frac{1}{d^2}$	Diversity rank	Quality rank	
1 4	4	0.040	1	1	$1/[(5-1)^2+(1-4)^2]$
3 1	3	0.250	5	2	$1/[(5-3)^2+(1-1)^2]$
1 2	2	0.059	3	3	$1/[(5-1)^2+(1-2)^2]$
1 1	1	0.062	4	4	$1/[(5-1)^2+(1-1)^2]$
7 5	0	0.050	2	5	$1/[(5-7)^2+(1-5)^2]$

- Sort the *n* individuals by quality.
- Sort the *n* individuals by the sum of their inverse squared distances to already selected candidates.
- Use the rank method, but sort on the sum of the quality rank and the diversity rank, rather than on quality rank.

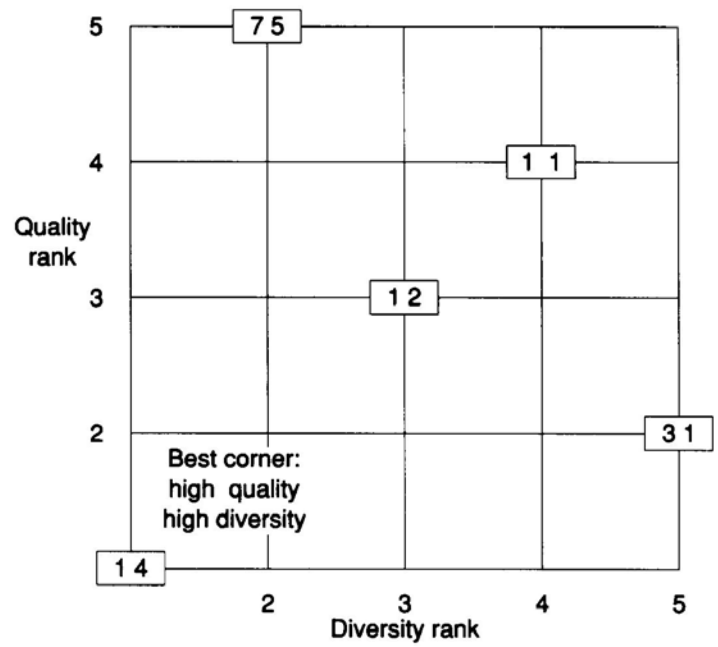


Figure 3.1-1 Rank space enables two chromosomes to be compared in both quality and diversity.

Chromosome	Rank sum	Combined rank	Fitness
1 4	2	1	0.667
3 1	7	4	0.025
1 2	6	2	0.222
1 1	8	5	0.012
7 5	7	3	0.074

- For the chromosome with same rank sum, sort by diversity rank.

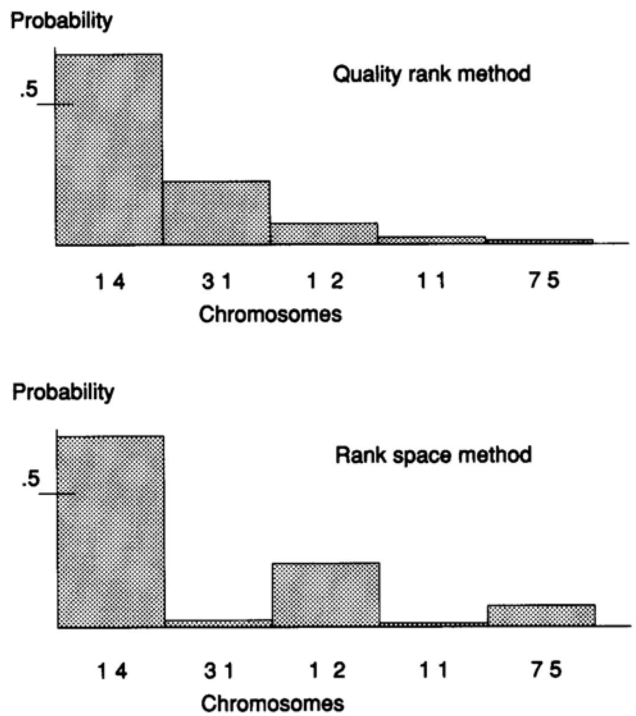


Figure 3.1-2 Fitness for five chromosomes, as measured by the quality-rank method and the rank-space method.

Assume 1-4 is also selected.

Chromosome	$\sum_i \frac{1}{d_i^2}$	Diversity rank	Quality rank	Combined rank	Fitness	
3 1	0.327	4	1	4	0.037	$1/\{[(5\text{-}3)^2+(1\text{-}1)^2]+[(1\text{-}3)^2+(4\text{-}1)^2]\}$
1 2	0.309	3	2	3	0.074	$1/\{[(5\text{-}1)^2+(1\text{-}2)^2]+[(1\text{-}1)^2+(4\text{-}2)^2]\}$
1 1	0.173	2	3	2	0.222	$1/\{[(5\text{-}1)^2+(1\text{-}2)^2]+[(1\text{-}1)^2+(4\text{-}2)^2]\}$
7 5	0.077	1	4	1	0.667	$1/\{[(5\text{-}1)^2+(1\text{-}2)^2]+[(1\text{-}1)^2+(4\text{-}2)^2]\}$

Assume 7-5 is also selected.

Chromosome	$\sum_i \frac{1}{d_i^2}$	Diversity rank	Quality rank	Combined rank	Fitness
3 1	0.358	3	1	3	0.111
1 2	0.331	2	2	2	0.222
1 1	0.190	1	3	1	0.667

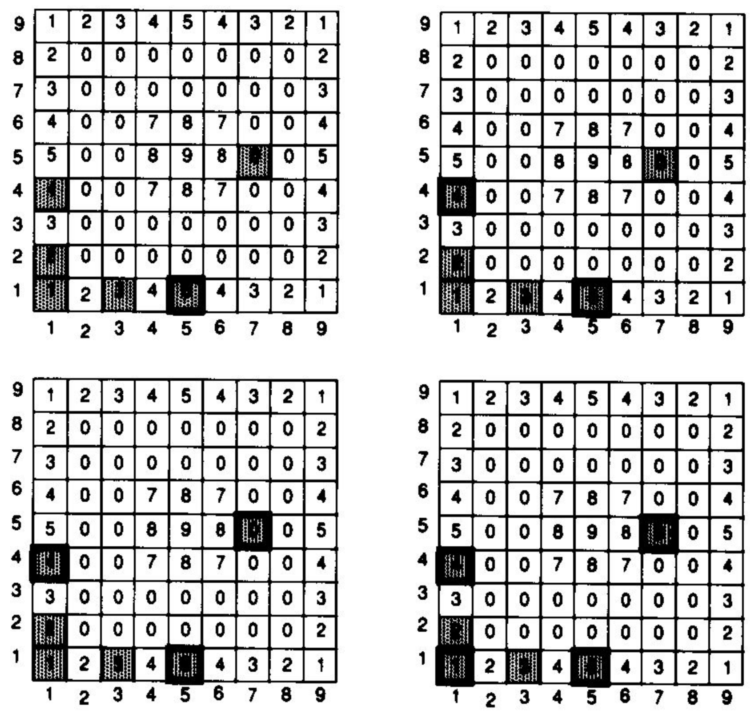


Figure 3.1-3 The sequence of selection by rank-space method.

Tournament selection: select *n* chromosomes, always select the best among *n*, then repeat until population size

3.2 The Rank-Space Method Does Well on Moat Mountain

Summary of average number of generation among 1000 simulation experiments by different methods.

Mountain	Standard method	Quality rank	Rank space
Bump	14	12	12
Moat	155	75	15

- Among 1000 simulation experiments performed by Kookie using the rank-space method, the luckiest produced the best combination after just seven generations, after starting at generation 0 with one 1-1 chromosome.

Generation 0:

Chromosome	Quality		
1 1	1	-> 2 1	

Generation 1:

Chromosome	Quality		
2 1	2	-> 3 1	2 1 x 1 1 -> 2 1, 11
1 1	1	-> 2 1	

Generation 2:

Chromosome	Quality		
3 1	3	-> 4 1	3 1 x 2 1 -> 3 1, 2 1
2 1	2	-> 2 2	
1 1	1		

Generation 3:

Chromosome	Quality		
4 1	4	-> 5 1	2 2 x 4 1 -> 2 1, 4 2
3 1	3	-> 4 1	
1 1	1	-> 1 2	
2 2	0	-> 2 3	

Combined operations result:

Chromosome	Quality
4 1	4
3 1	3
1 1	1
2 2	0
5 1	5
1 2	2
2 3	0
2 1	2
4 2	0

Generation 4:

Chromosome	Quality		
5 1	5	-> 6 1	5 1 x 1 2 -> 5 2, 1 1
3 1	3	-> 3 2	5 1 x 2 3 -> 5 3, 2 1
1 2	2	-> 2 2	3 1 x 1 2 -> 3 2, 1 1
2 3	0	-> 2 4	

Combined operations result:

Chromosome	Quality
5 1	5
3 1	3
1 2	2
2 3	0
6 1	4
2 2	0
3 2	0
2 4	0
2 1	2
1 1	1
5 2	0
3 2	0
5 3	0

Generation 5:

Chromosome	Quality	
5 1	5	5 1 x 2 4 -> 5 4, 2 1
3 1	3	1 2 x 2 4 -> 1 4, 2 2
1 2	2	
2 4	0	

Generation 6:

Chromosome	Quality	
5 4	8	-> 5 5
1 4	4	
3 1	3	
1 2	2	

Generation 7:

Chromosome	Quality
5 5	9
1 4	4
1 2	2
5 2	0

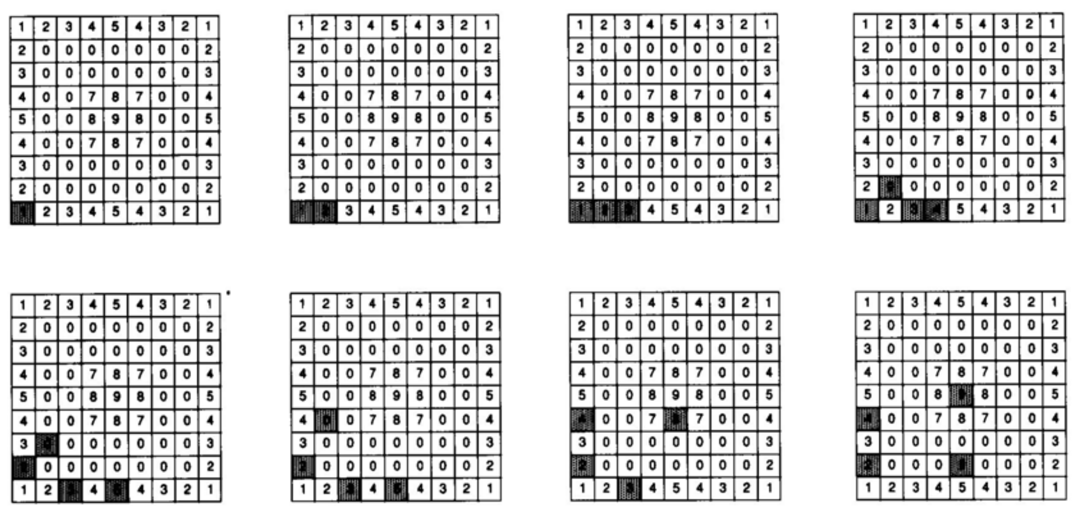


Figure 3.2-1 The results of a lucky experiment using the rank-space method.