Universal Cycles

2011

Yu She

Wirral Grammar School for Girls

Department of Mathematical Sciences

University of Liverpool

Supervisor: Professor P. J. Giblin

Contents

1 Introduction

De Bruijn sequences are sequences where each possible binary / ternary / quaternary... sequence of length n appears exactly once. Universal cycles are generalisations of De Bruijn sequences to other combinatorial structures such as permutations and partitions of a set. Although universal cycles do not exist for partitions of a number, this was an interesting extension to the project.

2 De Bruijn sequences and Eulerian Graphs

For a given alphabet where each digit can take k different values, a De Bruijn sequence $B(k, n)$ is a sequence of numbers in which every possible set of n digits appears exactly once.
Example
In the case of a binary De Bruijn sequence of order 3 (meaning each subsequence has length 3) where each digit is either 0 or 1:
The 8 possible sets of 3 digits are:
000, 001, 011, 111, 110, 101, 010, 100
A De Bruijn sequence $B(2, 3)$ would be: 00011101
By running a window of length 3 along the above sequence, we can see that each of the possible triplets appears exactly once (this includes going around the corner at the end of the sequence to obtain the triplets 010 and 100).
A De Bruijn sequence has length $=k^n$ as each digit can take k different values and there are n digits in each set (called an n-tuple). For the example above, the sequence has length $2^3 = 8$.

2.1 Constructing a De Bruijn Sequence

De Bruijn sequences can be constructed from directed Eulerian Graphs.
Considering the case $k = 2$ and $n = 3$:
To find a De Bruijn sequence of order 3, we write down all the possible binary sets of length 2 and use them as the vertices of the graph.
00, 01, 11, 10

We draw arrows from a vertex to a second vertex when the second digit of the first vertex is the same as the first digit of the second vertex. These arrows can also start and finish at the same vertex, as in the case of the vertices 00 and 11 as the second digit of both these vertices are the same as the first digit. By labelling an arrow AB (travelling from vertex A to vertex B) with the digits in vertex A as well as the last digit of vertex B or the first digit of vertex A as well as the digits in vertex B (the same sequence of digits) and finding an Eulerian cycle around the graph, a De Bruijn sequence can be found. For any vertex A, there are 2 arrows leaving the vertex because the sequence is binary, meaning there are 2 choices for the next digit in the sequence (i.e. 0 and 1), which will be satisfied by joining vertex A to 2 other vertices. Similarly, there will be 2 arrows entering vertex A because there are 2 choices for the previous digit in the sequence (i.e. 0 and 1), which will be satisfied by joining 2 other vertices to vertex A. Therefore, the number of arrows entering the vertex equals the number of arrows leaving the vertex – the vertices are all even.

We can write out all the binary 3-tuples by writing out each binary 2-tuple and adding the digit 0 to the end, then writing out each binary 2-tuple again and adding the digit 1 to the end. This is effectively the same process as the one used when drawing and labelling the arrows of the graph, so each of the arrows is distinct and represents a unique binary 3-tuple. As a result of the way we

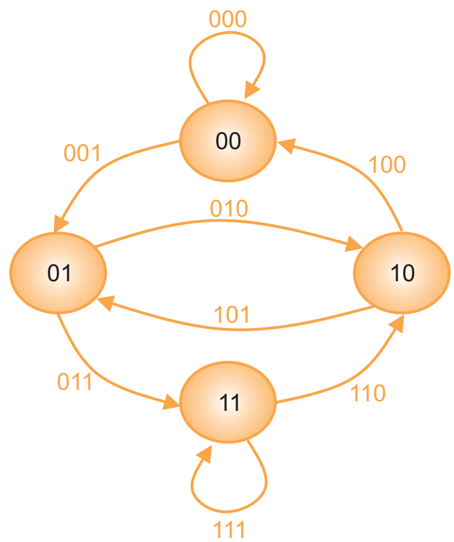


Figure 1: Digraph for the construction of $B(2,3)$

have constructed the graph, the last 2 digits of any arrow entering a given vertex are the same as the first 2 digits of any arrow leaving the same vertex. To write out a De Bruijn sequence from the graph, we find a route which travels along each arrow exactly once and returns to the starting vertex (this is called an Eulerian circuit). We write out the labels on each arrow as we travel along it, deleting the 2 digits from the end of the first arrow that are repeated at the start of the second arrow.

For the graph above, a suitable route would be: 000, 001, 011, 111, 110, 101, 010, 100.

This gives the De Bruijn sequence $B(2,3)$: 00011101

To find a De Bruijn sequence of order 4, we repeat the above process using the binary sets of length 3 as the vertices of the graph and labelling the arrows between vertices with sequences of 4 digits. There are still 2 paths entering and 2 paths leaving each vertex, but the vertices of the graph for $B(2,4)$ now have the same labels as the arrows of the graph for $B(2,3)$. This new graph $N*$ is formed by "doubling" the original graph N.

In this case, an Eulerian circuit for the graph would be: 0000, 0001, 0011, 0111, 1111, 1110, 1100, 1001, 0010, 0101, 1011, 0110, 1101, 1010, 0100, 1000.

This gives the De Bruijn sequence $B(2,4)$: 0000111100101101.

Similarly, the simplest graph for De Bruijn sequences for $k = 2$ (when $n = 2$) has only 2 vertices, with each vertex labelled with just 1 digit i.e. one is labelled '0' and the other '1'.

In this case there is only one Eulerian circuit (if the vertex from which we start does not matter): 00, 01, 11, 10.

This leads to the (only) De Bruijn sequence $B(2,2)$: 0011

This method can also be used when $k \neq 2$. For $k = 3$ and $n = 3$, the vertices of the graph are all the possible combinations of 2 digits where each digit can take one of the 3 values 0, 1 or 2. There are $3^2 = 9$ such combinations so the graph has 9 vertices. The arrows are drawn and labelled in the same way as before and an Eulerian circuit around the graph is found. The De Bruijn sequence in this case would have length $3^3 = 27$.

An Eulerian circuit would be: 011, 111, 112, 121, 210, 101, 012, 120, 201, 010, 102, 021, 211, 110, 100, 002, 022, 221, 212, 122, 222, 220, 202, 020, 200, 000, 001.

This gives the De Buijn sequence $B(3,3)$: 011121012010211002212220200.

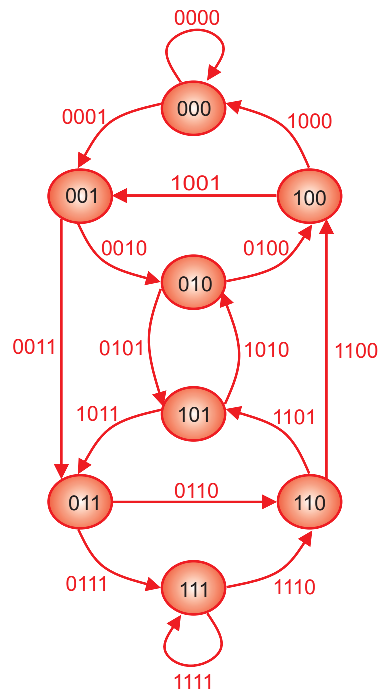


Figure 2: Digraph for the construction of $B(2,4)$

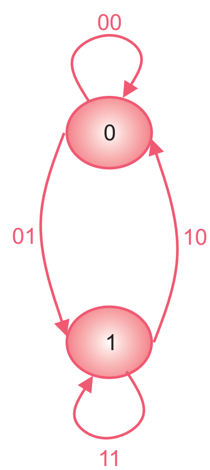


Figure 3: Digraph for the construction of $B(2,2)$

A graph to find $B(k,n)$ can be constructed in the same way for any k and n values and will always have the following properties:

- k^{n-1} vertices
- k^n arrows
- k arrows entering each vertex (in degree $= k$)
- k arrows leaving each vertex (out degree $= k$)

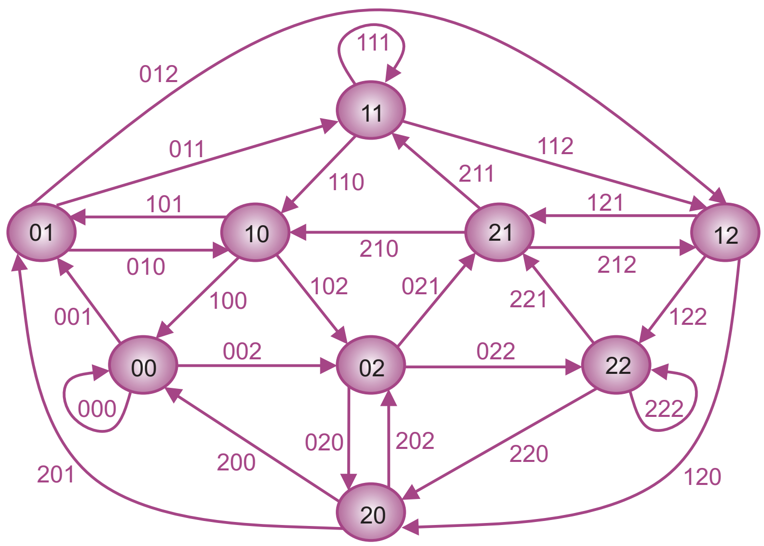


Figure 4: Digraph for the construction of $B(3,3)$

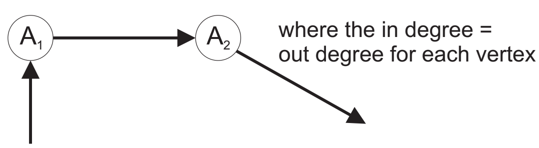
- each vertex is even

2.2 Eulerian Graphs

An Eulerian graph is a graph in which an Eulerian circuit can be found. To ensure that this method works for De Bruijn sequences of any k and n values, we must prove that an Eulerian circuit can always be found for the graphs we use to construct De Bruijn sequences.

Theorem 1 *Let D be a connected digraph. Then D is Eulerian if and only if the out degree of each vertex equals the in degree.*

Proof 1. If D is Eulerian, the out degree of each vertex equals the in degree.
If D has an Eulerian circuit, we can travel along the circuit using each arrow exactly once and return to our starting point. Whenever we pass through a vertex of G there is a contribution of 1 to both the in degree and the out degree of the vertex (meaning they are the same) - this includes the initial vertex, which we return to at the end of the circuit. Since each in and out arrow of D is used just once, the in and out degree of each vertex is the same.
2. If the out degree of each vertex equals the in degree, the graph is Eulerian.
Consider a connected digraph where each vertex is even:



Start at any vertex and follow directed paths to other vertices. When a vertex A_i is reached, there are 2 possible situations:

1. There is an arrow out of A_i, in which case follow this arrow.
2. There is no edge out of A_i – as the in degree of each vertex equals the out degree, this vertex A_i must have been reached before (i.e. the path started at A_i) as we have followed all edges out of A_i. This means a directed cycle C has been formed in the graph.

As there are a finite number of vertices, situation 2 must at some point arise, so it is always possible to find a directed cycle C in the graph.

Let D be a graph where the out degree and in degree of each vertex are equal and the total number of arrows = m. D contains a directed cycle C.
When $m = 0$, D consists of one vertex only and it is therefore Eulerian.
Assume true for any connected digraph where $m < n$.
Now consider a digraph with n arrows. Delete the edges of the directed cycle C from D. The resulting digraph H has $m < n$ and every vertex H has an in degree = out degree, therefore each component of H is Eulerian. By following the directed cycle C, taking Eulerian trails for the components of H when we meet them and returning to cycle C, we can find an Eulerian trail for the graph.
This explains why De Bruijn sequences always exist for any k, n as they are formed from Eulerian circuits of graphs where the vertices are even and of order $n - 1$.

2.3 Fleury's Algorithm

Fleury's algorithm is a way of finding an Eulerian circuit in an Eulerian graph. In the case of De Bruijn sequences, this will be a directed graph with arrows instead of edges connecting the vertices.

1. Choose a starting vertex u.
2. At each stage, traverse any available edge, choosing a bridge (an edge whose removal disconnects a vertex from the rest of the graph) only if there is no alternative.
3. After traversing each edge, erase it (also erase any vertices of degree 0 which result) and then traverse another available edge.
4. Stop when there are no more edges - an Eulerian circuit has been found.

2.4 The number of De Bruijn sequences

There is more than one De Bruijn sequence when $n > 2$ or $k > 2$ and these can be found by taking different Eulerian circuits in the appropriate digraph. The number of distinct binary sequences, where the reverse of a sequence is counted as a different sequence but cyclic permutations are regarded as the same, is $P_n = 2^{2^{n-1}-n}$.

Let the number of different Eulerian circuits in an Eulerian graph N be denoted by $\mid N \mid$. Similarly, the number of Eulerian circuits in $N*$ (the doubled graph of N) is denoted by $\mid N* \mid$.

It was proved by N.G. de Bruijn in [2] that

$$\mid N* \mid = 2^{m-1} \cdot \mid N \mid$$

where m is the order of N (N has m vertices and $2m$ arrows).

This result can be used to prove the number of De Bruijn sequences for any value of n when $k = 2$.

Theorem 2 *The number of binary De Bruijn sequences* $P_n = 2^{2^{n-1}-n}$

Proof When $n = 1$, $P_1 = 1$(this is evident from the graph)
Using the formula: $P_1 = 2^{2^{1-1}-1} = 2^0 = 1$
Therefore the formula is true for $n = 1$.
Assume true for $n \leq y$.

We must now prove it is true for $n = y + 1$, i.e. $P_{y+1} = 2^{2^y - y - 1}$
A De Bruijn sequence of order n can be found from an Eulerian graph of order 2^{n-1}, so for a De Bruijn sequence where $n = y + 1$ we must consider an Eulerian graph of order 2^y.
N_{y-1} has order 2^{y-1} and can can be used to construct a De Bruijn sequence of order y
N_y has order 2^y and is therefore the 'doubled' version of N_{y-1}
$P = \mid N_y \mid = \mid N*_{y-1} \mid = 2^{m-1} \cdot \mid N_{y-1} \mid$
$= 2^{2^{y-1}-1} \cdot 2^{2^{y-1}-y}$
$= 2^{2 \cdot 2^{y-1}-y-1}$
$= 2^{2^y - y - 1}$

We have shown that the formula for P is correct for $n = 1$ and that if it is true for $n = y$, then it is also true for $n = y + 1$. Therefore it is true, by induction, for all $n \geq 1$.

This result was extended by T. van Aardenne-Ehrenfest and N. G. de Bruijn in [1] and there is now a formula for the number of distinct De Bruijn sequences for any values of k and n:

$$P = k!^{k^{n-1}-n}$$

2.5 Construction using Modular Arithmetic

An alternative way of finding a De Bruijn sequence when $k = 2$ was given in [6]. For a given n and starting with $a_1 = 2^n - 1$, a sequence of numbers can be generated by repeatedly substituting the previous number in the sequence into the formula:

$$a_{i+1} \equiv 2a_i \pmod{2^n}$$

If for some $i \leq j$, $a_i = 2a_j$, then in this case:

$$a_{i+1} \equiv 2a_i + 1 \pmod{2^n}$$

This means that if the first formula gives a number already generated, then add 1 to this number to obtain the next number in the sequence, substituting this new number back into the first equation to obtain further numbers in the sequence. All these numbers of the sequence should be written in binary form (to base 2 with 3 digits - put 0s before the number if there are fewer than 3 digits) and consecutive numbers in the sequence will join together to form a De Bruijn sequence $B(2, n)$.
For $n = 3$, we have mod $2^3 = 8$. This means that if the number ≥ 8, we take the remainder when it is divided by 8. Another way to think of this would be a clock with 8 hours on the face - when it is 9 o'clock the clock face looks the same as it does when it is 1 o'clock i.e. 9=1 (mod 8).
$a_1 = 2^3 - 1 = 8 - 1 = 7 = 111$(in binary form) $\quad a_2 = 2 \times 7 = 14 = 6 = 110$
$a_3 = 2 \times 6 = 12 = 4 = 100$ (here we can do 2×6 instead 2×14 as it is simpler and both calculations give the same answer mod 8)
$a_4 = 2 \times 4 = 8 = 0 = 000$ $\quad a_5 = 2 \times 0 = 0$ which has already been generated $\Rightarrow a_5 = 0 + 1 = 1 = 001$
$a_6 = 2 \times 1 = 2 = 010$ $\quad a_7 = 2 \times 2 = 4$ which has already been generated $\Rightarrow a_7 = 4 + 1 = 5 = 101$
$a_8 = 2 \times 5 = 10 = 2$ which has already been generated $\Rightarrow a_8 = 2 + 1 = 3 = 011$
We can now stop as we have obtained the first 2^n numbers in the sequence needed to form a De Bruijn sequence. Putting these numbers together and omitting the overlapping digits between consecutive numbers (including the last 2 digits which overlap with the first 2 digits) gives the sequence $B(2, 3)$: 11100010.

Why does this method work?
Effectively, for a given n, we are generating the numbers 0 to $2^n - 1$ in a special order which means that when they are written in binary form, the first $n-1$ digits of a number are the same as the last $n-1$ digits of the number generated before it. This property must hold in order for the construction of a De Bruijn sequence to be possible. When using this method only the first 2^n numbers need to be generated as the De Bruijn sequence will have length 2^n. Because we are using modular arithmetic mod 2^n and do not allow repeated numbers to be counted in the sequence, this means the first 2^n numbers generated must be the numbers 0 to $2^n - 1$. Multiplying a_i by 2 to find a_{i+1} shifts the last $n-1$ digits of a_i one place to the left (like multiplying by 10 in base 10), so consecutive numbers generated by the first formula always overlap by $n-1$ digits. The first formula alway results in a number ending in '0' so to obtain alternative endings i.e. a last digit of '1' we add '1' to $2a_i$ (formula 2). This does not affect the first $n-1$ digits of the number, so it still overlaps with the previous number generated.

We must be careful with the choice of a_1 because otherwise we may need to add '1' twice consecutively to a_i to get a number which has not previously been generated. This would mean the number generated would not overlap with the previous number in the sequence. Let us consider the largest number to be generated: $2^n - 1$, which is a number consisting of n '1's when written in binary form. If the integer $2^n - 1$ is not at the beginning, then it must be preceded by $2^{n-1} - 1$ (binary form: a '0' followed by $n-1$ '1's) and $2^n - 2$ (binary form: $n-1$ '1's followed by a '0') must have already occurred in the sequence so that we add '1' using formula 2. However, the number $2^n - 2$ must be preceded by either $2^n - 1$ or $2^{n-1} - 1$. This means that for this method of construction to work, we must have $a_1 = 2^n - 1$ or $a_1 = 2^n - 2$. If we were to start with the second option for a_1 rather than the first, we would obtain the same sequence of numbers, except that the number $2^n - 1$ would be the last number generated as opposed to the first. Both values for a_1 therefore give the same De Bruijn cycle.

For $n = 4$, we have mod $2^4 = 16$.
$a_1 = 2^4 - 1 = 16 - 1 = 15 = 1111$ $\quad a_2 = 2 \times 15 = 30 = 14 = 1110$
$a_3 = 2 \times 14 = 28 = 12 = 1100$ $\quad a_4 = 2 \times 12 = 24 = 8 = 1000$
$a_5 = 2 \times 8 = 16 = 0 = 0000$ $\quad a_6 = 2 \times 0 = 0 \Rightarrow a_6 = 0 + 1 = 1 = 000$
$a_7 = 2 \times 1 = 2 = 0010$ $\quad a_8 = 2 \times 2 = 4 = 0100$
$a_9 = 2 \times 4 = 8 \Rightarrow a_9 = 8 + 1 = 9 = 1001$ $\quad a_10 = 2 \times 9 = 18 = 2 \Rightarrow a_10 = 2 + 1 = 3 = 001$
$a_{11} = 2 \times 3 = 6 = 0110$ $\quad a_{12} = 2 \times 6 = 12 \Rightarrow a_{12} = 12 + 1 = 13 = 110$
$a_{13} = 2 \times 13 = 26 = 10 = 1010$ $\quad a_{14} = 2 \times 10 = 20 = 4 \Rightarrow a_{14} = 4 + 1 = 5 = 010$
$a_{15} = 2 \times 5 = 10 \Rightarrow a_{15} = 10 + 1 = 11 = 1011$ $\quad a_{16} = 2 \times 11 = 22 = 6 \Rightarrow a_{16} = 6 + 1 = 7 = 011$
We now have the 2^n numbers needed to form a De Bruijn sequence $B(2, 4)$: 1111000010011010.

Does this method work for $k \geq 3$?
To see if modular arithmetic could be used to construct De Bruijn sequences with larger alphabets i.e. $k \geq 3$, I tried using this method to construct $B(3, 2)$ by starting with $a_1 = 3^n - 1$, writing numbers generated in the sequence in ternary form and adapting the previous formulae:

$$a_{i+1} \equiv 3a_i \pmod{3^n}$$

If for some $i \leq j$, $a_i = 3a_j$, then in this case:

$$a_{i+1} \equiv 3a_i + 1 \pmod{3^n}$$

For $k = 3, n = 2$, we have mod $3^2 = 9$.
$a_1 = 3^2 - 1 = 9 - 1 = 8 = 22$ $a_2 = 3 \times 8 = 24 = 6 = 20$
$a_3 = 3 \times 6 = 18 = 0 = 00$ $a_4 = 3 \times 0 = 0 \Rightarrow a_4 = 0 + 1 = 1 = 01$
$a_5 = 3 \times 1 = 3 = 10$ $a_6 = 3 \times 3 = 9 = 0 \Rightarrow a_6 = 0 + 1 = 1 \Rightarrow a_6 = 1 + 1 = 2 = 02$
$a_7 = 3 \times 2 = 6 \Rightarrow a_7 = 6 + 1 = 7 = 21$ $a_8 = 3 \times 7 = 21 = 3 \Rightarrow a_8 = 3 + 1 = 4 = 11$
$a_9 = 3 \times 4 = 12 = 3 \Rightarrow a_9 = 3 + 1 = 4 \Rightarrow a_9 = 4 + 1 = 5 = 12$
We now have the 3^n numbers needed to form a De Bruijn sequence $B(3,2)$: 220010211. Therefore the method has been successful in producing a De Bruijn sequence when $k \neq 2$. Note how in this example, we were able to add '1' twice consecutively, whereas in the previous example we only added '1' once to obtain a particular number in the sequence. This is because in this example we wrote the numbers in base 3, so could add '1' or '2' after multiplying by 3 without changing the first 2 digits of the final number generated. This meant that the first 2 digits of this number would still overlap with the final 2 digits of the previous number in the sequence. Adding 3 in this situation would have changed the second digit in the sequence, meaning it would no longer fully overlap with the previous number (a property required for the formation of a De Bruijn sequence), but there was no need to do this.

Remark 3 I believe that this method could be extended for all values of k, with $a_1 = k^n - 1$ and:

$$a_{i+1} \equiv ka_i \pmod{k^n}$$

Unless for some $i \leq j$, $a_i = ka_j$, then in this case:

$$a_{i+1} \equiv ka_i + 1 \pmod{k^n}$$

However, a disadvantage of this method is that only one De Bruijn sequence for a given n and k can be formed, whereas the method of construction using Eulerian graphs can be used to find all possible distinct De Bruijn cycles by finding all the different Eulerian circuits around the graph.

2.6 Applications of De Bruijn sequences

De Bruijn sequences have many uses, including in the positioning of robots. If a robot is moving along a track marked with a De Bruijn sequence e.g.$B(2,3)$, then by looking at the nearest 3 numbers in the sequence, the robot can determine its location on the track as each triplet is unique.

Eulerian graphs were used to solve the famous problem the Seven Bridges of Konigsberg. The Russian city of Konigsberg is located on the River Pregel. The city has 2 large islands, which were connected to the mainland area by 7 bridges. The people of the city often wondered whether it was possible to cross each of the 7 bridges exactly once on a single route, but Leonhard Euler proved that this was impossible. By drawing each land mass as a vertex of a graph and the bridges as edges of the graph, he showed that the resulting graph was not Eulerian and so crossing each bridge exactly once was not possible.

De Bruijn sequences can be used to minimise the effort needed to guess a code in locks that do not have an 'enter' key but instead accept the last n digits entered. For example, to try all the possible combinations of a 4 digit PIN like code, a De Bruijn sequence $B(10,4)$ could be used. This universal cycle would have length 10^4 and entering digits in the order given by the cycle would require only $10^4 + 3 = 10003$ presses, whereas trying all the possible combinations for the code separately would take $10^4 \times 4 = 40000$ presses.

De Bruijn sequences also have applications in generating sequences in DNA, with an alphabet consisting of the 4 types of nucleotide which make up DNA: adenine(A), thymine(T), cytosine(C), guanine(G).

The uses of De Bruijn sequences in card tricks are given in the section on multiplying universal cycles.

3 Permutations

When considering permutations, it is the relative size of each digit which we are concerned with rather than the actual value of a digit. For example, for permutations of 3 digits, we can think of each of the 3 digits as being low (L), medium (M) or high (H). In this case the number of permutations would be $3! = 6$ and can be expressed by the letters above as LMH, LHM, MLH, MHL, HLM, HML, which correspond to the number permutations 123, 132, 213, 231, 312, 321. For permutations of n objects, there are $n!$ distinct permutations as we have a choice of n values for the first digit, $n-1$ values for the second digit and so on until there is only one possible number left for the nth digit.

3.1 Universal Cycles for Permutations

We say that the n-tuples a and b are order-isomorphic, written $\bar{a} \sim \bar{b}$ if $a_i < a_j \Leftrightarrow b_i < b_j$. Universal cycles for permutations are cycles of length $n!$ where each of the $n!$ permutations of n distinct integers is order-isomorphic to exactly one n-tuple in the cycle. That is to say, any n consecutive digits in the cycle will have a distinct relative size order.
A universal cycle for permutations of n digits where $n \geq 3$ will consist of at least $n+1$ digits. This is because if, for example, we try constructing a universal cycle for $n = 3$ using just 3 digits:
123 $\leftarrow$ the next digit must be 1 as each symbol in a window of length 3 must be different
1231 $\leftarrow$ the next digit must be 2
12312 $\leftarrow$ the next digit must be 3, but this means we have repeated the first window of 3 symbols, so this is not a universal cycle.
The same problem arises if we try constructing a universal cycle for any $n \geq 3$ with just n symbols – the sequence always ends up repeating itself, so this cannot be done. It was proved in [5] that for $n \geq 3$, it is always possible to find a universal cycle for permutations using exactly $n+1$ symbols.

The method of construction of universal cycles for permutations is similar to that used to form De Bruijn sequences as graphs are used. This time we use the $n!$ permutations of n integers as the vertices of the graph. This graph is called a transition graph. Next we work out which vertices are connected by drawing arrows from each vertex to the vertices where the last $n-1$ digits of the original vertex and the first $n-1$ digits of the second vertex are order-isomorphic.
When $n = 3$, there are $3! = 6$ permutations and so the graph has 6 vertices. We can start with any 3 numbers (even non-integers) as we are only concerned with the relative size of the integers. e.g. $679, 679 \sim 123$. We consider the possible size of the next digit by looking at the last $n-1 = 3-1 = 2$ digits: $79x$. x could lie in $n-1+1 = n = 3$ different ranges:
1. $x < 7 < 9 \rightarrow 79x \sim 231$
2. $7 < x < 9 \rightarrow 79x \sim 132$
3. $7 < 9 < x \rightarrow 79x \sim 123$

This means there are n arrows leaving and n arrows entering each vertex and a total of $n \cdot n!$ arrows in the transition graph.

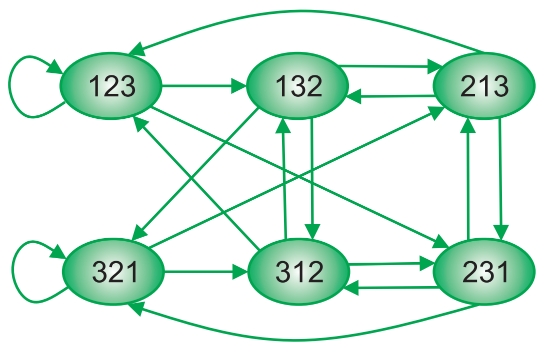


Figure 5: Transition graph for $n = 3$

To form a universal cycle, we would need to take a Hamiltonian circuit in the graph (a path which visits each vertex exactly once, returning to the starting vertex). However, Hamiltonian circuits are difficult to find, especially for more complicated graphs and there is no way of knowing if such a path exists without trying out all the possible paths. To avoid this problem, we convert the transition graph to an Eulerian graph by grouping together permutations where the first $n-1$ digits are order-isomorphic into new larger vertices. There should be $(n-1)!$ vertices in the new graph.

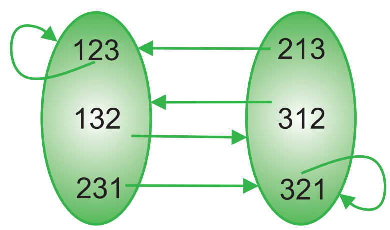
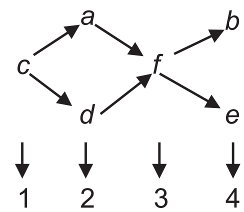


Figure 6: Eulerian graph for $n = 3$

In the Eulerian graph, an arrow from a permutation to a vertex means that it is possible to travel from that permutation to all the permutations in the vertex.

Next we find an Eulerian circuit in the graph, e.g. 231, 312, 123, 132, 321, 213, 231. Clearly, the maximum number of symbols we would need for a sequence of length $n!$ is $n!$ and for this example we will use the $3! = 6$ letters: $abcdef$. By running windows of length n along the sequence of letters and comparing this to the permutations of length 3 in the Eulerian circuit, inequalities can be written down to compare the relative sizes of the letters.
$abc : 231 \rightarrow c < a < b$
$bcd : 312 \rightarrow c < d < b$
$cde : 123 \rightarrow c < d < e$
$def : 132 \rightarrow d < f < e$
$efa : 321 \rightarrow a < f < e$
$fab : 213 \rightarrow a < f < b$
Combining these inequalities together gives:

a and d can be combined and b and e can be combined whilst still satisfying all the inequalities,



so only 4 symbols were needed for the permutation cycle (which is the least possible number as we need at least $n+1 = 3+1 = 4$ symbols). Finally we write down the numbers corresponding to the sequence $abcdef$ to obtain the universal cycle for permutations of length 3: 241243.
This combining of letters is not always possible and sometimes we are forced to use more than $n+1$ symbols. For the Eulerian circuit: 312, 231, 213, 123, 132, 321, 312 we have the inequalities:
$abc : 312 \to b < c < a$
$bcd : 231 \to d < b < c$
$cde : 213 \to d < c < e$
$def : 123 \to d < e < f$
$efa : 132 \to e < a < f$
$fab : 321 \to b < a < f$

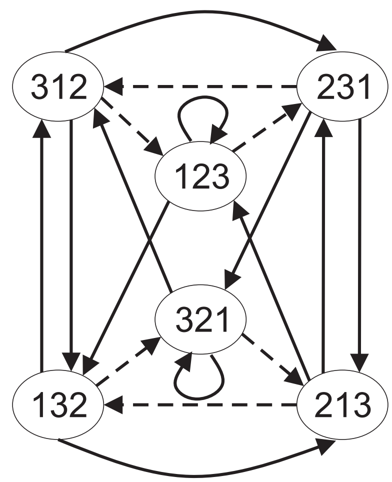
Combining these inequalities together gives:
$d < b < c < e < a < f$
so in this case none of the letters can be combined and we get the universal cycle: 523146.

Remark 4 I found several Eulerian circuits in the graph and converted these into universal cycles for permutations when $n = 3$ and an interesting pattern began to emerge. As has already been mentioned, the number of symbols used in a universal cycles for $n = 3$ is a minimum of 4 and a maximum of 6. However, I noticed that there did not seem to be any cycles consisting of exactly 5 different symbols.

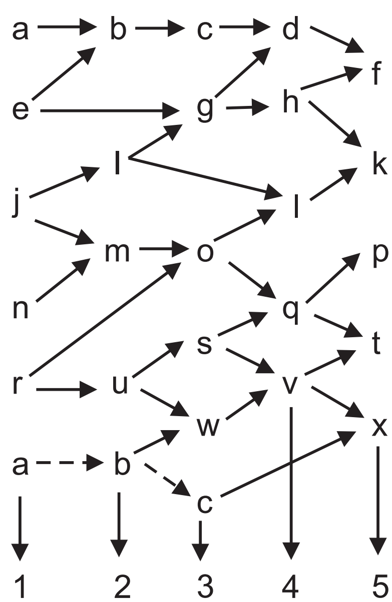
Looking at this problem in more detail, I found that the Eulerian circuits which led to a cycle using 4 symbols had 4 of what I thought of as 'good' connections and circuits with 4 'good' connections always led to a cycle with 4 symbols. These 'good' connections occurred when the final 2 digits of the first permutation were exactly the same as the first 2 digits of the second permutation as opposed to being just order-isomorphic. These 'good' connections seemed to lead to fewer symbols being needed for the cycle as it meant some of the 6 letters initially used could be combined. I also found that the Eulerian circuits which led to a cycle using 6 symbols had only 1 'good' connection and circuits with 1 'good' connection always led to a cycle with 6 symbols. I thought that perhaps Eulerian circuits with 2 or 3 'good' connections would lead to a universal cycle with 5 symbols. However, I was unable to find such a circuit in the graph and therefore think that perhaps it is not possible to obtain a 5 symbol cycle for permutations (obviously this excludes the case where one of the number 4s in a cycle with 4 symbols is simply changed to a bigger number).

In the transition graph below, the 'good' connections are those with dashed arrows.

This method can be used to construct universal cycles for permutations of 4 or more objects. For $n = 4$, the transition graph has $3! = 6$ vertices and the permutations where the first $n-1 = 4-1 = 3$ digits are order-isomorphic are grouped in the same vertex. Arrows are drawn in the same way as before, with a single arrow to a vertex representing paths to all the permutations in that vertex. In this case, an Eulerian circuit would be: 1234, 2341, 2314, 3142, 1423, 4231, 3421, 3214, 2143, 1432,



4321, 4213, 2134, 1243, 2431, 4312, 3124, 1342, 2413, 4132, 1324, 3241, 3412, 4123, 1234. Using this to write inequalities gives:



This in turn gives the universal cycle: 12341534215421354135243 5. This cycle using the smallest possible number of different symbols for $n = 4$: 5.

[5] gave an alternative method of constructing universal cycles for permutations which guaranteed that the resulting cycle had the smallest possible number $(n + 1)$ of different symbols. This involves constructing sub cycles where the permutations had what I previously called 'good' connections. These sub cycles are then connected together to form a cycle of $n!$ permutations which can then be converted into a universal cycle of length $n!$.

4 Multiplying Universal Cycles

Multiplying together 2 universal cycles x and y which have lengths R and S respectively is possible when one of the cycles with window length k finishes with a block of k repeated symbols. The method for doing this was given in [4]. The 2 universal cycles being multiplied together need not have the same window length but in most cases they do. Multiplying these cycles together gives a cycle of 'pairs' of length RS, with values from x on the top row and values of y on the bottom. If both universal cycles have window length k, each k consecutive pairs in the product cycle will be unique (i.e. a unique combination of the top and bottom rows). The cycles multiplied together can be of different types, e.g. a De Bruijn sequence can be multiplied with a universal cycle for permutations. When a permutation cycle for 3 objects: 241243 is multiplied by a De Bruijn sequence for binary strings of length 3: 11101000, we get:

2	4	1	2	4	3	2	4	1	2	4	3	2	4	1	2	4	3	2	4	1	2	4	3
1	1	1	0	1	0	0	0	1	1	1	0	1	0	0	0	1	1	1	0	1	0	0	1

2	4	1	2	4	3	2	4	1	2	4	3	2	4	1	2	4	3	2	4	1	2	4	3
1	1	0	1	0	0	0	1	1	1	0	1	0	0	0	1	1	1	0	1	0	0	0	0

Running a window of length 3 along the product cycle gives a unique coupling of a low, medium, high permutation with a binary string of length 3. This product cycle can be used in card tricks. It has a length of 48 so the 4 kings must be removed from a deck of 52 cards. Next, the permutation cycle for window length k on the top row of the product is 'lifted' so that it contains 6 distinct symbols rather than just 4 (lifting is the process of increasing the alphabet of a universal cycle whilst retaining the cycle's properties). This is done by finding the highest digit (it may appear several times) and changing one appearance of this digit to the number $k!$. Then we take the next highest digit (not including the $k!$ digit) and change it to the digit $k! - 1$. This process continues until there are as many distinct digits in the cycle as required.

For example, a universal cycle for permutations of 3 objects is: 241243. Working from left to right in the case of the same digit appearing more than once, we get: 261243 $\rightarrow$ 261253 $\rightarrow$ 261254 $\rightarrow$ 361254 (which now contains 6 different digits). We now look at 2 adjacent copies of this cycle: 361254361254. Each digit '1' can be assigned either '1' or '2', each digit 'n' can be assigned either the number '$2n$-1' or the number '$2n$' up until the digit '6' can be assigned either '11' or '12'. Choosing the lower number first in each case gives:

3	6	1	2	5	4	3	6	1	2	5	4
5	11	1	3	9	7	6	12	2	4	10	8

(We can change whether we choose the lower or higher of the 2 numbers to make the sequence look more random.) In the product cycle above, this pattern is repeated 4 times so each of the 12 cards in each suit is included exactly once, with red cards (Hearts and Diamonds) for '1's in the De Bruijn sequence on the bottom row and black cards (Clubs and Spades) for '0's in the bottom row. An example of a sequence which could be used for a card trick would be:

5	11	1	3	9	7	6	12	2	4	10	8	5	12	2	3	9	8	6	11	1	4	10	7
D	H	H	S	D	C	C	C	D	D	D	S	H	S	S	C	H	D	H	C	D	S	S	D

6	12	1	4	10	8	5	11	2	3	9	7	5	11	2	3	10	8	6	12	1	4	9	7
D	H	C	H	C	C	S	D	H	H	S	H	C	S	C	D	H	H	S	D	S	C	C	S

By asking people to choose 3 adjacent cards in the sequence and finding out which cards are red, as well as which cards are the highest and lowest, we can deduce the value and suit of each of the 3 cards chosen by referring to the original product sequence and then the lifted sequence.

Consider 2 cycles x and y of lengths R and S. If R and S are relatively prime (have a highest common factor of 1), then to multiply the 2 cycles, simply write down x S times and below write down y R times. For example, to multiply a universal cycle for partitions of set of 4 numbers (x): *daabbbbcbccbadb* ($R = 15$) with a De Bruijn sequence for binary strings of length 4 (y): 1111001011010000 ($S = 16$), we simply write out x 16 times and y 15 times as 15 and 16 are co-prime. This gives the product consisting of $15 \times 16 = 240$ pairs:

$$\begin{matrix} d & a & a & b & b & b & b & c & b & c & c & b & a & d & b & d & a & a & b & b & b & b & c & b & c & c & b & a & d & b\ldots \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0\ldots \end{matrix}$$

Running a window of length 4 along this product gives a unique partition of a set of 4 elements plus binary string of length 4 combination.

If R and S are not relatively prime, and their highest common factor is d ($d \neq 1$), then we have: $R = rd$ and $S = sd$ where r and s are co-prime. To obtain a cycle of RS pairs:

1. write down x S times
2. write down y r times
3. remove the last repeated digit in the repeated y sequence, forming the sequence $y*$
4. write down $y*$ a total of d times below the repeated sequence of x
5. add d of the digit removed to the end of the repeated $y*$ sequence

This gives a sequence of length: $(Sr - 1) \cdot d + d = Srd - d + d = Srd = RS$ as required.

For example, to multiply a De Bruijn sequence for binary strings of length 2: 0011 by itself, $R = S = 4$, $d = 4$, $R = S = 1 \times d$. First, we write down x 4 times:

$$\begin{matrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \end{matrix}$$

Next, write down y once and remove the last '1' to form $y*$:

$$\begin{matrix} 0 & 0 & 1 \end{matrix}$$

Finally, write down $y*$ 4 times below the repeated sequence of x and add 4 '1's at the end to give the product sequence:

$$\begin{matrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{matrix}$$

Running a window of length 2 along the product gives 16 unique 2×2 squares.

This process can be continued and the product cycle can be multiplied by another universal cycle, for example multiplying the De Bruijn sequence for $n = 1$: 01 by itself gives:

$$\begin{matrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{matrix}$$

Multiplying the product by 01 again gives:

0	1	0	1	0	1	0	1
0	0	1	1	0	0	1	1
0	1	0	0	1	0	1	1

Multiplying the product by 01 again gives:

0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
0	1	0	0	1	0	1	1	0	1	0	0	1	0	1	1
0	1	0	1	0	1	0	0	1	0	1	0	1	0	1	1

and so on.

5 Partitions of a Set

A partition of any set A consisting of n elements $\{1\ 2\ 3\ 4\ \ldots n\}$ is formed when A is divided into subsets which are mutually exclusive (have no elements in common) and collectively contain all the elements in the set. Each subset must be non-empty (must contain at least 1 element from the set). A partition of a set can be expressed by using vertical bars to show the separation of elements of the set into subsets. The subsets are independent of order, meaning the order in which the separate subsets of a partition are written and the order of the elements within a subset do not matter. For example, the set $A\{1\ 2\ 3\ 4\}$ has 15 partitions, which can be written as:

1 |234	1 |2 |34	1 |2 |3 |4	1234
2 |134	1 |3 |24		
3 |124	1 |4 |23		
4 |123	2 |3 |14		
12 |34	2 |4 |13		
13 |24	3 |4 |12		
14 |23			

The number of partitions of a set with n elements is given by the Bell numbers B_n.

5.1 Universal Cycles of Partitions of a Set

Universal cycles of partitions of a set with n elements can be formed. These universal cycles are very different from those formed for permutations and De Bruijn sequences. They usually consist of a sequence of B_n letters. When a window of length n is moved along the sequence, we number each letter in the window 1 to n from left to right. If the letters of 2 or more numbers are the same, this means these numbers are in the same subset in the partition. This means that if 2 numbers correspond to different letters in the sequence, these 2 numbers are in separate subsets. The method used to construct these universal cycles is similar to that used to produce universal cycles of permutations and De Bruijn sequences as we must again use Eulerian graphs.

First, we construct a transition graph by writing down the partitions of the set as the vertices of the graph. Arrows are drawn from one vertex to another vertex when the relationship between the numbers 2 to n of the first vertex is the same as the relationship between the numbers 1 to $n-1$ in the second vertex (in terms of whether or not they are in the same subset of the partition).

For the case where $n = 3$, we are constructing a cycle for partitions of the set $\{1\ 2\ 3\}$. There are 5 partitions of this set:

1 |23	1 |2 | 3
2 |12	123
3 |12	

These partitions are the 5 vertices of the transition graph. The vertex 123 would be represented by the letters *aaa* as all 3 numbers are in the same partition and when drawing arrows from this vertex we look at the the numbers 2 to $n = 3$ which are represented by *aa*. This can be followed by either *aaa* or *aab* which correspond to the vertices 123 and 12 |3 so arrows can be drawn from the original vertex to these vertices. It is important to note that the actual letters that represent a partition are not of great importance - it is whether these letters are the same or different that matters. For example, an arrow could be drawn from 1 |2 |3 (*abc*) to 1 |23 (*abb*).

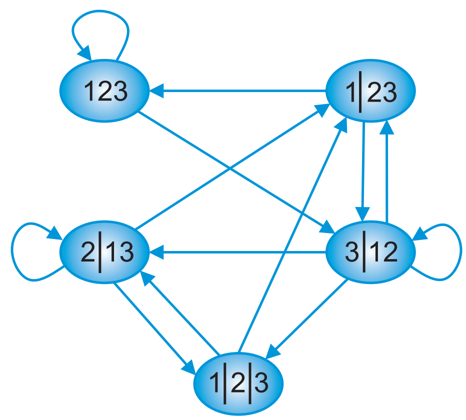


Figure 7: Transition graph for partitions of a set with 3 elements

This transition graph is then converted to another graph by grouping together partitions where the first $n - 1$ numbers have the same relationship to form a large single vertex. This means the new graph will have B_{n-1} vertices. Once again an arrow from a partition to a vertex means that it is possible to travel from that permutation to all the permutations in the vertex.

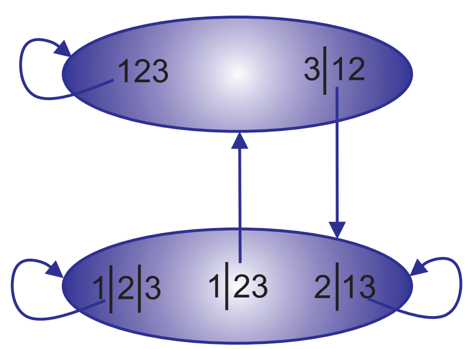


Figure 8: Eulerian graph for partitions of a set with 3 elements

Next, we must find an Eulerian circuit in this graph. For $n = 3$ the only 2 possible circuits are:
123, 3 |12, 2 |13, 1 |2 |3, 1 |23, 123 and
123, 3 |12, 1 |2 |3, 2 |13, 1 |23, 123.

Finally we need to 'lift' the Eulerian circuit to a universal cycle. This is done by assigning x_{B_n} symbols to the numbers in the partitions and writing equalities to show whether or not these symbols are equal (i.e. whether they are in the same subset of a partition). For each partition, the numbers are rearranged in increasing order, then we look at whether or not their corresponding symbols should be equal.

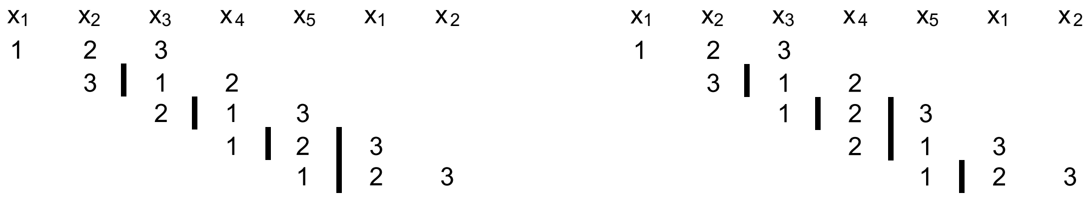


Figure 9: 'Lifting' Eulerian circuits for $n = 3$

For the first Eulerian circuit, we obtain the following set of equalities:
$x_1 = x_2 = x_3$
$x_2 = x_3 \neq x_4$
$x_3 = x_5 \neq x_4$
$x_4 \neq x_5$ $\quad x_5 \neq x_1$ $\quad x_1 \neq x_4$
$x_1 = x_2 \neq x_5$

However, this leads to a contradiction, as we can deduce that $x_5 \neq x_1 = x_3 = x_5$. Therefore we conclude that this particular Eulerian circuit cannot be 'lifted' successfully to form a universal cycle.

For the second Eulerian circuit, we obtain the following set of equalities:
$x_1 = x_2 = x_3$
$x_2 = x_3 \neq x_4$
$x_3 \neq x_4$ $\quad x_4 \neq x_5$ $\quad x_5 \neq x_3$
$x_4 = x_1 \neq x_5$
$x_1 = x_2 \neq x_5$

However, this also leads to a contradiction, as we can deduce that $x_4 = x_1 = x_3 \neq x_4$. We conclude that this Eulerian circuit cannot form a universal cycle. As there are only 2 Eulerian circuits for the graph of partitions of a set of 3 elements and we have shown that neither can be used to form a universal cycle, we have proved by exhaustion that no universal cycles exist for partitions of a set of 3 elements.

Universal cycles for partitions of a set of 4 elements do exist and can be constructed by skipping the transition graph and drawing the Eulerian graph straight away and combining the partitions in which the numbers 1 to $n - 1 = 4 - 1 = 3$ have the same relationship into a single large vertex. Drawing arrows between the vertices in the same way as in the previous example gives an Eulerian graph with 5 vertices and 15 arrows.

In [3] it was stated that in order to prevent equalities leading to a contradiction and to guarantee an Eulerian circuit can be 'lifted' to form a universal cycle, a sequence of partitions called a 'breaker' must occur in the Eulerian circuit. Using the 'breaker' they gave: 1 |4 |23, 12 |34, 1 |234, 1234, 4 |123, I found an Eulerian circuit for the graph and used it to produce a universal cycle for $n = 4$. The Eulerian circuit was:

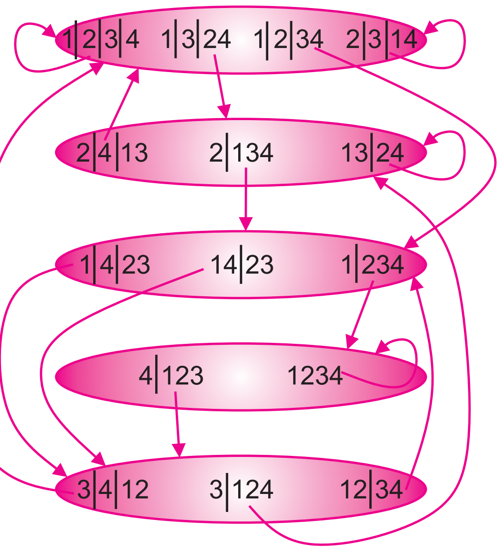


Figure 10: Eulerian graph for partitions of a set with 4 elements

1 |4 |23, 12 |34, 1 |234, 1234, 4 |123, 3 |124, 13 |24, 2 |134, 14 |23, 3 |4 |12, 1 |2 |3 |4, 2 |3 |14, 1 |3 |24, 2 |4 |13, 1 |2 |34, 1 |4 |23.

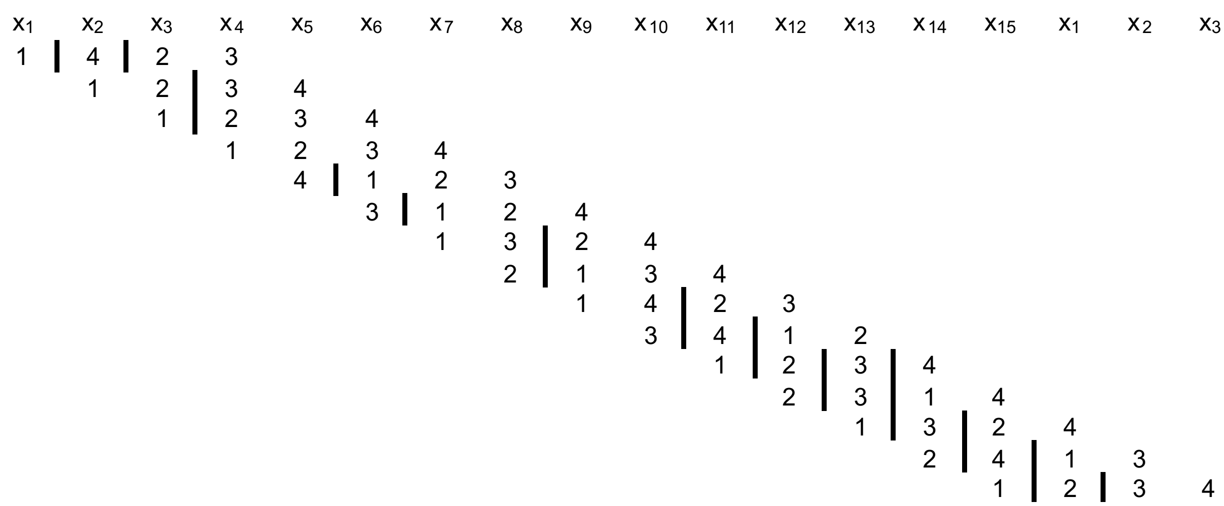


Figure 11: ‘Lifting’ an Eulerian circuit for $n = 4$

This gave the set of equalities:

$x_1 \neq x_2 = x_3 \neq x_4 \qquad x_1 \neq x_4$
$x_2 = x_3 \neq x_4 = x_5$
$x_3 \neq x_4 = x_5 = x_6$
$x_4 = x_5 = x_6 = x_7$
$x_5 = x_6 = x_7 \neq x_8$
$x_6 = x_7 = x_9 \neq x_8$
$x_7 = x_9 \neq x_8 = x_{10}$
$x_8 = x_{10} = x_{11} \neq x_9$

$x_9 = x_{12} \neq x_{10} = x_{11}$
$x_{10} = x_{11} \neq x_{12} \neq x_{13}$ $\quad x_{10} = x_{11} \neq x_{13}$
$x_{11} \neq x_{12} \neq x_{13} \neq x_{14}$ $\quad x_{11} \neq x_{13}$ $\quad x_{11} \neq x_{14}$ $\quad x_{12} \neq x_{14}$
$x_{12} = x_{15} \neq x_{13} \neq x_{14}$ $\quad x_{12} = x_{15} \neq x_{14}$
$x_{13} \neq x_{14} = x_1 \neq x_{15}$ $\quad x_{13} \neq x_{15}$
$x_{14} = x_1 \neq x_{15} \neq x_2$ $\quad x_{14} = x_1 \neq x_2$
$x_2 = x_3 \neq x_{15} \neq x_1$ $\quad x_2 = x_3 \neq x_1$

I assigned a letter to each group of x symbols which were equal:
$a : x_2 = x_3$
$b : x_4 = x_5 = x_6 = x_7 = x_9 = x_{12} = x_{15}$
$c : x_8 = x_{10} = x_{11}$
$d : x_{14} = x_1$
$e : x_{13}$
I found that e could be combined with a whilst still satisfying all the equalities. This meant that the universal cycle for $n = 4$ would only contain 4 letters as opposed to 5. Finally, I wrote out the letters corresponding to each x symbol when the x symbols were arranged in order. This gave the universal cycle: *daabbbbcbccbadb*.

Remark 5 Based on the properties of the 'breaker' given in [3], I tried to construct a different 'breaker' for $n = 4$.
The original 'breaker' has the properties:

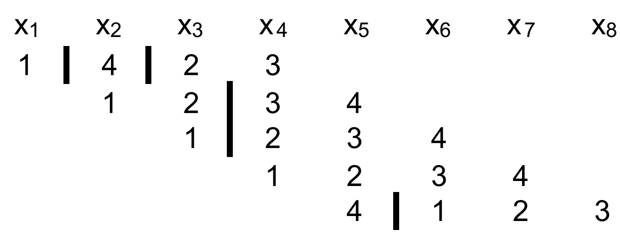


Figure 12: Properties of the original 'breaker'

$x_1 \neq x_2 = x_3 \neq x_4$ $\quad x_1 \neq x_4$
$x_2 = x_3 \neq x_4 = x_5$
$x_3 \neq x_4 = x_5 = x_6$
$x_4 = x_5 = x_6 = x_7$
$x_5 = x_6 = x_7 \neq x_8$

In [3], it is claimed that the inclusion of such a 'breaker' in an Eulerian circuit will always lead to a circuit which can be 'lifted' to form a universal cycle. The reason given for this is that a x symbol occurring before x_1 e.g. x_{15} cannot be forced to be either equal or unequal to an x symbol after x_7. This is true when looking at the 'breaker' alone. However, the partitions form a circuit, so when the full circuit of partitions are written out, it will be possible to deduce whether or not most, if not all the x symbols are equal or unequal. For example, in the construction of a universal cycle for $n = 4$ above, it was possible to deduce from the full set of equalities that $x_{15} \neq x_8$. Therefore it would seem that this claim is not true.

I constructed the following 'breaker' which has the same properties as the original 'breaker', except the equalities are reversed:
1 |234, 1234, 4 |123, 12 |34, 1 |4 |23

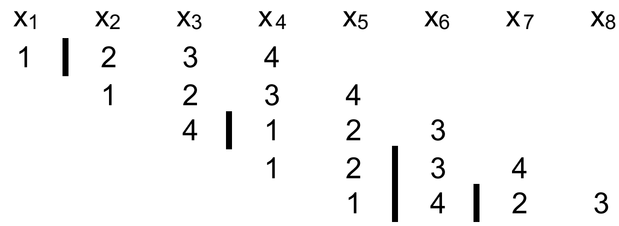


Figure 13: Properties of different 'breaker'

$x_1 \neq x_2 = x_3 = x_4$
$x_2 = x_3 = x_4 = x_5$
$x_3 = x_4 = x_5 \neq x_6$
$x_4 = x_5 \neq x_6 = x_7$
$x_5 \neq x_6 = x_7 \neq x_8 \qquad x_5 \neq x_8$

This 'breaker' was successful when used in the Eulerian circuit:
1 |234, 1234, 4 |123, 12 |34, 1 |4 |23, 3 |124, 13 |24, 2 |4 |13, 1 |3 |24, 2 |134, 14 |23, 3 |4 |12, 1 |2 |3 |4, 2 |3 |14, 1 |2 |34, 1 |234
giving the universal cycle: *daaaabbcbcaccab*.
However, when put in the Eulerian circuit:
1 |234, 1234, 4 |123, 12 |34, 1 |4 |23, 3 |4 |12, 2 |3 |14, 1 |2 |3 |4, 1 |3 |24, 2 |4 |13, 1 |2 |34, 14 |23, 3 |124, 13 |24, 2 |134, 1 |234
this circuit could not be 'lifted' successfully.

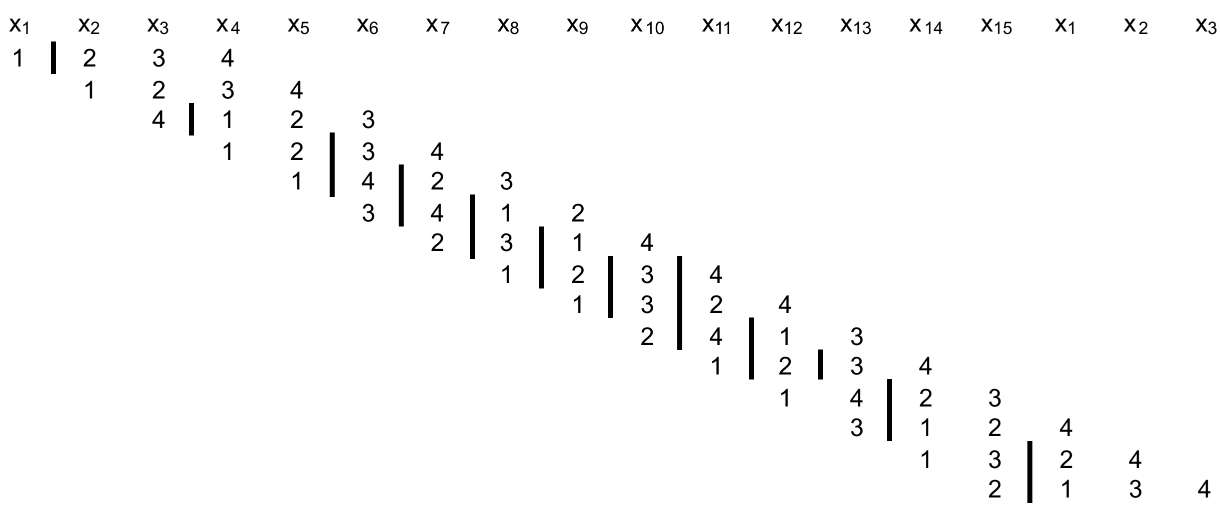


Figure 14: Failure using the new 'breaker'

Which gave the equalities:
$x_1 \neq x_2 = x_3 = x_4$

$x_2 = x_3 = x_4 = x_5$
$x_3 = x_4 = x_5 \neq x_6$
$x_4 = x_5 \neq x_6 = x_7$
$x_5 \neq x_6 = x_7 \neq x_8$ $x_5 \neq x_8$
$x_6 = x_7 \neq x_8 \neq x_9$ $x_6 = x_7 \neq x_9$
$x_7 = x_{10} \neq x_8 \neq x_9$ $x_7 = x_{10} \neq x_9$
$x_8 \neq x_9 \neq x_{10} \neq x_{11}$ $x_8 \neq x_{10}$ $x_8 \neq x_{11}$ $x_9 \neq x_{11}$
$x_9 \neq x_{10} = x_{12} \neq x_{11}$ $x_9 \neq x_{11}$
$x_{10} = x_{12} \neq x_{11} \neq x_{13}$ $x_{10} = x_{12} \neq x_{13}$
$x_{13} = x_{14} \neq x_{11} \neq x_{12}$ $x_{13} = x_{14} \neq x_{12}$
$x_{12} = x_{15} \neq x_{13} = x_{14}$
$x_{13} = x_{14} = x_1 \neq x_{15}$
$x_{14} = x_1 \neq x_{15} = x_2$
$x_1 \neq x_{15} = x_2 = x_3$

However, this leads to the contradiction that $x_5 \neq x_6 = x_7 = x_{10} = x_{12} = x_{15} = x_2 = x_3 = x_4 = x_5$. We conclude that this Eulerian circuit cannot form a universal cycle. The reversal of equalities compared to the original 'breaker' should not affect the 'breaker's' effectiveness, yet in this case it failed to work. This casts further doubt on the effectiveness of 'breakers' and it would seem that 'breakers' need to be defined in more detail to guarantee they always work.

5.2 Bell Numbers

We already know that the number of partitions of a set with n elements is given by the Bell numbers B_n. Consider a set of $n+1$ numbers: $\{1\ 2\ 3\ \ldots n{+}1\}$.
If $n+1$ is on its own in a subset, there are B_n ways to partition the remaining n elements in the set. This can be written as: $\binom{n}{0}B_n$ as by definition $\binom{n}{0}$ is equal to 1.
If $n+1$ is with 1 other element in a subset, there are $\binom{n}{1} = n$ ways of choosing this other element from the n available elements in the set and there are B_{n-1} ways of partitioning the remaining $n-1$ elements in the set. This means the number of partitions which contain such a subset can be written as: $\binom{n}{1}B_{n-1}$.
If $n+1$ is with 2 other elements in a subset, there are $\binom{n}{2}$ ways of choosing the other 2 elements and there are B_{n-2} ways of partitioning the remaining $n-2$ elements in the set. This means the number of partitions which contain such a subset can be written as: $\binom{n}{2}B_{n-2}$.
$\vdots$
There are $\binom{n}{n-1}B_1$ partitions which contain a subset where $n+1$ is with $n-1$ other elements in a subset.
There are $\binom{n}{n}B_0$ partitions which contain a subset where $n+1$ is with n other elements in a subset. (By definition, $B_0 = 1$, so there is only one partition containing all the elements in a single subset.)
Therefore the Bell numbers satisfy the recurrence relation:

$$B_{n+1} = \sum_{r=0}^{n} \binom{n}{n-r} B_r = \sum_{r=0}^{n} \binom{n}{r} B_r$$

$$\left(\text{due to the symmetry of combinations: } \binom{n}{n-r} = \frac{n!}{r!(n-r)!} = \binom{n}{r} \right)$$

Using this formula, we can calculate the first few Bell numbers:
$B_0 = 1$

$B_1 = B_0 = 1$
$B_2 = B_0 + B_1 = 1 + 1 = 2$
$B_3 = B_0 + 2B_1 + B_2 = 1 + 2 \times 1 + 2 = 1 + 2 + 2 = 5$
$B_4 = B_0 + 3B_1 + 3B_2 + B_3 = 1 + 3 \times 1 + 3 \times 2 + 5 = 1 + 3 + 6 + 5 = 15$
$B_5 = B_0 + 4B_1 + 6B_2 + 4B_3 + B_4 = 1 + 4 \times 1 + 6 \times 2 + 4 \times 5 + 15 = 1 + 4 + 12 + 20 + 15 = 52$
and so on

The Bell triangle can be used to generate Bell numbers (like Pascal's triangle can be used to generate binomial coefficients).

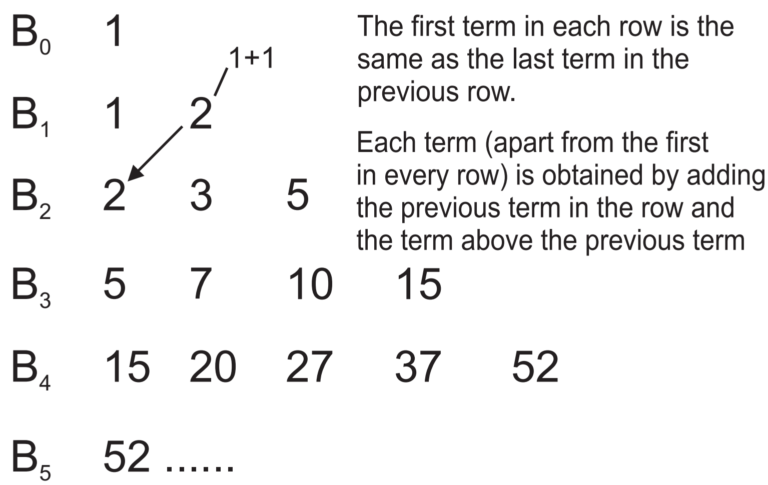


Figure 15: Finding Bell numbers using the Bell triangle

5.3 The Exponential Generating Function of Bell Numbers

A generating function is a formal power series where the coefficients of x^n give information about the numbers in a sequence a_n – an infinite series where the variable x is generally regarded as a place holder rather than assigned an actual value. Generating functions are very useful as they can represent sequences as functions and can be used to solve counting problems.

There are 2 main types of generating functions: ordinary generating functions and exponential generating functions.
The ordinary generating function $G(x)$ for an infinite sequence $(a_0, a_1, a_2, a_3 \ldots)$ would be:

$$G(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \ldots$$

The exponential generating function $E(x)$ for the same sequence would be:

$$E(x) = a_0 + \frac{a_1}{1!}x + \frac{a_2}{2!}x^2 + \frac{a_3}{3!}x^3 \ldots$$

In this case, there is an exponential generating function for the Bell numbers B_n.

Theorem 6 *The exponential generating function of the Bell numbers is e^{e^x-1}, i.e. the coefficient of $\frac{x^n}{n!}$ in the power series expansion of e^{e^x-1} is the number of partitions of a set of n elements.*

Proof We already know the recurrence relation for Bell numbers:

$$B_{n+1} = \sum_{r=0}^{n} \binom{n}{r} B_r \qquad B_n = \sum_{r=0}^{n-1} \binom{n-1}{r} B_r$$

To find B_1 (which is the coefficient of x), we differentiate the exponential generating function $y = E(x)$ and take the value of the function when $x = 0$.
The kth Bell number B_k should equal the kth derivative $y^{(k)}$ of e^{e^x-1} when $x = 0$.

When $n = 1$, we know that $B_1 = 1$

$$y = e^{e^x-1}$$

$$lny = ln(e^{e^x-1})$$

$$lny = (e^x - 1) \times 1$$

$$\frac{1}{y} \cdot \frac{dy}{dx} = e^x$$

$$\frac{dy}{dx} = ye^x$$

When $x = 0$,

$$\frac{dy}{dx} = e^{e^x-1} \cdot e^x = e^{e^0-1} \cdot e^0 = e^0 \cdot e^0 = 1 \cdot 1 = 1$$

$\Rightarrow$ true for $n = 1$
Assume true for $n = k$, i.e.

$$y^{(k)} = e^x \sum_{r=0}^{k-1} \binom{k-1}{r} y^{(r)}$$

here $e^x = 1$ as we put $x = 0$ to obtain the Bell numbers We need to prove it is true for $n = k+1$, i.e.

$$y^{(k+1)} = e^x \sum_{r=0}^{k} \binom{k}{r} y^{(r)}$$

$$y^{(k+1)} = \frac{d}{dx}\left(e^x \sum_{r=0}^{k-1} \binom{k-1}{r} y^{(r)}\right)$$

$$= e^x \sum_{r=0}^{k-1} \binom{k-1}{r} y^{(r)} + e^x \sum_{r=0}^{k-1} \binom{k-1}{r} y^{(r+1)}$$

$$= e^x \left[\binom{k-1}{0} y^{(0)} + \sum_{r=1}^{k-1} \binom{k-1}{r} y^{(r)} + \binom{k-1}{k-1} y^{(k)} + \sum_{r=0}^{k-2} \binom{k-1}{r} y^{(r+1)}\right]$$

$$= e^x \left[y^{(0)} + \sum_{r=1}^{k-1} \binom{k-1}{r} y^{(r)} + \sum_{r=1}^{k-1} \binom{k-1}{r-1} y^{(r)} + y^{(k)}\right]$$

$$= e^x \left\{y^{(0)} + \sum_{r=1}^{k-1} \left[\binom{k-1}{r} + \binom{k-1}{r-1}\right] y^{(r)} + y^{(k)}\right\}$$

$$= e^x \left[y^{(0)} + \sum_{r=1}^{k-1} \binom{k}{r} y^{(r)} + y^{(k)} \right] *$$

$$= e^x \sum_{r=0}^{k} \binom{k}{r} y^{(r)}$$

We have shown that the result is true for $n = 1$ and that if it is true for $n = k$ then it is also true for $n = k + 1$. Therefore it is true, by induction, for all $n \geq 1$.

*Here the property used was:

$$\binom{k}{r} = \binom{k-1}{r-1} + \binom{k-1}{r}$$

Proof

$$\binom{k-1}{r-1} + \binom{k-1}{r}$$

$$= \frac{(k-1)!}{(r-1)!(k-r)!} + \frac{(k-1)!}{r!(k-r-1)!}$$

$$= \frac{r(k-1)!}{r!(k-r)!} + \frac{(k-1)!(k-r)}{r!(k-r)!}$$

$$= \frac{(k-r+r)(k-1)!}{r!(k-r)!}$$

$$= \frac{k(k-1)!}{r!(k-r)!}$$

$$= \frac{k!}{r!(k-r)!}$$

$$= \binom{k}{r}$$

5.4 Stirling Numbers of the Second Kind

A second way of finding the number partitions of a set are using the Stirling numbers of the second kind.
$\left\{ {n \atop k} \right\}$ = the number of ways of partitioning a set with n elements into k non-empty subsets.
Think of an element n in a set with n elements:
If n is on its own, the number of partitions $= \left\{ {n-1 \atop k-1} \right\}$.
If n is not on its own, it is part of one of the k subsets formed from $n - 1$ elements, the number of partitions $= k\left\{ {n-1 \atop k} \right\}$.
Therefore the Stirling numbers of a second kind have the property:

$$\left\{ {n \atop k} \right\} = k \left\{ {n-1 \atop k} \right\} + \left\{ {n-1 \atop k-1} \right\}$$

$$B_n = \sum_{k=0}^{n} \left\{ {n \atop k} \right\}$$

6 Partitions of a number

A partition of a number is a way of representing a non negative integer n as the sum of positive integers called parts, with these parts written in non increasing order. e.g. $n = a + b + c$ where $a \geq b \geq c$ The partition function $p(n)$ is the number of distinct ways of writing n as the sum of positive integers, where the order of the parts does not matter. To find $p(n)$, we can use the generating function of partitions of a number:

$$\prod_{k=1}^{\infty} \frac{1}{1-x^k}$$

Using the binomial expansion, this can be written as:

$$(1 + x + x^2 + x^3 \ldots)(1 + x^2 + x^4 + x^6 \ldots)(1 + x^3 + x^6 + x^9 \ldots) \ldots$$

This expansion is only valid when $-1 < x < 1$, but because we do not usually assign values to x in generating functions, we do not need to worry about the issue of convergence.
The coefficient of x^k is the number of partitions of the number k.

Partitions can be represented by Ferrers diagrams which show the parts of each partition. For example, the Ferrers diagrams below show the 22 partitions for when k=8.

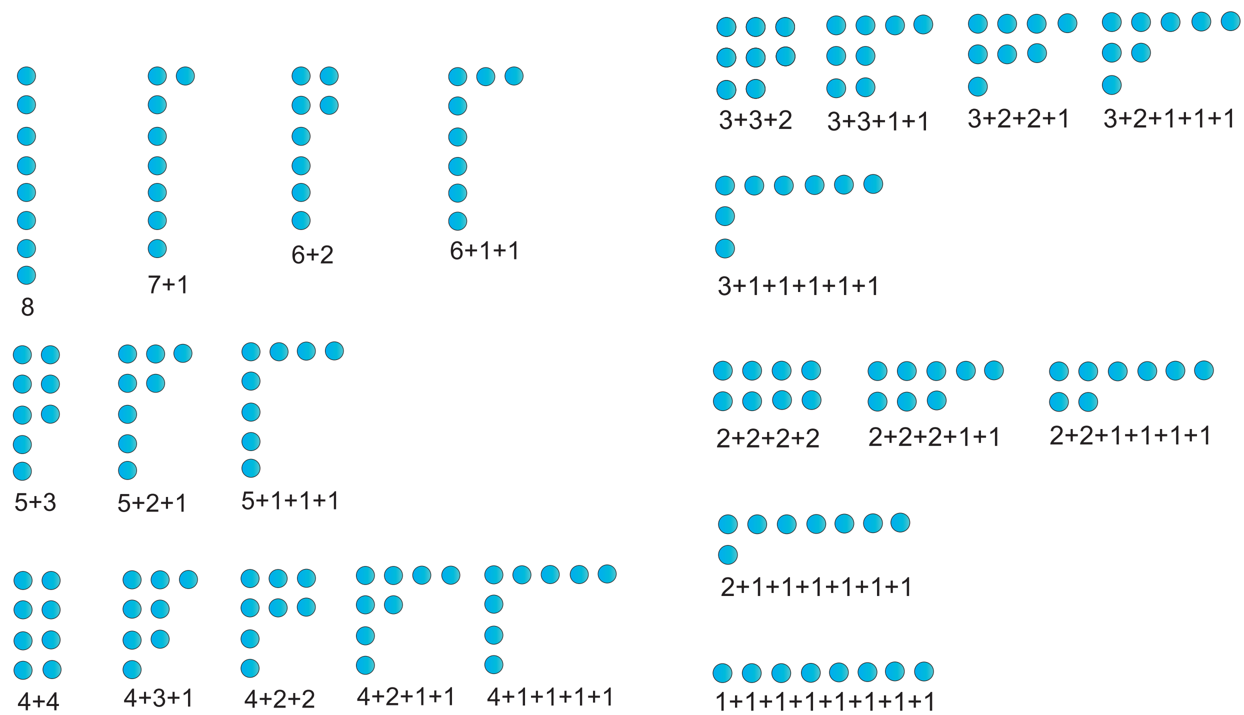


Figure 16: Partitions of 8

The above partitions can be categorised:

(i) The parts are all odd i.e. each part of the partition is an odd number:
7+1, 5+3, 5+1+1+1, 3+3+1+1, 3+1+1+1+1+1, 1+1+1+1+1+1+1+1
There are 6 such partitions.

(ii) The parts are all distinct:
8, 7+1, 6+2, 5+3, 5+2+1, 4+3+1
There are 6 such partitions.

(iii) The parts are all odd and distinct:
7+1, 5+3
There are 2 such partitions.

(iv) There are at most 2 parts:
8, 7+1, 6+2, 5+3, 4+4
There are 5 such partitions.

(v) There are at most 3 parts:
8, 7+1, 6+2, 6+1+1, 5+3, 5+2+1, 4+4, 4+3+1, 4+2+2, 3+3+2
There are 10 such partitions.

(vi) All parts are ≤ 2:
2+2+2+2, 2+2+2+1+1, 2+2+1+1+1+1, 2+1+1+1+1+1+1, 1+1+1+1+1+1+1+1
There are 5 such partitions.

(vii) All parts are ≤ 3:
3+3+2, 3+3+1+1, 3+2+2+1, 3+2+1+1+1, 3+1+1+1+1+1, 2+2+2+2, 2+2+2+1+1, 2+2+1+1+1+1, 2+1+1+1+1+1+1, 1+1+1+1+1+1+1+1
There are 10 such partitions.

We notice that:

$$\text{the number of partitions with } \leq 2 \text{ parts} = \text{the number of partitions with all parts } \leq 2$$

and

$$\text{the number of partitions with } \leq 3 \text{ parts} = \text{the number of partitions with all parts } \leq 3$$

These results can be proved by looking the the Ferrers diagrams. For any Ferrers diagram of a partition of k, the rows can be turned into columns and the columns into rows by reflecting the diagram along the diagonal from the top left of the diagram to the bottom right. This always gives us the Ferrers diagram of another partition of k. These 2 partitions are referred to as being 'conjugate'. Some partitions are symmetrical about this diagonal so a reflection just results in the original partition. These partitions are called 'self-conjugate'.
For partitions of 8, the conjugate pairs are:
$8 \leftrightarrow 1+1+1+1+1+1+1+1$
$7+1 \leftrightarrow 2+1+1+1+1+1+1$
$6+2 \leftrightarrow 2+2+1+1+1+1$
$6+1+1 \leftrightarrow 3+1+1+1+1+1$
$5+3 \leftrightarrow 2+2+2+1+1$
$5+2+1 \leftrightarrow 3+2+1+1+1$
$5+1+1+1 \leftrightarrow 4+1+1+1+1$
$4+4 \leftrightarrow 2+2+2+2$
$4+3+1 \leftrightarrow 3+2+2+1$
$4+2+2 \leftrightarrow 3+3+1+1$
The self-conjugate pairs are:
$4+2+1+1$ and $3+3+2$

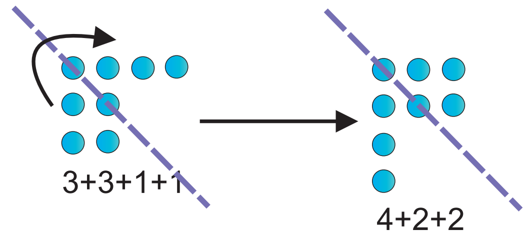


Figure 17: Conjugate partitions

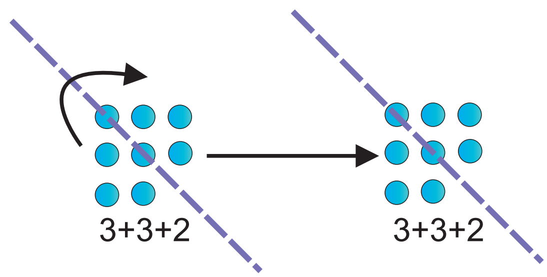


Figure 18: A self-conjugate partition

Because the columns in a Ferrers diagram represent the value of each part and the rows represent the number of parts, this means that when a partition with all parts $\leq l$ is reflected in the diagonal, its conjugate is a partition with $\leq l$ parts. As all partitions can be reflected in this way, this means that for any given n:

$$\text{the number of partitions with } \leq l \text{ parts} = \text{the number of partitions with all parts } \leq l$$

Sometimes Ferrers diagrams can be drawn the other way around, with the columns representing the number of parts and the rows representing the value of each part. In this case, the proof still applies.

For the partitions of 8, we can also see that:

$$\text{the number of partitions with all odd and distinct parts} = \text{the number of self-conjugate parts}$$

This can also be proved using Ferrers diagrams. For partitions with odd and distinct parts where each part has value m, the top $\frac{m+1}{2}$ dots of each part can be 'folded over' 90 degrees to form an upside down 'L' shape which is symmetrical about the top left to bottom right diagonal. As each part is also distinct and is therefore a different size, these folded over 'parts' fit together to form a different partition of n which due to being symmetrical about the diagonal, is self conjugate. This proves the observation above is true for any n.
Finally, for partitions of 8 we notice that:

$$\text{the number of partitions with all odd parts} = \text{the number of partitions with distinct parts}$$

This property can be proved using the generating functions of partitions of a number. The generating function for the number of partitions with all odd parts is:

$$(1 + x + x^2 + x^3 \ldots)(1 + x^3 + x^6 + x^9 \ldots)(1 + x^5 + x^{10} + x^{15} \ldots) \ldots = \prod_{k=1}^{\infty} \frac{1}{1 - x^{2k-1}}$$

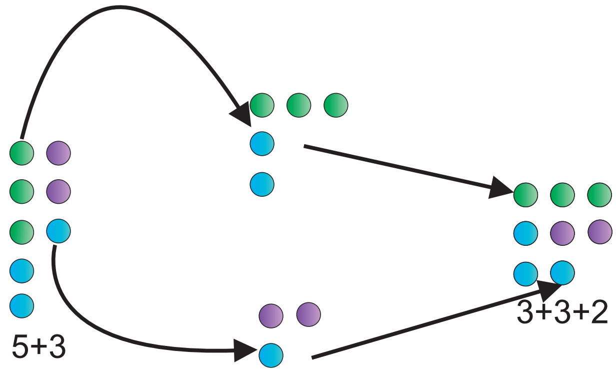


Figure 19: Transforming partitions with odd and distinct parts into self-conjugate partitions

The generating function for the number of partitions with distinct parts is:

$$(1+x)(1+x^2)(1+x^3)\ldots = \prod_{k=1}^{\infty}(1+x^k)$$

This can be rewritten using the difference of two squares as:

$$\frac{1-x^2}{1-x}\cdot\frac{1-x^4}{1-x^2}\cdot\frac{1-x^6}{1-x^3}\cdot\frac{1-x^8}{1-x^4}\ldots = \prod_{k=1}^{\infty}\frac{1-x^{2k}}{1-x^k}$$

However, this simplifies for any n as the terms in the numerators with even powers of x cancel with the every other term in the denominator, leaving just terms with an odd power of x in the denominator and terms with powers of x greater than n in the numerator (which we can ignore as they do not contribute to the coefficient of x^n and therefore do not affect the number of partitions with distinct parts of n). This is written as:

$$\prod_{k \text{ odd } k\geq 1}\frac{1}{1-x^k}$$

which is the same as the generating function for the number of partitions with all odd parts.

Generating functions can be applied to real life problems. For example, Bob is at a pick and mix counter choosing n sweets:

- there can be at most 1 gobstopper
- there can be at most 3 bonbons
- the number of cola bottles must be a multiple of 4
- the number of jelly beans must be even

In how many ways can Bob choose n sweets for his pick and mix?
The generating function for gobstoppers (the number of ways of choosing gobstoppers) is:

$$G(x) = 1 + x$$

The generating function for bonbons is:

$$B(x) = 1 + x + x^2 + x^3 = \frac{1-x^4}{1-x}$$

The generating function for cola bottles is:

$$C(x) = 1 + x^4 + x^8 + x^{12} \ldots = \frac{1}{1 - x^4}$$

The generating function for jelly beans is:

$$J(x) = 1 + x^2 + x^4 + x^6 \ldots = \frac{1}{1 - x^2}$$

The generating function for choosing sweets is:

$$\begin{aligned} G(x)B(x)C(x)J(x) &= (1 + x)\frac{1 - x^4}{1 - x} \cdot \frac{1}{1 - x^4} \cdot \frac{1}{1 - x^2} \\ &= \frac{1}{(1 - x)^2} \\ &= 1 + 2x + 3x^2 + 4x^3 \ldots nx^{n-1} + (n + 1)x^n \end{aligned}$$

This means the number of ways in which Bob can choose a selection of n sweets is $n + 1$.

References

[1] T. van Aardenne-Ehrenfest and N. G. de Bruijn, 'Circuits and trees in oriented linear graphs', *Simon Stevin* 28 (1951) 203–217

[2] N. G. de Bruijn, 'A Combinatorial Problem', *Indagationes Math.* 8 (1946) 461–467.

[3] F. Chung, P. Diaconis and R. Graham, 'Universal cycles for combinatorial structures', *Discrete Mathematics* 110 (1992) 43–59

[4] P. Diaconis and R. Graham, 'Products of Universal Cycles', *A Gathering for Gardner* (2005)

[5] J. R. Johnson, 'Universal cycles for permutations', *Discrete Mathematics* 309 (2009) 5264–5270

[6] C. Lessard, 'Cycles of n-digit binary integers', *American Mathematical Monthly* 58 (1951) 573–575

7 Acknowledgements

I would like to thank Professor P. J. Giblin for his help and guidance throughout the project and for making the project so interesting and enjoyable.

I would also like to thank Sharon Burke, the Nuffield Foundation and Maestro for giving me this amazing opportunity.

Finally, I would like to thank my teachers at school for allowing me to apply for a Nuffield Bursary.

I have learnt so much during the project and am extremely grateful to everyone who made it possible. Universal cycles was a new area of maths to me which was very interesting to study and I will definitely try to keep up to date with any future developments in this exciting field! The project has confirmed my choice to study mathematics at university and has given me the wonderful experience of working with Professor Giblin. I have learned to work more independently and have benefitted massively from the project.