Kami

Shevonne Chia (h1710124)

January 23, 2019

1 Introduction

This problem was inspired by Kami 2, a mobile game on iOS.

Consider a lattice of cells divided into R regions where each region is assigned 1 colour (not necessarily distinct), with a total of 0 <C ≤ R colours. We call a move altering the colour of one region.

Now two regions are merged if they are adjacent (share a common side) and have the same colour. Merged regions behave as a single region, i.e. their colour can be set simultaneously with
one move.

The objective is to find the minimum number of moves m required to make all regions the same colour. (Note the background is not a region and cannot be coloured.) As such we call an ideal move one which allows the number of moves m to be fulfilled.

Figure 1: Example of one valid move.

1.1 Observations

We can see that m ≥ C − 1 since each move can completely eliminate at least one colour, and the total number of colours must decrease from C to 1. For the equality to be reached, each move must completely remove one colour.

Additionally, note that it is always possible to set the colour of each vertex to c0. We let N(c)= number of regions with colour c. By pigeonhole principle, ∃ ci | N(ci) ≥⌈ R

C ⌉.

1

Kami

Shevonne Chia (h1710124)
January 23, 2019

Mathematical Medley • Volume 45 No. 1 • August 2019 7

SMPF Project: Kami NUS High School

This gives us:

m ≤ R −⌈R

C ⌉ = ⌊R − R

C ⌋ = ⌊R(1 − 1

C)⌋

As such, these inequalities form our upper and lower bounds.

C − 1 ≤ m ≤⌊R(1 − 1

C)⌋

These bounds apply for all graphs and can be used in initially narrowing down possible m.

1.2 Lattices

We are investigating planes in square lattices, such that each vertex (grid square) is adjacent to that above, left of, right of and below it.

We will be referring to individual squares as cells, and to groups of adjacent cells with the same colour as regions.

Figure 2: Diagram of cells and regions.

For a lattice of width w and height h, we index the cells from (0, 0) (bottom left corner) to (w − 1,h − 1) (top right corner). Now (x, y) is adjacent to (x − 1,y), (x +1,y), (x, y − 1) and (x, y + 1).

2 Binary Colouring

In this project, we will mainly be investigating cases where C = 2 i.e. ci ∈ {0, 1}. Here we define toggling a region as flipping its colour (from 0 to 1, or vice versa).

Given a w by h grid with R ≤ wh regions, we can apply the earlier bound to get 1 ≤ m ≤ wh/2. However, we can attain a tighter bound.

Shevonne Chia 2

8 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami

We consider the point (⌊w/2⌋, ⌊h/2⌋). When the colour of this point is toggled repeatedly, the region will expand at least one cell in each direction, hence it will take at most m ≤⌊w/2⌋+⌊h/2⌋
moves.

2.1 Checkerboards

We claim this upper bound is met when a checkerboard pattern is used; colour of (x, y)= x + y mod 2.

0 1 0 1 0
1 0 1* 0 1
0 1 0 1 0
1 0 1 0 1

0 1 0 1 0
1 0 0* 0 1
0 1 0 1 0
1 0 1 0 1

0 1 1 1 0
1 1 1* 1 1
0 1 1 1 0
1 0 1 0 1

0 0 0 0 0
0 0 0* 0 0
0 0 0 0 0
1 0 0 0 1
*This cell, (2, 2), is toggled repeatedly until there is only one region. This fulfils m = 4.

As such, we will be investigating the checkerboard and its variants. Now we must prove that this upper bound is achieved.

2.1.1 Simple Definitions

The distance between two regions is defined as the minimum number of moves required to merge them. In the context of a checkerboard, we note that this is the Manhattan distance between them, i.e. |x1 − x2| + |y1 − y2|.

Let d(a, b) denote the distance between the regions a and b.

The diameter of a board is equal to the maximum distance between any two regions. In a checkerboard this is equivalent to the distance between two opposite corner cells, which is simply w + h − 2.

Hence, the centroid is the region with minimum distance to the furthest region.

2.1.2 Centre Cells

We define a centre cell to be one where

x = ⌈(w − 1)/2⌉, ⌊(w − 1)/2⌋
y = ⌈(h − 1)/2⌉, ⌊(h − 1)/2⌋

As such, the number of centre cells depends on the parity of w and h.

We notice that the centre cells serve as centroids and have minimal distance from the furthest point.

If w, h are even, each of the four centre cells is in one quadrant.

Mathematical Medley • Volume 45 No. 1 • August 2019 9

Kami

0 1 0 1
1 0* 1* 0
0 1* 0* 1
1 0 1 0
w = h =4

If exactly one of w, h is even, we have two centre cells.

0 1 0 1 0
1 0 1* 0 1
0 1 0* 1 0
1 0 1 0 1
w =5,h =4

If both w, h are odd, there is only one centre cell.

0 1 0 1 0
1 0 1 0 1
0 1 0* 1 0
1 0 1 0 1
0 1 0 1 0
w = h =5

A notable property of centre cells is that repeatedly toggling any one of them yields this minimum m. We can see in the previous diagram how the region expands as a square in a flood-fill fashion. We define the size of the square to be the maximum distance between its centre and outermost cell, which is the number of moves required to obtain it.

2.1.3 Diagonals and Links

A diagonal with equation y = ±x+k is a set of points (xi,yi) which fulfil the equation. We note that there are two sets of diagonals; upward diagonals with equation y = x + k (up, right) and downward diagonals with y = −x + k (down, right). Note that there are w + h − 1 of each type. Each point (xi,yi) belongs to exactly two diagonals.

We note that a diagonal is characterised by two perpendicular edge directions (up, down, left, right).

Two adjacent diagonals are linked when a region containing cells in both diagonals is formed. A link refers to the formation of a common region between two unlinked diagonals.

As such the distance between two cells must be equal to the number of links created in one direc-
tion. This is because the distance is defined by the shortest path via moves in two perpendicular directions which can be characterised by one type of diagonal.

It is easy to see that the number of links in the other diagonal is equal to the difference between
the number of moves in each direction.

10 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami

Note that all diagonals must be linked for the board to be complete, where w + h − 2 links are formed. In particular, any two cells must be merged, hence all diagonals between them must be linked; the required number of links is the maximum distance between the cells.

Each move causes the formation of at most two new links. This will give m ≥⌈ w+h−2

2 ⌉ =
⌊ w+h−1

2 ⌋.

For this maximum to be achieved, the colour of the final board must be either (0, 0) or (w − 1,h − 1). Similarly, considering downward diagonals, we must have the final colour as that of (0,h − 1) or (w − 1, 0).

However, if both w and h are even, then we have

(0, 0) = (w − 1,h − 1) = 0
(0,h − 1) = (w − 1, 0) = 1

so another move is required. This can be given with the expression ⌊w/2⌋ + ⌊h/2⌋−⌊ w+h−1

2 ⌋,
which returns 1 if this is fulfilled. Otherwise, it returns 0.

Hence, our bound is as follows:

m ≥⌊w + h − 1

2 ⌋ +(⌊w/2⌋ + ⌊h/2⌋−⌊w + h − 1

2 ⌋)
= ⌊w/2⌋ + ⌊h/2⌋

and thus proven.

For optimal m to be attained, we observe that we must maximise the number of links (new diagonals reached) per move. As such, we refer to the links which are not made in a move as omitted links. In particular,

1. If w and h are both odd, (w − 1) + (h − 1) links are required in each direction i.e. there are no omitted links.

2. If exactly one of w and h is odd, (w − 1) + (h) − 1 links are required i.e. in each direction there may be one omitted link.

3. If w and h are both even, (w)+(h) − 2 links are required i.e. in each direction there may be two omitted links.

2.1.4 Starting Cells

There are certain cells which, if they are toggled in the first move, allow this minimum m to be obtained. We call these starting cells. As such, we aim to identify such cells.

Note that all centre cells (centroids) of checkerboards are starting cells as we can simply toggle
them repeatedly. There may be other starting cells depending on the parity of w, h, which we
may now investigate.

Here we may assume that all cell sequences involve two or more toggled cells.

Mathematical Medley • Volume 45 No. 1 • August 2019 11

Kami

2.2 Interactions

First, we observe the interactions between regions.

Toggling two cells on either side of another cell will result in two connected squares being formed, which will henceforth expand as a region.

1 0 1 0 1
0 1 0 1 0
1 1* 1 0 1
0 1 0 1 0
1 0 1 0 1

1 0 1 0 1
0 1 0 1 0
1 1 1 1* 1
0 1 0 1 0
1 0 1 0 1

1 0 1 0 1 0 1
0 1 0 1 0 1 0
1 0 0 0 0 0 1
0 0 0 0* 0 0 0
1 0 0 0 0 0 1
0 1 0 1 0 1 0
1 0 1 0 1 0 1
Interactions of two cells with maximum distance 2 from each other.

Here we have two intersecting squares, each with size 2. This has similar effects to toggling the centre cell twice.

0 1 0 1 0
1 0 1 0 1
0 1 1* 1 0
1 0 1 0 1
0 1 0 1 0

1 0 1 0 1
0 1 1 1 0
1 1 1* 1 1
0 1 1 1 0
1 0 1 0 1

1 0 1 0 1 0 1
0 1 0 0 0 1 0
1 0 0 0 0 0 1
0 0 0 0* 0 0 0
1 0 0 0 0 0 1
0 1 0 0 0 1 0
1 0 1 0 1 0 1
Expansion of a single cell.

We note that here a region is comprised of one or several overlapping squares. When a region is toggled (expanded), the size of all squares in the region increases by exactly 1.

In the above example, the first two moves cause each square (of initial size 0) to expand to size
1. The next move causes the size of both squares to increase to 2.

The following example now illustrates two cells with maximum distance 3. Before the regions merge, one square has been toggled twice and the other only once, hence the squares are of different sizes.

0 1 0 1 0 1
1 0 1 0 1 0
0 1 0 1 0 1
1 1* 1 0 0* 0
0 1 0 1 0 1
1 0 1 0 1 0
0 1 0 1 0 1

0 1 0 1 0 1
1 0 1 0 1 0
0 0 0 1 0 1
0 0* 0 0 0* 0
0 0 0 1 0 1
1 0 1 0 1 0
0 1 0 1 0 1
An interaction between two cells with maximum distance 3.

Now, the following example demonstrates how these interactions are similar even when the cells
are not in the same row or column.

12 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami

0 1 0 1 0 1
1 0 1 0 1 0
0 1 1* 1 0 1
1 0 1 1* 1 0
0 1 0 1 0 1
1 0 1 0 1 0

0 1 0 1 0 1
1 0 0 0 1 0
0 0 0* 0 0 1
1 0 0 0* 0 0
0 1 0 0 0 1
1 0 1 0 1 0
Interaction between two cells with maximum distance 2. In the first two moves, they form squares with size 1 each. The subsequent move causes both squares to reach size 2.

Immediately after the squares merge, we consider the two sets of diagonals. In at least one set, the squares share exactly one diagonal; we define the cross of the squares as one less than the number of diagonals shared in the other set (≥ 1).

We note that the number of links created at the time immediately after merging is equal to 2(total size of squares) - (cross of squares) in one direction and 2(total size of squares) in the other; the number of moves is equal to the total size of squares.

Hence number of links = 4(total moves) - (cross of squares).

Noting that each move creates a maximum of four links, number of omitted links = (cross of squares).

As such there are no omitted links if and only if the cells are in the same column or row, since exactly one diagonal will contain both squares when the squares are merged and their cross is 0.

In addition, we note that at the point of merging, the two squares share a contiguous segment of cells (of length l ≥ 1) along a diagonal. Since there is a shared diagonal (perpendicular to the segment) on each such cell and in between such cells, we must have

cross = l +(l − 1) − 1=2l − 2

As such the cross is guaranteed to be even.

2.3 Graph Properties and Definitions2, 3

Now we shall consider some prior definitions in a graph theory context.

When we model each region as a vertex with edges joining it to each adjacent vertex, we may treat the board as an unweighted graph. Each edge has an equal cost, referring to one moves required to link the two vertices. We may classify each edge as horizontal or vertical.

As such, the distance between two regions is defined as the minimum number of edges on a shortest path between them. Every vertex (region) on one such path must be merged.

The diameter of a board the maximum distance between any two vertices, and the centroid the vertex with lowest distance to the furthest vertex. Both of these follow formal definitions in graph theory.

Mathematical Medley • Volume 45 No. 1 • August 2019 13

Kami

2.4 Reduction

The following analysis is derived from the paper The Complexity of Flood-Filling Games.

4 We
may note that this proof is based purely on properties of unweighted graphs, and thus may be used in any binary general graph; it is no longer limited to merely checkerboards.

We consider two regions a and b. (In a checkerboard context, these regions are cells.)

The regions are toggled ma and mb times respectively until they are merged. Note that each move decreases the distance to all other regions by 1. Hence the distance between the regions is ma + mb. Without loss of generality ma ≤ mb.

Note each region c which is merged after ma +mb moves. There are two possible cases: d(a, c) ≤ ma or d(b, c) ≤ mb.

Now, we consider another region e midway between a and b such that d(e, a)= ma and d(e, b)= mb. This can be achieved by considering a shortest path between a and b and taking the region with distance ma from a. Since the graph is unweighted, the region must have a distance mb from b.

We toggle this region ma + mb times.

Again we note each region c from earlier. Note that this region c is only merged with e if d(e, c) ≤ ma + mb. Hence we aim to prove that this is the case.

By triangle inequality, we have

d(e, c) ≤ d(a, e)+ d(a, c) ≤ ma + ma ≤ ma + mb

Hence d(e, c) ≤ ma + mb and c is necessarily linked within ma + mb moves.

The paper also mentions that since a solution using two regions may be reduced to toggling a single region, this can be applied through induction to attain a solution with only one region for all boards. As such, finding a minimum m is simply a matter of finding the minimum number of moves when only toggling one region.

Thus we may conclude that for i> 1, a sequence using i regions may be reduced to a sequence of the same length using i − 1 regions.

2.4.1 Centroids

We claim that the toggled region e is the centroid of the board. This is true since the number of moves required is equal to its maximum distance from any vertex, which we must minimise; the centroid is defined as the vertex with the minimal maximum distance, and results in minimum m.

In particular, in a checkerboard, we consider the sequence of moves (xi,yi where 0 ≤ i ≤ m − 1).

We claim that the coordinates of the centroid are (

!m−1

i=0 xi
m ,

!m−1

i=0 yi
m ). This definition may be
derived through the proof previously detailed.

Note that the centroid must be a centre cell.

14 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami

2.4.2 Converse

Note that the converse of this theorem is not always true i.e. for some region e merged with some cell c in ma + mb moves, two regions a and b (d(a, e)= ma ≤ d(b, e)= mb, e is on a shortest path between a and b) may not necessarily be merged with c in ma + mb moves.

In particular, the converse is false when for some cell c,

d(a, c) >ma
d(b, c) >mb
d(e, c) ≤ ma + mb

3 Analysis on Parity of Dimensions

Recall that the number of centre cells as well as omitted links are dependent on the parity of the width and height. As such, we may divide the different checkerboards into subcases which we examine individually.

3.1 Odd Width and Height

Recall that m = ⌊w/2⌋ + ⌊h/2⌋ = w+h

2 − 1.

We note that there are no omitted links when the width and height are both odd. As such, any starting cells must be within the same column or row such that they share exactly one diagonal after merging. Without loss of generality we assume them to be in the same vertical column.

Since the moves are symmetrical, we may assume the starting cells to be in the centre column to ensure the left and right corners are merged simultaneously.

1 0 1* 0 1
0 1 0* 1 0
1 0 1* 0 1
0 1 0* 1 0
1 0 1* 0 1
0 1 0* 1 0
1 0 1* 0 1
Centre column is indicated.

Additionally, we note that each move decreases the distance from the centre cell to two of the
four corners. Specifically, if the toggled region extends to or above the centre cell, it decreases distance to the upper corners; if it extends to or below the centre cell, it decreases distance to
the lower corners. Hence we need only keep track of the region positions and ensure that the corners are merged simultaneously.

These two conditions guarantee that all four corners are merged within m moves; however,

Mathematical Medley • Volume 45 No. 1 • August 2019 15

Kami

the edge cells may not be merged. To verify this, we use central(a) to denote the minimum distance from cell a to the central column.

At any point, we consider two cells a and b where

1. a is the most recently merged cell in the top row, and

2. b is the most recently merged cell with minimum central(b).

In order for all edge cells to be merged when corner cells are merged, we must have

central(a) ≤ central(b)

. This is sufficient to confirm that the solution is valid.

Note that checkerboards with dimensions (w − 1) × h, w × (h − 1) or (w − 1) × (h − 1) have the same value of m. As such, a solution which applies on a w × h checkerboard would also be valid on these checkerboards (provided the cells are within bounds and both dimensions are nonzero).

3.2 Odd Width, Even Height

Note that this case mirrors that where the width is even and the height odd.

We first consider a solution sequence with exactly two regions a and b. Let ma and mb be the number of moves on each region respectively before they are merged.

Note that only one link may be omitted in each direction. We do not consider the sequences in which no link is omitted, as these are simply variations of sequences when w, h are odd.

As such, for one link to be formed, the cross of a, b is 1, which is odd. Hence there is no such solution.

Since any solution with more than two regions may be reduced to two regions, no such solution involving more than two regions may exist.

Therefore, the only valid solutions are those which satisfy the larger checkerboard with dimen-sions (w + 1) × h.

3.3 Even Width and Height

With even width and height, two links may be omitted. This significantly increases the range of possible move sequences, which may not be realistic to fully solve. The range of valid starting
cells is also significantly greater, allowing us to note several basic valid starting cells.

When the width and height are even, we claim that any cell is a valid initial starting cell. Recall
that two omitted links in each direction are permitted for ideal m.

When a centre cell is toggled m − 1 times, a square with size m − 1 is formed and every cell except one is merged as a single region.

16 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami

1 0 1 0
0 1 0 1
1 0* 1 0
0 1 0 1

1 0 1 0
0 0 0 1
0 0* 0 0
0 0 0 1

1 1 1 0
1 1 1 1
1 1* 1 1
1 1 1 1
Lattice after m − 1 = 2 moves. One link (of two) has been omitted so far. One more move is required to merge with the last cell, which a square of size m − 1 is not large enough to contain.

As such, one corner cell will remain unmerged depending on which centre cell was toggled.

Note that the starting cell will form two links in a certain quadrant of the board. Hence, the centre cell in the opposite quadrant may be toggled m − 1 times, satisfying ideal m regardless.

1 0 1* 0
0 1 0 1
1 0 1 0
0 1 0 1

1 0 0* 0
0 1 0 1
1 0* 1 0
0 1 0 1

1 0 0* 0
0 0 0 1
0 0* 0 0
0 0 0 1

1 1 1* 1
1 1 1 1
1 1* 1 1
1 1 1 1

This strategy is valid since the starting cell occupies exactly 3 diagonals, hence fulfilling a cross of 2.

4 Small Cases

Beyond binary lattices, some investigation can be done for small general grids.

It has been proven4 that n × n boards with C ≥ 3 and n × 3 boards with unbounded C are NP-hard. Hence extensive analysis is difficult to obtain. Nonetheless, we may experiment with some small cases to reach some minor results.

Specifically, we will be investigating some small values of w, h, C and whether there exist any starting lattices meeting each value m within the earlier obtained bounds. We will also make some general observations.

4.1 w =2,h =2

Note R = 4 so the only non-extreme case is C = 2. From the bounds, we have the following:

C C − 1 ⌊R(1 − 1

C)⌋ m
2 1 2 1, 2

Mathematical Medley • Volume 45 No. 1 • August 2019 17

Kami

4.1.1 C =2

For m = 1: 00

1 1

For m = 2: 01

1 0

4.2 w =3,h =2

Note R = 6.

C C − 1 ⌊R(1 − 1

C)⌋ m
2 1 3 1, 2, 3
3 2 4 2, 3, 4
4 3 4 3, 4

4.2.1 C =2

For m = 1: 000

1 1 1 (as long as all 1s are connected)

For m = 2: 011

1 0 0
For m = 3:
we must have at least 3 regions of each colour, and hence a total of 6 regions.

Then the only such pattern is: 010

1 0 1

which can be solved with m = 2 moves (by flipping the colour of (1, 1) twice).

4.2.2 C =3

For m = 2: 012

0 1 2

For m = 3: 012

1 2 0
For m = 4:
We must have

N(0) + N(1) ≥ 4
N(0) + N(2) ≥ 4
N(1) + N(2) ≥ 4

but N(0) + N(1) + N(2) = 6 and as such, all cells must be separate regions with N(0) = N(1) =

N(2) = 2. This means that only 012

1 2 0 and its reflection are valid, but they can be solved
with m = 3.

18 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami

4.2.3 C =4

For m = 3: 012

3 1 2

For m = 4: 012

1 0 3

4.3 Observations

1. If there is a grid with size w × h and C = i, m = j, then a grid with the same size w × h and some C ∈ {i−1,i}, m = j −1 exists. This grid is reached by executing one ideal move
on the original.

(a) The contrapositive is also true i.e. if there is no grid with size w×h and C ∈ {i, i+1},
m = j, there is no grid with the same size w × h and C = i + 1, m = j + 1.

2. Any grid with fixed size can be extended arbitarily (by repeating the last row and/or column), thereby obtaining a larger grid with the same C and m.

5 Summary of Findings

In this project so far, we have modelled the problem as a lattice and discovered tight upper/lower bounds for m depending on R, C.

We have conducted further research on binary grids, having found a tighter upper bound and proven that this is satisfied in checkerboards. Beyond analysis of a single expanding square and the interactions between multiple such squares, we have related these to the properties of the minimum number of moves m and used this to narrow down the possible first moves. We have also applied certain graph theory techniques in to attain further findings.

In addition to this, we have done some analysis on the parity of the width and height of the board, and their effects on the move sequences themselves.

Nonetheless, there is a significant amount of potential future extension.

5.1 Further Investigation and Extension

This problem is a broad one with many unexplored facets.

As such, besides completing investigation for grids with odd width and height, we aim to devise
an algorithm to find m and the sequence of ideal moves for binary grids.

We also hope to eventually expand our research to general grids and possibly general graphs.

Mathematical Medley • Volume 45 No. 1 • August 2019 19

Kami

SMPF Project: Kami NUS High School

5.2 Graphs2

While the project is primarily focused on the lattice model, a more general way of modelling the problem is via general graphs. We take one point within each region which we define as the centre. Now we draw a graph, with a vertex at each centre of a region.

For every two adjacent regions A and B, we can select a point P on the common line. Now we join an edge from the centre of A to that of B passing through P. (Note that this is always possible.)

Hence, we now have the plane expressed as a planar undirected graph where each vertex repre-sents a region, and every two vertices representing adjacent regions are connected by an edge.

The colour of each such vertex is hence an integer ci;0 ≤ ci <C, corresponding to the colour of the region. As such, we label each vertex a0,a1,..., aR−1 where c0 ≤ c1 ≤... ≤ cR−1.

Figure 3: Demonstration of one move on graph.

References

1 State of Play. Kami 2.

2 Xiong Bin and Zheng Zhongyi. Graph Theory: Volume 3 (Mathematical Olympiad Series). World Scientific, East China Normal University Press.

3 Steven Halim and Felix Halim. Competitive Programming 3: The New Lower Bound of Pro-gramming Contests.

4 David Arthur, Rapha¨el Clifford, Markus Jalsenius, Ashley Montanaro, and Benjamin Sach.
The complexity of flood filling games. CoRR, abs/1001.4420, 2010.

Acknowledgements

I would like to thank NUS High School for giving me the opportunity to do this project and my mentor, Mr Chia, for guiding me every step of the way.

Shevonne Chia 14

20 Mathematical Medley • Volume 45 No. 1 • August 2019

Kami