Elementary proof techniques for the maximum number of islands

János Barát*[,1]

Péter Hajnal[2]

Eszter K. Horváth[3]

Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary

Abstract

Islands are combinatorial objects that can be intuitively defined on a board consisting of a finite number of cells. It is a fundamental property that two islands are either containing or disjoint. The maximum number of rectangular, brick and triangular islands have been recently determined by Czédli; Pluhár; Horváth, Németh and Pluhár. Here, we give short proofs for some previous theorems, and also for new, analogous results on toroidal and some other boards.

Key words: Island, rooted binary tree, induction, grid
2000 MSC: 05D99, 05C05

*Current address: Department of Computer Science, University of Pannonia, Egyetem u. 10, 8200 Veszprém, Hungary

Email addresses: barat@dcs.vein.hu (János Barát), hajnal@server.math.u-szeged.hu (Péter Hajnal), horeszt@math.u-szeged.hu (Eszter K. Horváth)

[1]Research is supported by OTKA Grant PD 75837 and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

[2]Research is supported by OTKA Grant K 76099.

[3]This research was partially supported by OTKA Grant T 049433, and by the Provincial Secretariat for Science and Technological Development, Autonomous Province of Vojvodina, grant "Lattice methods and applications".

1. Introduction, preliminaries

We start with an intuitive notion. Let a rectangular $m \times n$ board be given. We associate a number (real or integer) to each cell of the board. We can think of this number as a height above see level. A rectangular part of the board is called a *rectangular island* if and only if there is a possible water level such that the rectangle is an island in the usual sense.

1	2	1	2
1	5	3	1
2	3	5	1
1	5	3	2
2	1	2	1

Figure 1: Rectangular landscape with heights

The notion of an island turned up recently in information theory. The characterization of the lexicographical length sequences of binary maximal instantaneous codes in Földes and Singhi [4] uses the notion of *full segments*, which are one-dimensional islands. Several generalizations led to interesting combinatorial problems. Czédli [2] discovered a connection between islands and weakly independent subsets of finite distributive lattices. He determined the maximum number of rectangular islands on a rectangular board. The method is based on weak bases of a finite distributive lattice [3]. Pluhár [7] gave upper and lower bounds in higher dimensions. The third author together with Németh and Pluhár [5] gave upper and lower bounds for the maximum number of triangular islands on a triangular board. Lengvárszky [6] determined the minimal size of a maximal system of islands. In the present paper, we list some related problems with exact formulae. In each case, we present the proof, which we believe to be the shortest.

In full generality, we denote the set of all cells of some board by $\mathcal{C}$. A *height function* is a mapping $h : \mathcal{C} \to \mathbb{R}$, $c \mapsto h(c)$. We have to specify a neighborhood relation on the cells. If not otherwise stated, two cells are *neighbors* if they share a point. Let R be a subset of cells. The neighbors of R can be defined naturally as the set of cells not in R but having a neighbor in R. A connected subset R of cells is called an *island*, if the minimum height in R is greater than the maximum height on the neighbors of R. That is, a water level makes it an island in the usual sense. We always fix a geometric shape, and consider the islands of this shape only. If h is a height function, then we denote the induced set of islands by $\mathcal{I}(h)$. Let us consider rectangular islands. We say that rectangles R and S are *far from each other,* if no cell of R is the neighbor of any cell of S. We denote by $P(\mathcal{C})$ the power set of $\mathcal{C}$, that is the set of all subsets of $\mathcal{C}$. The following statement in a different form was proved in [2].

Lemma 1. *Let $\mathcal{C}$ be the set of all cells of some board, and let $\mathcal{B}$ denote the entire board as an island. Let $\mathcal{I}$ be a set of islands. The following two conditions are equivalent:*

(i) *there exists a mapping $h : \mathcal{C} \to \mathbb{R}$, $c \mapsto h(c)$ such that $\mathcal{I} = \mathcal{I}(h)$.*

(ii) *$\mathcal{B} \in \mathcal{I}$, and for any $R_1 \neq R_2 \in \mathcal{I}$ either $R_1 \subset R_2$, or $R_2 \subset R_1$, or R_1 and R_2 are far from each other.*

A subset of $P(\mathcal{C})$ satisfying the equivalent conditions of Lemma 1 is called *a system of islands.* The set of maximal elements of $\mathcal{I} \setminus \{\mathcal{B}\}$ is denoted by $\max \mathcal{I}$.

2. Methods

We list three effective proof techniques for island problems. We give detailed demonstration of the latter two, the original method can be read in [2]. We recall the following result of [2]:

Theorem 2. *The maximum number of rectangular islands of an $m \times n$ rectangular board is $r(m, n) = [(m + 1)(n + 1)/2] - 1$.*

Let $\mathcal{C}$ be the set of unit squares of the $m \times n$ board. The proof in [2] exploits that the islands form a weakly independent set in the distributive lattice of $P(\mathcal{C})$. In a distributive lattice, maximal weakly independent subsets are called *weak bases.* By the main theorem of [3], any two weak bases have the same cardinality. The details are given in [2].

We use basic graph theory, see Bondy and Murty [1]. A graph without a cycle is called a *forest.* A connected forest is a *tree.* A forest with a distinguished vertex (root) in each component is called a *rooted forest.* If a vertex u is not a root, then u^+ denotes the vertex following u on the unique path from u to the root. It is called the *father of* u, while u is a *son* of u^+.

If v is on the path from u to a root, then v is an *ancestor* of u and u is a *descendant* of v. For any v the vertex v and its descendants span a *rooted subtree* T_v. Finally, *leaves* are vertices without sons. A rooted tree is *binary* if and only if any non-leaf vertex has precisely two sons.

Let a height function h be defined on the set of cells, fix an island shape, and consider $\mathcal{I}(h)$. In our discussion, the entire board $\mathcal{B}$ is always an island itself. Therefore, the Hasse diagram of $(\mathcal{I}(h), \subseteq)$ is a rooted tree with root $\mathcal{B}$, denoted $T_0(\mathcal{I}(h))$. The sons of $\mathcal{B}$ are the *maximal* islands, and they are disjoint. The leaves of $T_0(\mathcal{I}(h))$ are the *minimal* islands.

In what follows, we use $\mathcal{I}$ instead of $\mathcal{I}(h)$. For each vertex x with exactly one son y, we add another son $y'(x)$ to the graph $T_0(\mathcal{I})$. The new vertex $y'(x)$ is a *dummy* island. This way, we obtain the graph $T(\mathcal{I})$. The leaves of $T(\mathcal{I})$ are the minimal islands and the dummy islands, and each non-leaf vertex has at least two sons. See Figure 2 for an illustration.

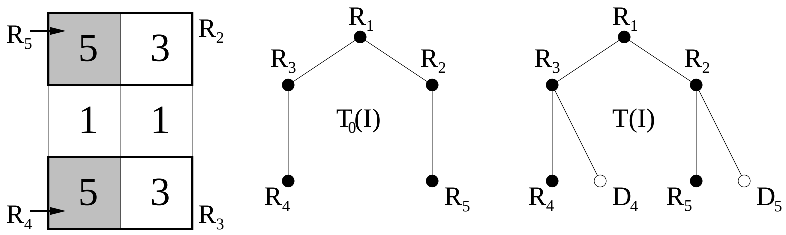

Figure 2: Hasse diagram of islands with respect to inclusion

In order to bound the number of islands, the following Lemma (folklore or an easy exercise in studying rooted trees) is very useful.

Lemma 3.

(i) Let T be a binary tree with ℓ leaves. Then the number of vertices of T depends only on ℓ and $|V| = 2\ell - 1$.

(ii) Let T be a rooted tree such that any non-leaf vertex has at least two sons. Let ℓ be the number of leaves in T. Then $|V| \leq 2\ell - 1$.

Our simple strategy is the following: if we know how to express the number of islands by the number of vertices and dummy vertices, then we apply Lemma 3.

A proof example. Let $B_{m,n}$ denote the set of mn unit squares of the $m \times n$ rectangular board. Let the island shape be rectangular. We call the vertices of the unit squares *grid points*, there are $(m+1)(n+1)$ of them.

Let $\mathcal{I}$ be a system of islands with s minimal islands and d dummy islands. Any island covers at least four grid points. If I_1 and I_2 are two islands, $I_1 \supset I_2$, and I_2 is the only son of I_1, then there is a dummy vertex joined to I_1. On the board, I_1 covers at least 2 more grid points than I_2. Therefore, we assign grid points to the leaves of $T(\mathcal{I})$: four points to the minimal islands, two points to the dummy leaves. These assigned sets of grid points are disjoint in the set of all $(m+1)(n+1)$ grid points. Therefore, $4s+2d \leq (m+1)(n+1)$. The number of leaves of $T(\mathcal{I})$ is $\ell = s+d$. By Lemma 3, the number of islands is $|V| - d \leq (2\ell - 1) - d = 2s + d - 1 \leq (m+1)(n+1)/2 - 1$. □

This proof is very suggestive, clear and short. Still, it needed some technical preparation. As it turns out, we can make the proof even more elementary.

The iterative description of $T_0(\mathcal{I})$ or $T(\mathcal{I})$ suggests a recursive proof technique: the mathematical induction. Actually, all known upper bounds on the number of islands [2, 5, 7] can be proved by induction.

A proof example. Let $r(m,n)$ be the maximum number of islands on $B_{m,n}$. We claim that $r(m,n) \leq (m+1)(n+1)/2 - 1$. Let us denote the covered grid points by $\|B_{m,n}\|$. For disjoint sub-boards $S_1, S_2, \ldots, S_k$ of $B_{m,n}$, we know that $\|B_{m,n}\| \geq \|S_1\| + \|S_2\| + \ldots + \|S_k\|$ holds.

We prove the claim by induction. The case of small boards can be easily checked. Let $\mathcal{I}^*$ be a system of islands realizing the number $r(m,n)$.
$r(m,n) = 1 + \sum_{R \in \max \mathcal{I}^*} r(R) \leq 1 + \sum_{R \in \max \mathcal{I}^*} (\|R\|/2 - 1) =$
$1 + 1/2 \sum_{R \in \max \mathcal{I}^*} \|R\| - |\max \mathcal{I}^*| \leq \|B_{m,n}\|/2 + 1 - |\max \mathcal{I}^*|$.

If $|\max \mathcal{I}^*| \geq 2$, then the induction is complete. If $|\max \mathcal{I}^*| = 1$, then one needs a minor technical remark to finish the proof. □

3. Applications

3.1. Peninsulas

A rectangular island P is a *peninsula* if it reaches at least one side of the board. We denote the maximum number of peninsulas in an $m \times n$ board by $p(m,n)$. The following is a slight strengthening of Theorem 2.

Theorem 4. *In a rectangular $m \times n$ board $p(m,n) = r(m,n)$.*

Proof. Let $p'(m,n)$ be the maximum number of peninsulas reaching the west or north side of the board. Clearly, $p'(m,n) \leq p(m,n) \leq r(m,n)$. We prove $p'(m,n) \geq r(m,n)$ by induction. For $m \in \{1,2\}$ and $n \in \{1,2\}$, we check that $p'(m,n) = r(m,n)$. For the induction step, let us delete two rows from south. This $2 \times n$ board contains n different peninsulas of width one reaching the west side of the original board. We can similarly delete two columns from east. In both cases, the entire $m \times n$ board is a new peninsula beyond the ones on the smaller boards. Therefore, $p'(m,n) \geq p'(m,n-2) + m + 1 = [(m(n-2) + m + (n-2) + 1)/2] - 1 + m + 1 = [(mn + m + n - 1)/2] - 1 = r(m,n)$. □

3.2. Cylindric board, rectangular islands

In this section, we put a square grid on the surface of a cylinder with height m and circumference of the base circle n. We get the same object by identifying the sides of length m of an $m \times n$ rectangle. We denote by $c_1(m,n)$ the maximum number of rectangular islands on this cylinder, assuming that the entire board is an island, but no other cylinders are.

Theorem 5. *If m, n are integers, $n \geq 2$, then $c_1(m,n) = [(m+1)n/2]$.*

Proof. By deleting a column of the cylinder, we get an $m \times (n-1)$ rectangle. Therefore, $c_1(m,n) \geq r(m,n-1) + 1 = [(m+1)n/2]$.

Let $\mathcal{I}^*$ be a maximum cardinality system of islands. For an island R, let $u(R)$ and $v(R)$ denote the length of its horizontal and vertical side. We drop R, and simply write u and v, when there is no danger of ambiguity. Now $c_1(m,n) = 1 + \sum_{R \in max\mathcal{I}^*} f(R) = 1 + \sum_{R \in max\mathcal{I}^*} ([(u+1)(v+1)/2] - 1) = 1 - |max(\mathcal{I}^*)| + \sum_{R \in max\mathcal{I}^*} [(u+1)(v+1)/2] \leq 1 - 1 + [(m+1)n/2] = [(m+1)n/2]$. We applied $|max(\mathcal{I}^*)| \geq 1$, and $\sum_{R \in max\mathcal{I}^*} [(u+1)(v+1)/2] \leq [(m+1)n/2]$. To see the latter, observe that we count the grid points of the entire board on the right-hand side and the grid points covered by the maximal islands on the left-hand side. □

3.3. Cylindric board, cylindric and rectangular islands

On a cylindric board, it is natural to consider cylindric islands as well. In this section, we allow two shapes for the islands, cylindric and rectangular. We denote by $c_2(m,n)$ the maximum cardinality of such a system of islands on the cylindric $m \times n$ board.

Theorem 6. *If $n \geq 2$, then $c_2(m,n) = [(m+1)n/2] + [(m-1)/2]$.*

Proof. We show by induction on m, that $c_2(m,n) \geq [(m+1)n/2] + [(m-1)/2]$. Notice, that $c_2(1,n) = n$ and $c_2(2,n) \geq r(2,n-1)+1 = [3n/2]$. Let $m > 2$. For the induction step, we remove a $2 \times n$ cylinder, which can contain n islands of height one and an empty row. Therefore, $c_2(m,n) \geq c_2(m-2,n)+n+1 = [(m-1)n/2] + [(m-3)/2] + n + 1 = [(m+1)n/2] + [(m-1)/2]$.

We show that $c_2(m,n) \leq [(m+1)n/2] + [(m-1)/2]$. Assume there is a maximum cardinality system given. By Theorem 5, there is a cylindric island in this system, Y say. Now Y is contained in a maximal cylindric island, M say, that is different from the entire board. Observe, that M is bordered with a strip of smaller heights from one side. By maximality, the rest of the board is a maximal cylindric island. Therefore, there exist $a, b \in \mathbb{N}_0$ so that $a+b+1 = m$ and $c_2(m,n) = c_2(a,n)+c_2(b,n)+1 = [(a+1)n/2]+[(a-1)/2]+[(b+1)n/2]+[(b-1)/2]+1 \leq [(a+b+1+1)n/2]+[(a+b+1-3)/2]+1 = [(m+1)n/2] + [(m-1)/2]$. □

3.4. Toroidal board, rectangular islands

With respect to the neighborhood relation, the most symmetric case appears on a toroidal board. Naturally, we get the most compact result of all.

Assume there is an $m \times n$ board on the torus, that is also known as $C_m \times C_n$. The island shape is fixed as rectangular, but we consider the entire board as an island. We denote by $t(m,n)$ the maximum number of rectangular islands on the $m \times n$ toroidal board.

Theorem 7. *If m and n are integers, $m, n \geq 2$, then $t(m,n) = [mn/2]$.*

Proof. We can cut off a horizontal and a vertical strip to get an $(m-1) \times (n-1)$ rectangular board. Therefore, $t(m,n) \geq r(m-1,n-1)+1 = [mn/2]$.

We denote by $\mathcal{I}^*$ a set of rectangular islands that realize the maximum cardinality. Analogously to the proof of Theorem 5, we obtain $t(m,n) = 1+\sum_{R \in max\mathcal{I}^*} r(R) = 1+\sum_{R \in max\mathcal{I}^*} ([(u+1)(v+1)/2]-1) = 1-|max(\mathcal{I}^*)|+\sum_{R \in \max\mathcal{I}^*} [(u+1)(v+1)/2] \leq 1-1+[mn/2] = [mn/2]$. □

4. Changing the neighborhood relation

In this section, we define two distinct cells as *neighbors* if and only if they have a side in common. For a height function h defined on a rectangular $m \times n$ board, the set of induced rectangular islands will be denoted by $\hat{\mathcal{I}}$. The maximum value of $|\hat{\mathcal{I}}|$ for all possible height functions h is denoted $\hat{r}(m,n)$. Notice, that R_1 and R_2 are far from each other now if their intersection consist of at most one point. As one may expect, Lemma 1 remains true. On the other hand, the following theorem is much less expectable.

Theorem 8. $\hat{r}(m,n) = [(m+1)(n+1)/2] - 1$, *the same as* $r(m,n)$.

Proof. Clearly, $\hat{r}(m,n) \geq r(m,n)$. We show $\hat{r}(m,n) \leq r(m,n)$ via induction on mn. We check the statement for $m \in \{1,2\}$ and $n \in \{1,2\}$. Let $\mathcal{I}^*$ be a maximum cardinality system of islands.

We mimic the proof of Theorem 5 and 7 with a sole difference: we associate all covered grid points to a rectangle except two. Namely, the northwest and northeast corner are left out from the counting. Assume that R is a rectangle with side lengths u and v. Let $\mu(R) = \mu(u,v) := (u+1)(v+1) - 2$. The induction step goes as before:
$\hat{r}(m,n) = 1 + \sum_{R \in max\mathcal{I}^*} \hat{r}(R) = 1 + \sum_{R \in max\mathcal{I}^*} ([(u+1)(v+1)/2] - 1) = 1 + \sum_{R \in max\mathcal{I}^*} ([\mu(u,v)/2]) \leq 1 + [\mu(m,n)/2] - 2$. In the last inequality, $\sum_{R \in \max\mathcal{I}^*} \mu(R) \leq \mu(m,n)$ would be trivial. The small improvement by 2 can be easily verified, calculating the number of northeast and northwest corners. □

5. Islands in hypercubes

We give an exact formula for the maximum number of hypercubic islands in a big hypercube. The board consists of all vertices of a hypercube, or in other words the elements of a Boolean algebra $\mathcal{A} = \{0,1\}^n$. Two cells are neighbors if their Hamming distance is 1. We denote the maximum number of islands in $\mathcal{A} = \{0,1\}^n$ by $b(n)$.

Theorem 9. $b(n) = 1 + 2^{n-1}$.

Proof. Consider the vertices with an odd number of 1's. They form a system of singleton islands. Therefore, $b(n) \geq 1 + 2^{n-1}$, if we consider the entire space as an island.

We prove the opposite direction by induction on n. For $n = 0,1$ the statement is easy to check. For $n \geq 2$, we cut the hypercube into two half-hypercubes of size 2^{n-1}. If one of them is an island, then the other part can not contain an island. If neither of them is an island, then by the induction hypothesis, in both half-hypercubes, the maximum cardinality of a system of islands is at most 2^{n-2}. This implies the claim: $b(n) \leq 1 + 2^{n-1}$. □

References

[1] A. Bondy and U.S.R. Murty, Graph Theory, *Graduate Texts in Mathematics*, **244.** Springer, New York, (2008), xii+651 pp.

[2] G. Czédli, The number of rectangular islands by means of distributive lattices, *European J. Combin.*, **30** (2009), 208–215.

[3] G. Czédli, A.P. Huhn and E.T. Schmidt, Weakly independent subsets in lattices, *Algebra Universalis*, **20** (1985), 194–196.

[4] S. Földes and N.M. Singhi, On instantaneous codes, *J. Combin. Inform. System Sci.*, **31** (2006), 307–316.

[5] E.K. Horváth, Z. Németh, G. Pluhár, The number of triangular islands on a triangular grid, *Period. Math. Hungar.*, **58** (2009), 25–34.

[6] Zs. Lengvárszky, The minimal cardinality of maximal systems of rectangular islands *European J. Combin.*, **30** (2009), 216–219.

[7] G. Pluhár, The number of brick islands by means of distributive lattices, *Acta Sci. Math.* **75**, No. 1-2, (2009), 3–11.