COMBINATORIAL SETTLEMENT PLANNING

MATE PULJIZ, STJEPAN ŠEBEK, AND JOSIP ŽUBRINIĆ

ABSTRACT. In this article, we consider a combinatorial settlement model on a rectangular grid where at least one side (east, south, or west) of each house must be exposed to sunlight without obstructions. We are interested in maximal configurations, where no additional houses can be added. For a fixed $m \times n$ grid, we explicitly calculate the lowest number of houses, and give close to optimal bounds on the highest number of houses that a maximal configuration can have. Additionally, we provide an integer programming formulation of the problem and solve it explicitly for small values of m and n.

1. INTRODUCTION

Consider the following problem: a rectangular $m \times n$ tract of land, whose sides are oriented north-south and east-west as in Figure 1, consists of mn square lots of size 1×1. Each 1×1 square lot can be either empty, or occupied by a single house. A house is said to be *blocked from sunlight* if the three lots immediately to its east, west and south are all occupied (it is assumed that sunlight always comes from the south[1]). Along the boundary of the rectangular $m \times n$ grid, there are no obstructions to sunlight. We refer to the models of such rectangular tracts of land, with certain lots occupied, as configurations. Of interest are maximal configurations, where no house is blocked from the sunlight, and any further addition of a house to the configuration on any empty lot would result in either that house being blocked from the sunlight, or it would cut off sunlight from some previously built house, or both.

We can encode any fixed configuration as a 0-1, $m \times n$ matrix C, with $C_{i,j} = 1$ if and only if a house is built on the lot (i, j) (i-th row and j-th column, counted from the top left corner). We can, equivalently, think of

---

[1]Our Southern Hemisphere friends are welcome to turn the page upside down when inspecting the figures in our paper.

FIGURE 1. An example of a tract of land ($m = 5$, $n = 7$).
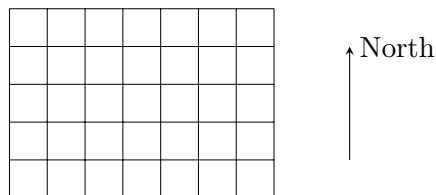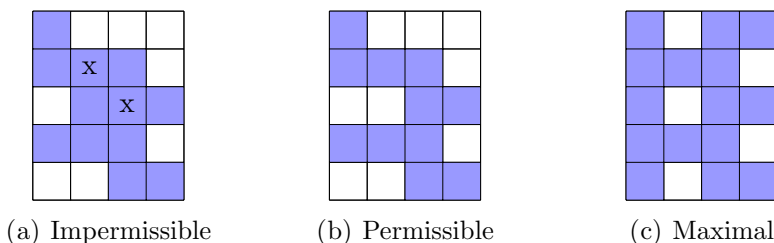

(a) Impermissible   (b) Permissible   (c) Maximal

FIGURE 2. Examples of impermissible, permissible and maximal configuration on a 5×4 tract of land.

C as a subset of $[m] \times [n] = \{(i, j) : 1 \le i \le m,\ 1 \le j \le n\}$, where, again, $(i, j) \in C$ if and only if a house is built on the lot (i, j).

It is natural to define the *building density* of a configuration C as

$$\frac{|C|}{mn},$$

where

$$|C| = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{i,j}$$

is the total number of occupied lots in the configuration C, i.e. the cardinality of C when C is interpreted as a subset of $[m] \times [n]$. We also call $|C|$ the *occupancy* of C.

A configuration C is said to be *permissible* if no house in it is blocked from the sunlight, otherwise it is called *impermissible*.

A configuration C is said to be *maximal* if it is permissible and no other permissible configuration strictly contains it, i.e. no further houses can be added to it, whilst ensuring that all the houses still get some sunlight. See Figure 2 for examples of impermissible, permissible and maximal configuration on a 5×4 tract of land. Shaded squares represent houses and unshaded squares represent empty lots on the tract of land. The houses that are blocked from the sunlight are marked with the letter x.

One is naturally interested in maximal configurations, especially those that achieve the highest and the lowest building density, or occupancy. One can think of those as solving one of two natural optimization problems

Problem 1. Maximize the revenue of a real estate investor, by building as many houses as possible, whilst ensuring that each house gets some sunlight.

Problem 2. Maximize the quality of living, by arranging the houses in order to achieve the lowest building density possible, whilst ensuring that no new houses can be added in the future without compromising access to sunlight.

Maximal configurations, for a particular $m, n \in \mathbb{N}$, with the highest building density possible we call *efficient*, while those with the lowest building density possible we call *inefficient*. We denote the number of occupied lots (occupancy) in any of the efficient configurations by $E_{m,n}$, and the number of occupied lots in any of the inefficient configurations by $I_{m,n}$.

One may also be interested in maximal configurations which exhibit occupancies in between those two extremes. We study these and related problems in an accompanying paper [8].

We were introduced to this problem by Juraj Božić who came up with it during his studies at the Faculty of Architecture, University of Zagreb. His main goal was to design a model for settlement planning where the impact of the architect would be as small as possible and people would have a lot of freedom in the process of building the settlement. This minimal intervention from the side of the architect is given through the condition that houses are not allowed to be blocked from the sunlight and that the tracts of land on which the settlements are built are of rectangular shapes.

Remark 1.1 (Boundary condition): As stated in the formulation of the problem, there are no obstructions to sunlight along the boundary of the rectangular $m \times n$ grid. However, it is possible to consider various other boundary conditions. One might be interested in the case where the whole border is bricked up and houses can get sunlight only from the empty lots within the grid. In fact, it is not hard to show that for each $m, n \geq 2$ there is an efficient configuration with all lots along the southern, eastern and western border occupied. This implies $E_{m,n} = \check{E}_{m-1,n-2} + 2m + n - 2$ where $\check{E}_{m-1,n-2}$ is the number of occupied lots in any efficient $(m-1) \times (n-2)$ configuration with the bricked up boundary. Note that the boundary condition along the northern side of the border is irrelevant.

The rest of the paper is organized as follows. In Section 2 we provide coarse bounds on the size of maximal configurations. In Section 3 we introduce periodic configurations on $\mathbb{Z}^2$ which, when restricted to a finite grid, come close to attaining the occupancy of efficient or inefficient configurations. In Section 4 we find a closed formula for $I_{m,n}$ and improve the bounds on $E_{m,n}$. In Section 5 we give an integer programming formulation of our optimization problems and compute $E_{m,n}$ explicitly, for values $m, n \leq 16$, using IBM ILOG CPLEX [2]. In Section 6, we provide alternative formulations of the problem of finding efficient configurations. Finally, in Section

7, we provide several possible directions for further research related to the model that we have studied.

2. TWO SIMPLE BOUNDS AND ASYMPTOTICS FOR LARGE m AND n

Lemma 2.1. *If C is any maximal configuration on the $m \times n$ grid, $m, n \geq 2$, then*

$$\frac{1}{2}mn \leq |C| \leq \frac{3}{4}mn + \frac{m-1}{2} + \frac{n}{4}.$$

Proof. We first prove the upper bound. If we interpret C as the set of all occupied lots, then its complement C^c is the set of all empty lots. Note that each occupied lot is either on the eastern, southern or western edge of the grid (there are $2m + n - 2$ such lots), or it gets sunlight from at least one empty lot in C^c, or possibly both. Furthermore, each empty lot in C^c gives light to at most 3 occupied lots in C. Therefore,

$$|C| \leq 3|C^c| + 2m + n - 2,$$

$$4|C| \leq 3(|C^c| + |C|) + 2m + n - 2 = 3mn + 2m + n - 2,$$

and the upper bound follows.

To prove the lower bound, it suffices to construct an injection from C^c to C, as then, the bound would follow from

$$mn = |C| + |C^c| \leq 2|C|.$$

To that end, consider any empty lot $(i, j) \in C^c$. Since C is maximal, the house cannot be built on the lot (i, j), i.e. (i, j) cannot be added to the configuration, without blocking some existing house (or itself) from the sunlight. This can happen in exactly two ways,

- either (i, j) is the only source of light to at least one of its occupied neighbors to the east, north or west: $(i, j+1)$, $(i-1, j)$, or $(i, j-1)$, (Note that we do not require that all of those three neighbors are occupied).
- or alternatively, its neighbors to the east, south and west: $(i, j+1)$, $(i+1, j)$, and $(i, j-1)$ are all occupied. (but (i, j) is not the only source of light to $(i, j+1)$, nor $(i-1, j)$, nor $(i, j-1)$, if occupied.)

In the first case, we map $(i, j) \in C^c$ to any of the neighbors: $(i, j+1)$, $(i-1, j)$, or $(i, j-1)$ that are occupied and for which (i, j) is the only source of light. In the second case, we map $(i, j) \in C^c$ to its east neighbor $(i, j+1)$. It is clear from the construction that this does define an injection from C^c to C, thus completing the proof of the lower bound. $\square$

Remark: Lemma 2.1 shows that for large grids, as both $m \to \infty$ and $n \to \infty$, the building density of maximal configurations must be between $\frac{1}{2}$ and $\frac{3}{4}$. We will later see that efficient configurations, in the limit, do approach building density $\frac{3}{4}$, while inefficient configurations, in the limit, do approach building density $\frac{1}{2}$. It is also interesting to note that these limit building densities are
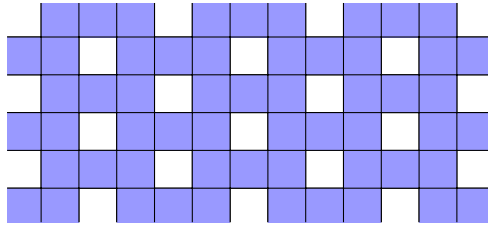
FIGURE 3. Brick pattern.

in fact, attained on the infinite grid by periodic configurations introduced in Section 3.

3. (NEAR-)OPTIMAL PATTERNS

In this section we describe periodic configurations on $\mathbb{Z}^2$ which yield optimal building densities, and the associated finite versions of these configurations which obey a similar periodic rule. The motivation for this is to construct somewhat regular configurations with near-optimal building densities. For this, we introduce the following definitions

A configuration on $\mathbb{Z}^2$ is simply a subset of $\mathbb{Z}^2$, or equivalently a function belonging to $\{0,1\}^{\mathbb{Z}^2}$. For a fixed configuration C on $\mathbb{Z}^2$, we define its building density as

$$\mathcal{D}(C) = \lim_{n \to \infty} \frac{1}{(2n+1)^2} \sum_{i=-n}^{n} \sum_{j=-n}^{n} C_{i,j},$$

when this limit exists.

A similar argument as in Lemma 2.1 yields the following bounds for the building density of maximal configurations on $\mathbb{Z}^2$

$$\frac{1}{2} \leq \mathcal{D}(C) \leq \frac{3}{4}, \quad \forall C \in \{0,1\}^{\mathbb{Z}^2}, \ C \text{ maximal.}$$

3.1. Patterns occurring in efficient configurations.

3.1.1. *Brick pattern.* Here we introduce the pattern for maximal configurations which we refer to as the brick pattern, see Figure 3. This pattern defines a maximal configuration C^{brick} on $\mathbb{Z}^2$ with the highest possible building density of $\mathcal{D}(C^{\mathrm{brick}}) = 3/4$.

Note that the configuration on $\mathbb{Z}^2$ with the brick pattern has the following two properties

- Each empty lot provides light to exactly three occupied neighbors.
- Each occupied lot receives light from exactly one neighbor.

The construction of finite maximal configurations with the brick pattern is depicted in Figure 4. The configurations are obtained by restricting the configuration C^{brick} to a finite grid, and building additional houses on the newly available lots. Among all such restrictions, we choose one that, in the
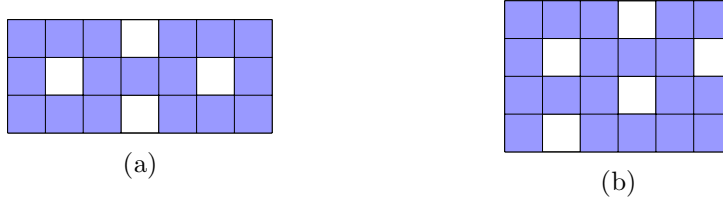
(a)

(b)

FIGURE 4. Finite configurations obtained from the brick pattern.

end, yields the highest occupancy. In order to determine the occupancy of such configurations, we start by fixing the dimensions $m \geq 2$ and $n \geq 3$. This configuration consists of $\left\lceil \frac{n}{2} \right\rceil$ fully filled vertical columns of height m. The remaining $\left\lfloor \frac{n}{2} \right\rfloor$ columns are filled in a shifted manner with only around half of the lots occupied. In fact, due to the necessity of shifting the half empty columns, we end up with around a half of them filled with $\left\lfloor \frac{m}{2} \right\rfloor$ houses and the remaining filled with $\left\lceil \frac{m}{2} \right\rceil$ houses. However, it is always more efficient to arrange the houses in such a way that there are

$$\left\lceil \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rceil$$

columns with $\left\lceil \frac{m}{2} \right\rceil$ houses and

$$\left\lfloor \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rfloor$$

columns with $\left\lfloor \frac{m}{2} \right\rfloor$ houses. We leave it to the interested reader to verify that in the cases $n \equiv 0 \pmod 4$, independently of m, and $n \equiv 2 \pmod 4$, $m \equiv 0 \pmod 2$, there is a possible minor improvement by choosing the restriction such that the bottom right lot can be filled, see Figure 4(b).

Definition 3.1 (Brick pattern function). *The function that gives us the occupancy $|C|$ of a configuration with the brick pattern is*

$$\mathcal{O}^{\text{brick}}(m, n) :=$$

$$\begin{cases} m \left\lceil \frac{n}{2} \right\rceil + \left\lceil \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rceil \left\lceil \frac{m}{2} \right\rceil + \left\lfloor \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rfloor \left\lfloor \frac{m}{2} \right\rfloor + 1, & \text{if } \begin{cases} n \equiv 0 \pmod 4; \text{ or} \\ n \equiv 2 \pmod 4; \text{ and} \\ m \equiv 0 \pmod 2. \end{cases} \\ m \left\lceil \frac{n}{2} \right\rceil + \left\lceil \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rceil \left\lceil \frac{m}{2} \right\rceil + \left\lfloor \frac{\left\lfloor \frac{n}{2} \right\rfloor}{2} \right\rfloor \left\lfloor \frac{m}{2} \right\rfloor, & \text{otherwise.} \end{cases}$$

Additionally, for $n = 2$, we have $\mathcal{O}^{\text{brick}}(m, 2) = 2m$.

3.1.2. *Comb pattern.* Another pattern occurring in efficient configurations is the comb pattern, see Figure 5. The configuration C^{comb} with such a pattern of built houses exhibits a building density of 2/3. Note that the configuration on $\mathbb{Z}^2$ with the comb pattern has the following two properties
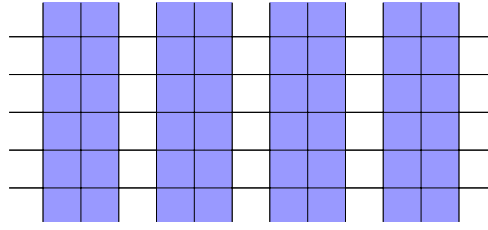
FIGURE 5. Comb pattern.

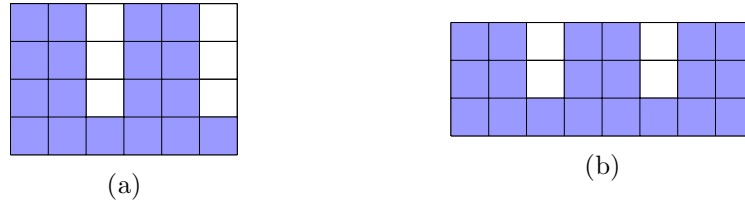(a)   (b)

FIGURE 6. Finite configurations obtained from the comb pattern.

- Each empty lot provides light to exactly two of its neighbors.
- Each occupied lot receives light from exactly one neighbor.

The way in which we obtain finite maximal configurations with the comb pattern is by restricting C^{comb} to a finite grid and filling all the southernmost lots, see Figure 6.

Definition 3.2 (Comb pattern function). *The function that gives us the occupancy $|C|$ of a configuration with the comb pattern is*

$$\mathcal{O}^{\text{comb}}(m,n) := \begin{cases} n + (m-1)\left(\frac{2}{3}n\right), & \text{if } n \equiv 0 \pmod 3; \\ n + (m-1)\left(\frac{2}{3}(n-1)+1\right), & \text{if } n \equiv 1 \pmod 3; \\ n + (m-1)\left(\frac{2}{3}(n-2)+2\right), & \text{if } n \equiv 2 \pmod 3, \end{cases}$$

where $m, n \geq 2$.

Remark: Based on its intermediate building density of $2/3$, it is expected that this pattern will not be occurring in either efficient nor inefficient configurations when both m and n are large. It does, however, occur in some efficient configurations when either m or n are small. This is confirmed in Table 2.

3.1.3. *Brick–comb combination.* We note here that the brick pattern and the comb pattern are compatible in the sense that one can design maximal configurations by alternating between these patterns. This is illustrated in Figure 7. For certain dimensions m and n, the combination of these patterns is more efficient than each of the patterns separately.

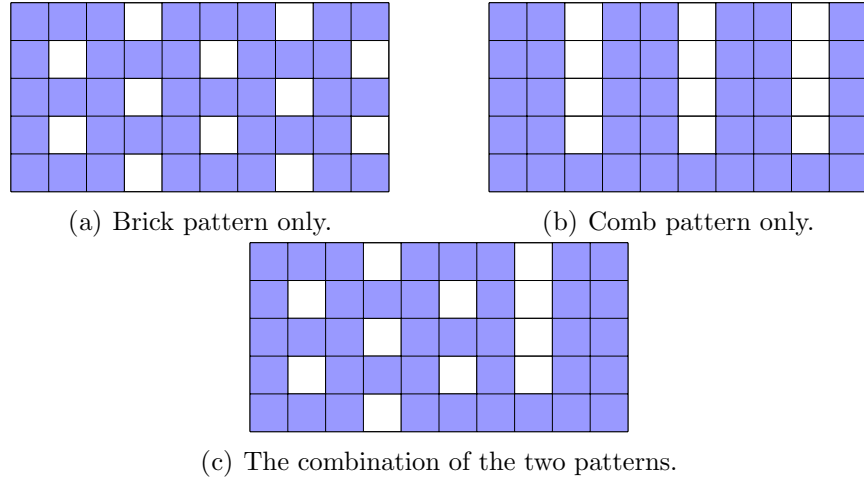### 3.2. **Patterns occurring in inefficient configurations.**

(a) Brick pattern only.   (b) Comb pattern only.

(c) The combination of the two patterns.

FIGURE 7. Consider the tract of land with dimensions 5×10. While the configuration with the brick pattern exhibits the occupancy of 38, as well as the configuration with the comb pattern, the combination, on the other hand, exhibits the occupancy of 39.
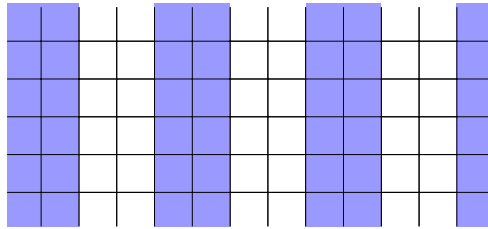

FIGURE 8. Rake pattern

3.2.1. *Rake pattern.* The rake pattern is a periodic pattern similar to the comb pattern but with bigger gaps. However, the configuration C^{rake} with the rake pattern, see Figure 8, yields the minimal possible building density of $1/2$. Note that the configuration on $\mathbb{Z}^2$ with the rake pattern has the following two properties

- Each empty lot provides light to only one of its neighbors.
- Each occupied lot receives light from only one of its neighbors.

The finite maximal configurations with the rake pattern are obtained by restricting C^{rake} to a finite grid, and building additional houses on the newly available lots. Among all such restrictions, we choose one that, in the end, yields the lowest occupancy, see Figure 9.

(a)   (b)

FIGURE 9. Finite configurations obtained from the rake pattern.

FIGURE 10. Stripe pattern.

Definition 3.4 (Rake pattern function). *The function that gives us the occupancy $|C|$ of a configuration with the rake pattern is*

$$\mathcal{O}^{\mathrm{rake}}(m,n) := \begin{cases} n + (m-1)\left(\frac{1}{2}n\right), & \text{if } n \equiv 0 \pmod 4; \\ n + (m-1)\left(\frac{1}{2}(n-1)+1\right), & \text{if } n \equiv 1 \pmod 4; \\ n + (m-1)\left(\frac{1}{2}(n-2)+2\right), & \text{if } n \equiv 2 \pmod 4; \\ n + (m-1)\left(\frac{1}{2}(n-3)+2\right), & \text{if } n \equiv 3 \pmod 4, \end{cases}$$

where $m, n \geq 2$.

3.2.2. *Stripe pattern.* The stripe pattern is a periodic pattern consisting of horizontal lines of built houses depicted in Figure 10. The configuration C^{stripe} on $\mathbb{Z}^2$ with the stripe pattern also yields the minimal possible building density of $1/2$, the same as the rake pattern. Note that the configuration on $\mathbb{Z}^2$ with the stripe pattern has the following two properties

- Each empty lot provides light to only one of its neighbors.
- Each occupied lot receives light from only one of its neighbors.

Remark: Notice that both of the above mentioned properties are shared by both the stripe and rake patterns. However, not all periodic patterns with the building density of $1/2$ obey these two conditions, see the check pattern in Section 3.2.4.

The finite maximal configurations with the stripe pattern are obtained by restricting C^{stripe} to a finite grid and building additional houses on the newly available lots. Among all such restrictions, we choose one that, in the end, yields the lowest occupancy, see Figure 11.
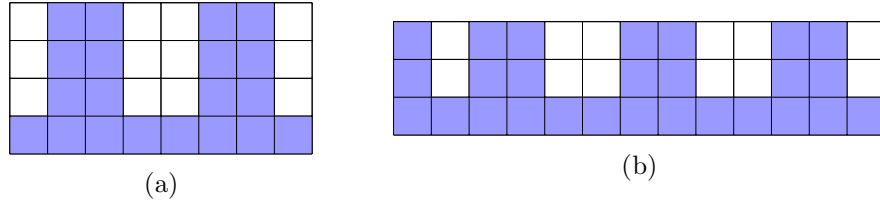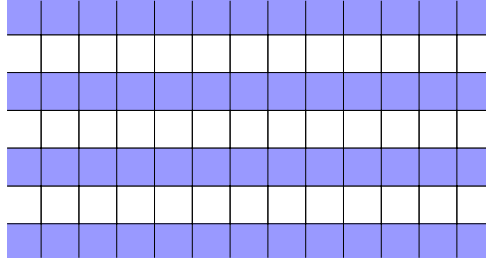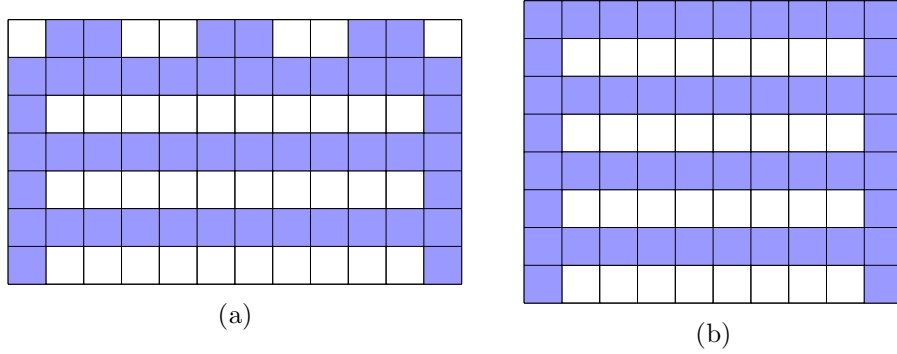
FIGURE 11. Finite configurations obtained from the stripe pattern.

Definition 3.6 (Stripe pattern function). *The function that gives us the occupancy $|C|$ of a configuration with the stripe pattern is*

$$\mathcal{O}^{\text{stripe}}(m, n) :=$$

$$\begin{cases} 2m + \left\lfloor \frac{m}{2} \right\rfloor (n-2), & \textit{if } m \equiv 0 \ (\text{mod } 2); \\ 2m + \left\lfloor \frac{m}{2} \right\rfloor (n-2) + \left(\frac{1}{2}n \right), & \textit{if } m \equiv 1 \ (\text{mod } 2), n \equiv 0 \ (\text{mod } 4); \\ 2m + \left\lfloor \frac{m}{2} \right\rfloor (n-2) + \left(\frac{1}{2}(n-1) + 1 \right), & \textit{if } m \equiv 1 \ (\text{mod } 2), n \equiv 1 \ (\text{mod } 4); \\ 2m + \left\lfloor \frac{m}{2} \right\rfloor (n-2) + \left(\frac{1}{2}(n-2) + 2 \right), & \textit{if } m \equiv 1 \ (\text{mod } 2), n \equiv 2 \ (\text{mod } 4); \\ 2m + \left\lfloor \frac{m}{2} \right\rfloor (n-2) + \left(\frac{1}{2}(n-3) + 2 \right), & \textit{if } m \equiv 1 \ (\text{mod } 2), n \equiv 3 \ (\text{mod } 4), \end{cases}$$

where $m, n \geq 2$.

Remark 3.7: Note that, in the case when m is odd, in order to yield the lowest occupancy, we need to fill the first row with a part of the rake pattern, see Figure 11(a).

3.2.3. *Rake–stripe combination.* Inspired by Remark 3.7 and forthcoming Lemma 4.1, we form a combination of the rake and stripe patterns, see Figure 12. This combination yields lower occupancy than each of the patterns separately. Moreover, in Theorem 4.5, we show that this combination attains the lowest occupancy possible among all maximal configurations.

Definition 3.8 (Rake–stripe pattern function). *The function that gives us the occupancy $|C|$ of a configuration with the rake–stripe pattern is*
(3.1)

$$\mathcal{O}^{\text{rake–stripe}}(m, n) := \begin{cases} n + 2 + (m-2) \left(\frac{1}{2}n \right), & \textit{if } n \equiv 0 \ (\text{mod } 4); \\ n + 2 + (m-2) \left(\frac{1}{2}(n-1) + 1 \right), & \textit{if } n \equiv 1 \ (\text{mod } 4); \\ n + 2 + (m-2) \left(\frac{1}{2}(n-2) + 2 \right), & \textit{if } n \equiv 2 \ (\text{mod } 4); \\ n + 2 + (m-2) \left(\frac{1}{2}(n-3) + 2 \right), & \textit{if } n \equiv 3 \ (\text{mod } 4), \end{cases}$$

when $m, n \geq 2$.

(a) Rake pattern only.   (b) Stripe pattern only.
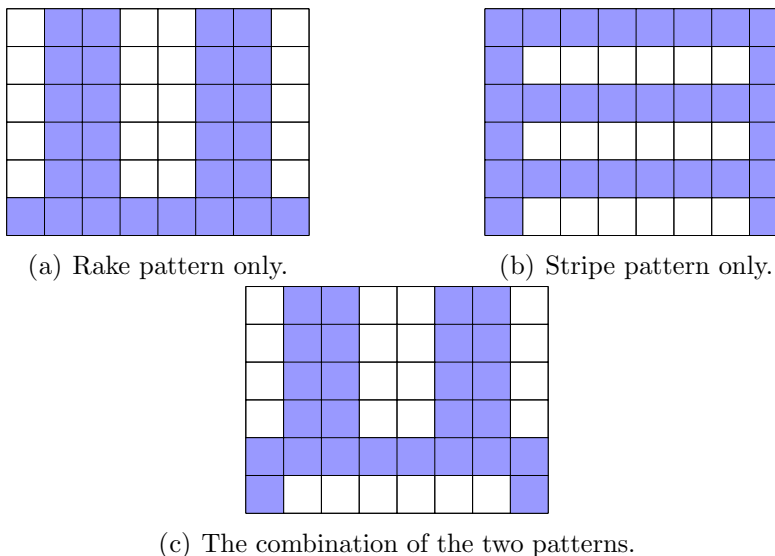

(c) The combination of the two patterns.

FIGURE 12. Consider the tract of land with dimensions 6×8. While the configuration with the rake pattern exhibits the occupancy of 28, and the stripe pattern exhibits the occupancy of 30, the combination, on the other hand, exhibits the occupancy of 26.
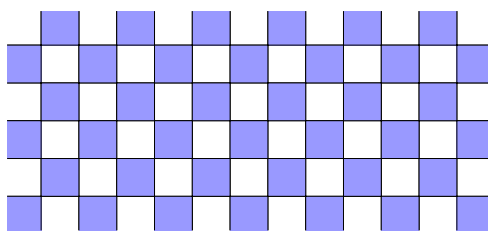
FIGURE 13. Check pattern.

3.2.4. *Check pattern.* Check pattern is a periodic pattern depicted in Figure 13. The maximal configuration C^{check} with the check pattern exhibits the building density of $1/2$. Note that the configuration on $\mathbb{Z}^2$ with the check pattern has the following two properties

- Each empty lot provides light to exactly three neighbors.
- Each occupied lot receives light from exactly three neighbors.

The finite maximal configurations with the check pattern are obtained by restricting C^{check} to a finite grid and building additional houses on the newly available lots. Among all such restrictions, we choose one that, in the end, yields the lowest occupancy, see Figure 14.

Figure 14. Finite configurations obtained from the check pattern.

Definition 3.9 (Check pattern function). *The function that gives us the occupancy $|C|$ of a configuration with the check pattern is*

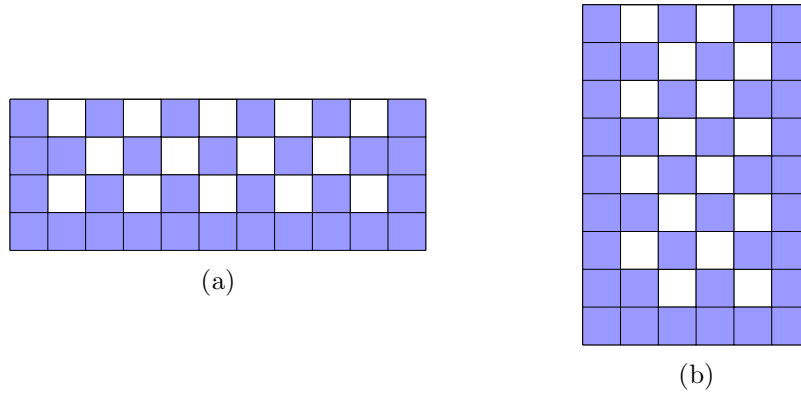$$\mathcal{O}^{\text{check}}(m,n) := 2(m-1) + n + \left\lfloor \frac{m-1}{2} \right\rfloor \left\lceil \frac{n-2}{2} \right\rceil + \left\lceil \frac{m-1}{2} \right\rceil \left\lfloor \frac{n-2}{2} \right\rfloor.$$

4. $I_{m,n}$ AND BOUNDS ON $E_{m,n}$

Recall that $I_{m,n}$ and $E_{m,n}$ are, respectively, the lowest and the highest occupancy attained among all of the maximal configurations on the $m \times n$ grid. In Lemma 2.1 we have proved

$$\frac{1}{2}mn \leq I_{m,n} \leq E_{m,n} \leq \frac{3}{4}mn + \frac{m-1}{2} + \frac{n}{4}.$$

Furthermore, the occupancy of any concrete maximal configuration C provides an upper bound on $I_{m,n}$ and a lower bound on $E_{m,n}$, since

$$I_{m,n} \leq |C| \leq E_{m,n}, \text{ for any maximal configuration } C.$$

In this section we obtain an explicit formula for $I_{m,n}$ (Theorem 4.5) and provide improved bounds on $E_{m,n}$.

4.1. **Explicit formula for $I_{m,n}$.** The main result of this section is Theorem 4.5. First we prove several auxiliary lemmas.

Lemma 4.1. *If C is any maximal configuration on the $m \times n$ grid, $m, n \geq 2$, then the number of occupied lots in the two southernmost rows is at least $n + 2$.*

Proof. We argue by contradiction. Suppose that there exist a maximal configuration C containing at most $n + 1$ occupied lots in rows $m - 1$ and m. Note that the lot $(m, 1)$ is occupied in any maximal configuration, as its only neighbors $(m, 2)$ and $(m - 1, 1)$ always get sunlight as they are on the

border. For the same reason, the lot (m, n) is also occupied in any maximal configuration.

We may also assume that the lot $(m - 1, 1)$ is occupied in C. If that was not the case, it would mean that $(m - 1, 2)$, $(m - 1, 3)$ and $(m, 2)$ are all occupied because of the maximality of C. We could then move the house from $(m, 2)$ to $(m - 1, 1)$, and we would still have a maximal configuration, with the same number of occupied lots as before. By similar reasoning, we may also assume that the lot $(m - 1, n)$ is occupied in C.

Therefore, of the remaining $2n - 4$ lots $\{m - 1, m\} \times \{2, 3, \ldots, n - 1\}$, at most $n - 3$ are occupied. By the pigeonhole principle, there must exist $2 \le i \le n - 1$ such that both lots $(i, m - 1)$ and (i, m) are empty. But this cannot happen in a maximal configuration, as there are no obstructions for (i, m) to be occupied — a contradiction. $\square$

Lemma 4.2. *Let C be any maximal configuration on the $m \times n$ grid, $m, n \ge 2$, and let $1 \le l \le m$ be a positive integer. The number of occupied lots in the restriction of the configuration C to the grid $S = \{1, 2, \ldots, l\} \times \{1, 2\}$ is at least one half of all the lots in S, i.e.*

$$|C \cap S| \ge \frac{1}{2}|S| = l.$$

The same is true if one takes S to be along the eastern border of the grid, instead of the western border, i.e. if $S = \{1, 2, \ldots, l\} \times \{n - 1, n\}$.

Proof. Let $1 \le i \le l$ and consider the lots $(i, 1)$ and $(i, 2)$. It is not possible that both of these are empty in a maximal configuration, since the only reason for leaving $(i, 1)$ empty is if $(i, 2)$ was occupied and, further, $(i, 1)$ was its only source of light. From here, the claim follows for the western border. Because of the mirror symmetry East↔West, the eastern border case also follows. $\square$

Lemma 4.3. *Let C be any maximal configuration on the $m \times n$ grid, $m \ge 2$, $n \ge 3$, and let $1 \le l \le m$ be a positive integer. The number of occupied lots in the restriction of the configuration C to the grid $S = \{1, 2, \ldots, l\} \times \{1, 2, 3\}$ is at least two thirds of all the lots in S, i.e.*

$$|C \cap S| \ge \frac{2}{3}|S| = 2l.$$

The same is true if one takes S to be along the eastern border of the grid, instead of the western border, i.e. if $S = \{1, 2, \ldots, l\} \times \{n - 2, n - 1, n\}$.

Proof. Let $1 \le i \le l$ and consider the lots $(i, 1)$, $(i, 2)$ and $(i, 3)$. We will show that either two out of three of these lots must be occupied or, failing that, there must be another row above them (i.e. $i > 1$), and taking into account the additional three lots in the row above $(i - 1, 1)$, $(i - 1, 2)$ and $(i - 1, 3)$, at least 4 out of the total of 6 lots must be occupied. From this, the statement of the lemma will follow.

First, we assume that $(i, 1)$ is empty. Arguing as in the proof of Lemma 4.2, since C is maximal, the lot $(i, 2)$ must be occupied, with $(i, 1)$ being its

only source of light. Therefore, $(i, 3)$ must also be occupied and we get two out of three occupied lots.

Otherwise, assume that $(i, 1)$ is occupied. If $(i, 2)$ or $(i, 3)$ is also occupied, we are done, so assume that both $(i, 2)$ and $(i, 3)$ are empty. Since C is maximal, the only reason for $(i, 2)$ being empty is

- either $(i, 2)$ is the only source of light to at least one of its occupied neighbors: $(i, 3)$, $(i - 1, 2)$, or $(i, 1)$;
- or alternatively, its neighbors to the east, south and west: $(i, 3)$, $(i + 1, 2)$, and $(i, 1)$ are all occupied.

Since $(i, 3)$ is assumed empty, and $(i, 1)$ has light coming from west border, the only possibility is that $(i - 1, 2)$ is occupied with $(i, 2)$ being its only source of light. This implies that $(i - 1, 1)$ and $(i - 1, 3)$ are also occupied. Hence, 4 out of 6 lots $\{i - 1, i\} \times \{1, 2, 3\}$ are occupied. This completes the proof of the western border version of the lemma. Again, the eastern border case follows from the mirror symmetry East$\leftrightarrow$West. □

Lemma 4.4. *Let C be any maximal configuration on the $m \times n$ grid, $m \geq 2$, $n \geq 4$, and let $1 \leq l \leq m$, $1 \leq t \leq n - 3$ be positive integers. The number of occupied lots in the restriction of the configuration C to the grid $S = \{1, 2, \ldots, l\} \times \{t, t + 1, t + 2, t + 3\}$ is at least one half of all the lots in S, i.e.*

$$|C \cap S| \geq \frac{1}{2}|S| = 2l.$$

Proof. Let $1 \leq i \leq l$ and consider the lots (i, t), $(i, t + 1)$, $(i, t + 2)$ and $(i, t + 3)$. We will show that either two out of four of these lots must be occupied or, failing that, there must be another row above them (i.e. $i > 1$), and taking into account the additional four lots in the row above $(i - 1, t)$, $(i - 1, t + 1)$, $(i - 1, t + 2)$ and $(i - 1, t + 3)$, at least four out of the total of eight lots must be occupied. From this, the statement of the lemma will follow.

We may assume that at most one out of four lots (i, t), $(i, t + 1)$, $(i, t + 2)$ and $(i, t + 3)$ is occupied, as otherwise we are done. Let us first assume that either (i, t) or $(i, t + 1)$ is the only occupied lot among those four. Since C is maximal, there are exactly four possible reasons for $(i, t + 2)$ being empty

- its neighbors to the east, south and west: $(i, t + 3)$, $(i + 1, t + 2)$, and $(i, t + 1)$ are all occupied;
- or $(i, t + 2)$ is the only source of light to its western neighbor $(i, t + 3)$;
- or $(i, t + 2)$ is the only source of light to its eastern neighbor $(i, t + 1)$;
- or $(i, t + 2)$ is the only source of light to its northern neighbor $(i - 1, t + 2)$.

In both cases (whether (i, t) or $(i, t + 1)$ is the only occupied lot among (i, t), $(i, t + 1)$, $(i, t + 2)$ and $(i, t + 3)$) the only possibility is that $(i - 1, t + 2)$ is occupied with $(i, t + 2)$ being its only source of light. This implies that $(i - 1, t + 1)$ and $(i - 1, t + 3)$ are also occupied. Hence, at least four out of eight lots $\{i - 1, i\} \times \{t, t + 1, t + 2, t + 3\}$ are occupied.

A similar argument shows that if $(i, t+2)$ or $(i, t+3)$ is the only occupied lot among (i, t), $(i, t+1)$, $(i, t+2)$ and $(i, t+3)$; then all of the lots $(i-1, t)$, $(i-1, t+1)$ and $(i-1, t+2)$ must also be occupied, and therefore again, at least four out of eight lots $\{i-1, i\} \times \{t, t+1, t+2, t+3\}$ are occupied.

Finally, assume that none of the lots (i, t), $(i, t+1)$, $(i, t+2)$ and $(i, t+3)$ are occupied. Arguing as in the previous cases, we conclude that all the lots: $(i-1, t)$, $(i-1, t+1)$, $(i-1, t+2)$ and $(i-1, t+3)$ must be occupied. Hence, once more, at least four out of eight lots $\{i-1, i\} \times \{t, t+1, t+2, t+3\}$ are occupied. $\qquad\square$

With this lemmas, we can improve the lower bound on the number of occupied lots in any maximal configuration and compute, exactly, the occupancies of inefficient configurations $I_{m,n}$.

Theorem 4.5 (Sharp lower bound on $I_{m,n}$). *If C is any maximal configuration on the $m \times n$ grid, $m, n \geq 2$, then*

$$|C| \geq I_{m,n} \geq \begin{cases} \frac{mn}{2} + 2, & \text{if } n \equiv 0 \pmod 4, \\ \frac{m(n+2)}{2}, & \text{if } n \equiv 2 \pmod 4, \\ \frac{m(n+1)}{2} + 1, & \text{if } n \equiv 1 \pmod 2. \end{cases}$$

Remark: These bounds are sharp since they are attained by configurations with the rake–stripe pattern combination introduced in Section 3.2.3. The expressions above match the occupancy of the rake–stripe configurations given in equation (3.1) and are, therefore, the values of $I_{m,n}$.

Proof of Theorem 4.5. If $n = 2$, all the lots must be occupied and we see that the bound holds. We now assume $n \geq 3$, and consider four cases depending on the remainder after dividing n by 4. In all four cases, we split the grid in the *lower part* consisting of the two southernmost rows $L = \{m-1, m\} \times [n]$ and the *upper part* consisting of the rest $U = [m-2] \times [n]$. The upper part we further divide depending on the said remainder. Note that, from Lemma 4.1, $|C \cap L| \geq n + 2$.

Case 1. $n \equiv 0 \pmod 4$ We divide U in equal blocks of width 4 and height $m - 2$ (Figure 15(a)). Applying Lemma 4.4 to each block, we get $|C \cap U| \geq \frac{1}{2}|U| = \frac{(m-2)n}{2}$. Thus, $|C| = |C \cap U| + |C \cap L| \geq \frac{mn}{2} + 2$.

Case 2. $n \equiv 1 \pmod 4$ We divide U in the western block of width 3, the eastern block of width 2, and the middle portion (if any) into equal blocks of width 4 (Figure 15(b)). Applying Lemma 4.3 to the western block, Lemma 4.2 to the eastern block, and Lemma 4.4 to the blocks in the middle portion, we get $|C \cap U| \geq \frac{2}{3}(m-2) \cdot 3 + \frac{1}{2}(m-2)(n-3) = \frac{(m-2)(n+1)}{2}$. Thus, $|C| = |C \cap U| + |C \cap L| \geq \frac{m(n+1)}{2} + 1$.

Case 3. $n \equiv 2 \pmod 4$ We divide U in the western block of width 3, the eastern block of width 3, and the middle portion (if any) into equal blocks of width 4 (Figure 15(c)). Applying Lemma 4.3 to the western and eastern block, and Lemma 4.4 to the blocks in the middle portion, we get
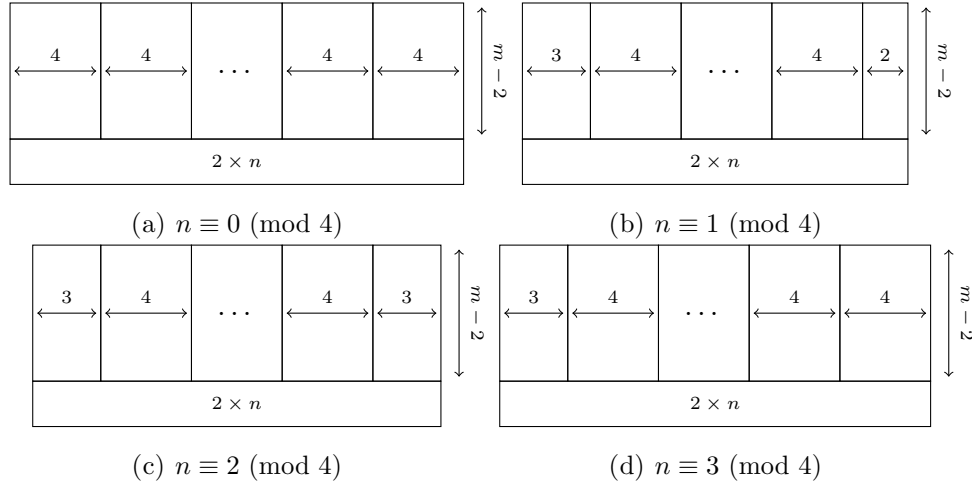
(a) $n \equiv 0 \pmod 4$ (b) $n \equiv 1 \pmod 4$

(c) $n \equiv 2 \pmod 4$ (d) $n \equiv 3 \pmod 4$

FIGURE 15. The subdivisions of the $m \times n$ grid used in the proof of Theorem 4.5.

$|C \cap U| \geq \frac{2}{3}(m-2) \cdot 6 + \frac{1}{2}(m-2)(n-6) = \frac{(m-2)(n+2)}{2}$. Thus, $|C| = |C \cap U| + |C \cap L| \geq \frac{m(n+2)}{2}$.

Case 4. $n \equiv 3 \pmod 4$ We divide U in the western block of width 3, and the rest (if any) into equal blocks of width 4 (Figure 15(d)). Applying Lemma 4.3 to the western block, and Lemma 4.4 to the blocks of width 4, we get $|C \cap U| \geq \frac{2}{3}(m-2) \cdot 3 + \frac{1}{2}(m-2)(n-3) = \frac{(m-2)(n+1)}{2}$. Thus, $|C| = |C \cap U| + |C \cap L| \geq \frac{m(n+1)}{2} + 1$.   □

4.2. **Improved bounds for $E_{m,n}$.** We now turn to the efficient configurations and improve bounds on $E_{m,n}$.

Proposition 4.7 (Upper bound on $E_{m,n}$). *If C is any maximal configuration on the $m \times n$ grid, $m, n \geq 2$, then*

$$|C| \leq E_{m,n} \leq \begin{cases} mn - \left\lfloor \frac{n}{4} \right\rfloor \cdot (m-1), & \text{if } n \not\equiv 3 \pmod 4, \\ mn - \left\lfloor \frac{n}{4} \right\rfloor \cdot (m-1) - \left\lfloor \frac{m}{2} \right\rfloor, & \text{if } n \equiv 3 \pmod 4. \end{cases}$$

Proof. The idea for this proof is taken from the solution of the problem posed in the 10$^{\text{th}}$ Middle European Mathematical Olympiad (see [6]). We begin the proof by dividing the tract of land, not including the bottom row, into $\left\lfloor \frac{n}{4} \right\rfloor \cdot (m-1)$ adjacent blocks of size 1×4, starting from the west. Now we claim that there exist an injection from the set of these blocks to the set of empty lots. Consider a single such block. Two situations are possible. Either this block itself contains an empty lot, in which case it can be mapped to its easternmost empty lot, or it is entirely occupied, which implies that the block directly below it now must have at least two empty lots. In that situation, we map the upper block to the westernmost empty lot of the lower

block, and the lower block to its easternmost empty lot. This mapping is clearly an injection. In the case $n \equiv 3 \pmod 4$, we can extend our injection by dividing the easternmost three columns into $\lfloor \frac{m}{2} \rfloor$ blocks of size 2×3. Since each of these contains at least one empty lot, we can injectively map those blocks into empty lots as well. This completes the proof.   □

Remark: Note that with the above proposition and the brick pattern function defined in Section 3.1.1, we have completely solved the case $m \times 3$, $m \geq 2$. We have

$$2m + \left\lceil \frac{m}{2} \right\rceil \leq E_{m,3} \leq 3m - \left\lfloor \frac{m}{2} \right\rfloor .$$

Before proceeding, we recall some well-known facts about the floor and ceiling functions which will be useful latter on

$$(4.1) \qquad\qquad \forall x \in \mathbb{R} \quad \lceil -x \rceil = - \lfloor x \rfloor ,$$

$$(4.2) \qquad\qquad \forall x \in \mathbb{R} \quad \lfloor -x \rfloor = - \lceil x \rceil ,$$

$$(4.3) \qquad\qquad \forall x \in \mathbb{R}, \forall k \in \mathbb{Z} \quad \lceil x + k \rceil = \lceil x \rceil + k,$$

$$(4.4) \qquad\qquad \forall x \in \mathbb{R}, \forall k \in \mathbb{Z} \quad \lfloor x + k \rfloor = \lfloor x \rfloor + k,$$

$$(4.5) \qquad\qquad \forall k, l \in \mathbb{N} \quad \left\lfloor \frac{k}{l} \right\rfloor = \left\lceil \frac{k - l + 1}{l} \right\rceil .$$

We will now give an improved upper bound on the size of maximal configurations. The bound is not going to be explicit, but given by the following recurrence relation. For a fixed $n \in \mathbb{N}$, we define the sequence $(R_{m,n})_{m \in \mathbb{N} \cup \{0\}}$ as follows

$$(4.6) \qquad \begin{cases} R_{0,n} = 0, \quad R_{1,n} = n, \\ R_{m,n} = R_{m-1,n} + n - \left\lfloor \frac{R_{m-1,n} - R_{m-2,n}}{3} \right\rfloor , & \text{if } m \geq 2. \end{cases}$$

Remark: From (4.6) it is not hard to see that

$$R_{m,n} = mn - \underbrace{\left\lfloor \frac{n}{3} \right\rfloor - \left\lfloor \frac{n - \left\lfloor \frac{n}{3} \right\rfloor}{3} \right\rfloor - \cdots - \left\lfloor \frac{n - \left\lfloor \frac{n - \lfloor \cdots \rfloor}{3} \right\rfloor}{3} \right\rfloor}_{m-1 \text{ terms}}$$

$$= mn - f_n(0) - f_n^2(0) - \cdots - f_n^{m-1}(0),$$

where $f_n(x) = \left\lfloor \frac{n - x}{3} \right\rfloor$, and f_n^k denotes the composition $\underbrace{f_n \circ f_n \circ \cdots \circ f_n}_{k \text{ times}}$.

Remark: Using (4.1) and (4.3) we can rewrite the recurrence relation (4.6) as

$$R_{m,n} = R_{m-1,n} + n - \left\lfloor \frac{R_{m-1,n} - R_{m-2,n}}{3} \right\rfloor$$

(4.7)
$$= n + \left\lceil R_{m-1,n} + \frac{-R_{m-1,n} + R_{m-2,n}}{3} \right\rceil$$

$$= n + \left\lceil \frac{2R_{m-1,n} + R_{m-2,n}}{3} \right\rceil, \text{ for } m \geq 2,$$

where we noted that $R_{m-1,n}$ is always an integer.

We will also need the following lemma.

Lemma 4.11. *Let $m \geq 2$, $n \geq 1$, and let C be any permissible configuration on the $m \times n$ grid. If $2 \leq r \leq m$ then*

$$|C \cap (\{r-1\} \times [n])| \leq n - \left\lfloor \frac{|C \cap (\{r\} \times [n])|}{3} \right\rfloor.$$

In other words, if there are k occupied lots in row r, then there are at most $n - \left\lfloor \frac{k}{3} \right\rfloor$ occupied lots in the row above it.

Proof. Let (r, j) be an arbitrary occupied lot in row $r \geq 2$ of a permissible configuration C, which is not on the eastern or western border ($1 < j < n$). It cannot happen that all three lots $(r-1, j-1)$, $(r-1, j)$, and $(r-1, j+1)$ in the row above are occupied, as this would mean that $(r-1, j)$ is blocked from the sun. We are, therefore, able to construct a map $f : C \cap (\{r\} \times \{2, \ldots, n-1\}) \to C^c \cap (\{r-1\} \times [n])$, which, to each occupied lot $(r, j) \in C \cap (\{r\} \times \{2, \ldots, n-1\})$, assigns an empty lot in $\{(r-1, j-1), (r-1, j), (r-1, j+1)\}$. By construction, each empty lot in row $r-1$ can be mapped to at most 3 occupied lots in row r, hence

$$|C \cap (\{r\} \times \{2, \ldots, n-1\})| \leq 3|C^c \cap (\{r-1\} \times [n])|.$$

Since $|C^c \cap (\{r-1\} \times [n])| = n - |C \cap (\{r-1\} \times [n])|$, and $|C \cap (\{r\} \times [n])| - 2 \leq |C \cap (\{r\} \times \{2, \ldots, n-1\})|$, we get

$$\frac{|C \cap (\{r\} \times [n])| - 2}{3} \leq n - |C \cap (\{r-1\} \times [n])|.$$

Therefore,

$$|C \cap (\{r-1\} \times [n])| \leq n - \frac{|C \cap (\{r\} \times [n])| - 2}{3},$$

and since the number on the left hand side is an integer

$$|C \cap (\{r-1\} \times [n])| \leq \left\lfloor n - \frac{|C \cap (\{r\} \times [n])| - 2}{3} \right\rfloor.$$

It remains to observe that

$$\left\lfloor n - \frac{|C \cap (\{r\} \times [n])| - 2}{3} \right\rfloor = n - \left\lceil \frac{|C \cap (\{r\} \times [n])| - 2}{3} \right\rceil$$

$$= n - \left\lfloor \frac{|C \cap (\{r\} \times [n])|}{3} \right\rfloor,$$

where we used (4.4), (4.2), and (4.5) for $l = 3$. □

Theorem 4.12 (Improved upper bound on $E_{m,n}$). *If C is any maximal configuration on the $m \times n$ grid, $m, n \geq 1$, then*

$$|C| \leq E_{m,n} \leq R_{m,n},$$

where $R_{m,n}$ is defined by the recurrence relation (4.6) (or (4.7)).

Remark: This bound seems to be sharp when $n = 1, 2, 3, 4, 5$ or 8, regardless of m (see Table 2), but it is not sharp in general (see Remark 4.14 below).

Proof of Theorem 4.12. Let the number of columns $n \in \mathbb{N}$ be fixed. We will argue by induction on the number of rows m.

The statement of the theorem is trivially true for $m = 0$ and $m = 1$ (and any $n \in \mathbb{N}$). The claim for $m = 0$ seems artificial, but it allows us for a simpler proof. To complete the inductive step, let us assume that $m \geq 2$, that the theorem holds for $m - 1$ and $m - 2$, and that C is a maximal configuration on $m \times n$ grid.

We set $S = \{2, \dots, m\} \times [n]$ and $T = \{3, \dots, m\} \times [n]$ (if $m = 2$ then $T = \emptyset$). Note that the number of occupied lots in C in row 2 is $|C \cap S| - |C \cap T|$. By Lemma 4.11, the number of occupied lots in row 1 is at most $n - \left\lfloor \frac{|C \cap S| - |C \cap T|}{3} \right\rfloor$. Hence,

$$|C| \leq n - \left\lfloor \frac{|C \cap S| - |C \cap T|}{3} \right\rfloor + |C \cap S|$$

$$= n + \left\lceil \frac{|C \cap T| - |C \cap S|}{3} + |C \cap S| \right\rceil$$

$$= n + \left\lceil \frac{|C \cap T| + 2|C \cap S|}{3} \right\rceil.$$

Above we used (4.1) and (4.3). Note that by assumption $|C \cap S| \leq R_{m-1,n}$ and $|C \cap T| \leq R_{m-2,n}$, as $|C \cap S|$ and $|C \cap T|$ are, perhaps not maximal but certainly, permissible configurations on grids of dimension $(m - 1) \times n$ and $(m - 2) \times n$, respectively. Therefore,

$$|C| \leq n + \left\lceil \frac{R_{m-2,n} + 2R_{m-1,n}}{3} \right\rceil = R_{m,n}.$$

This completes the inductive step and the proof of the theorem. □

TABLE 1. $R_{m,7}$ and $E_{m,7}$. Changing $R_{3,7}$ to 17 and $R_{4,7}$ to 22 and updating all the remaining values $R_{m,n}$, $m \geq 5$, according to (4.7), gives the correct values $E_{m,7}$, for $5 \leq m \leq 16$, which we were able to check using computer assisted exhaustive search.

m	$R_{m,7}$		m	$E_{m,7}$
2	12		2	12
3	**18**		3	17
4	**23**		4	22
5	29		5	28
6	34		6	33
7	40		7	39
8	45		8	44
9	51		9	50
10	56		10	55
$\vdots$	$\vdots$		$\vdots$	$\vdots$

Remark 4.14: The argument used in the proof of Theorem 4.12 actually shows

$$E_{m,n} \leq n + \left\lceil \frac{E_{m-2,n} + 2E_{m-1,n}}{3} \right\rceil, \text{ for } m \geq 2.$$

Since the right-hand side of this expression is increasing in both $E_{m-2,n}$ and $E_{m-1,n}$, it is possible to obtain even better bounds on $E_{m,n}$ than those achieved by $R_{m,n}$ simply by computing explicitly (using e.g. exhaustive search) few strategically chosen values of $E_{m,n}$, and then letting the recurrence relation take over. The bound is, in a way, self-improving. As an example, in Table 1 are listed values of $R_{m,n}$ and $E_{m,n}$, for $2 \leq m \leq 10$ and $n = 7$. The values $R_{m,n}$ are clearly not matching $E_{m,n}$, but if we change $R_{3,7}$ into 17 and $R_{4,7}$ into 22, and if we update all the remaining values $R_{m,n}$, $m \geq 5$, according to (4.7), we obtain the correct values $E_{m,7}$, for $5 \leq m \leq 16$. We were able to check these using computer assisted exhaustive search.

Remark: It is straightforward to check, by induction, that the bound on $E_{m,n}$, given by the recurrence relation $R_{m,n}$, in Theorem 4.12 is better than the one previously obtained in Proposition 4.7.

5. IP FORMULATION AND EXPLICIT SOLUTIONS

In this section we describe integer programming (IP) formulations for the problems of finding efficient and inefficient maximal configurations. By solving these problems we gain insight in the shape of the explicit solutions and the associated occupancies. The decision variables consist of a $m \times n$ matrix $X \in \{0,1\}^{m \times n}$ representing the configuration C on the tract of land with dimensions $m \times n$.

For the sake of elegance of the formulations, we introduce the following notation

Definition 5.1. *Let C be a configuration of houses on a tract of land with dimensions $m \times n$. For a lot (i, j) we define the following propositions (if applicable)*

- $\mathbb{P}_{i,j}^E := $ "(i, j) is empty and it is the only source of light to its eastern immediate neighbor." $\iff C_{i,j+1} + C_{i,j+2} + C_{i+1,j+1} = 3$.

- $\mathbb{P}_{i,j}^W := $ "(i, j) is empty and it is the only source of light to its western immediate neighbor." $\iff C_{i,j-1} + C_{i,j-2} + C_{i+1,j-1} = 3$.

- $\mathbb{P}_{i,j}^N := $ "(i, j) is empty and it is the only source of light to its northern immediate neighbor." $\iff C_{i-1,j+1} + C_{i-1,j-1} + C_{i-1,j} = 3$.

- $\mathbb{P}_{i,j}^C := $ "(i, j) is itself blocked from the sun." $\iff C_{i,j+1} + C_{i,j-1} + C_{i+1,j} = 3$.

Remark: Consider a maximal configuration C. We stress here that there are two distinct reasons why there would occur $C_{i,j} = 0$,

- either (i, j) is itself blocked from the sun;
- or (i, j) is the only source of light to one of its immediate neighbors.

This can be summarized with the following characterization of maximal configurations pertaining to the algebraic formulas.

Lemma 5.3. *Let C be a maximal configuration. For every $1 < i < m$, $2 < j < n - 1$ we have*

$$C_{i,j} = 0 \Rightarrow \mathbb{P}_{i,j}^E \vee \mathbb{P}_{i,j}^W \vee \mathbb{P}_{i,j}^N \vee \mathbb{P}_{i,j}^C,$$

with the similar formulas for the indices $i = 1, m$, $j = 1, 2, n - 1, n$ while keeping in mind the existence of its neighbors.

Efficient configurations (IP formulation). Let $m, n \in \mathbb{N}$ be the dimensions of the tract of land. The IP formulation of the problem of finding efficient maximal configurations is the following

(5.1)
$$
\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} X_{i,j} \\
\text{subject to} \quad & \neg \mathbb{P}_{i,j}^C, \quad i = 1, \dots m - 1, \quad j = 2, \dots, n - 1 \\
& X_{i,j} \in \{0, 1\}, \quad i = 1, \dots m, \quad j = 1, \dots, n
\end{aligned}
$$

Inefficient configurations (IP formulation). Let $m, n \in \mathbb{N}$ be the dimensions of the tract of land. The IP formulation of the problem of finding

TABLE 2. Table of occupancies of efficient configurations as calculated by IBM ILOG CPLEX.

m/n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	4	5	7	9	10	12	14	15	17	19	20	22	24	25	27
3	6	8	10	13	15	17	20	22	24	27	29	31	34	36	38
4	8	10	13	17	19	22	26	28	31	35	37	40	44	46	49
5	10	13	16	21	24	28	32	35	39	43	47	50	54	58	62
6	12	15	19	25	28	33	38	42	46	51	55	60	64	69	73
7	14	18	22	29	33	39	44	49	54	60	65	70	75	81	86
8	16	20	25	33	37	44	50	56	61	68	73	80	85	92	97
9	18	23	28	37	42	50	56	63	69	77	83	90	96	104	110
10	20	25	31	41	46	55	62	70	76	85	91	100	106	115	121
11	22	28	34	45	51	61	68	77	84	94	101	110	117	127	134
12	24	30	37	49	55	66	74	84	91	102	109	120	127	138	145
13	26	33	40	53	60	72	80	91	99	111	119	130	138	150	158
14	28	35	43	57	64	77	86	98	106	119	127	140	148	161	169
15	30	38	46	61	69	83	92	105	114	128	137	150	159	173	182
16	32	40	49	65	73	88	98	112	121	136	145	160	169	184	193

Efficient configurations obtained by the brick pattern

Efficient configurations obtained by the comb pattern

Efficient configurations obtained by both the brick and the comb pattern

inefficient maximal configurations is the following

(5.2)
$$\text{minimize} \quad \sum_{i=1}^{m}\sum_{j=1}^{n} X_{i,j}$$
$$\text{subject to} \quad X_{i,j}=0 \Rightarrow \mathbb{P}^E_{i,j} \vee \mathbb{P}^W_{i,j} \vee \mathbb{P}^N_{i,j} \vee \mathbb{P}^C_{i,j},$$
$$X_{i,j} \in \{0,1\}, \quad i=1,\ldots m, \quad j=1,\ldots,n$$

Remark: The constraint $\neg\mathbb{P}^C_{i,j}$ in (5.1) ensures that the resulting configuration is permissible, while the constraint $X_{i,j}=0 \Rightarrow \mathbb{P}^E_{i,j} \vee \mathbb{P}^W_{i,j} \vee \mathbb{P}^N_{i,j} \vee \mathbb{P}^C_{i,j}$ in (5.2) ensures that the resulting configuration is maximal. Although we can put both constraints in both optimization problems, it turns out that it is sufficient to use only one of them in each.

In order to rephrase the above logical constraints into the equivalent algebraic constraints according to the standard integer programming convention, one can consult [14, Chapters 8 and 9]. However, we have used IBM ILOG CPLEX for solving the above integer programs, and IBM Optimization Programming Language (OPL) supports the definitions of constraints in the form of logical constraints. The reference guide for integer programming with IBM ILOG CPLEX can be found in [2].

In Theorem 4.5 we have concluded that the rake–stripe pattern configurations are examples of inefficient configurations. However, we still do not know examples of efficient configurations for all $m, n \in \mathbb{N}$. The only insight into the set of efficient configurations comes from the explicit solutions to the problem (5.1). We have gathered the computed occupancies of efficient configurations in Table 2.
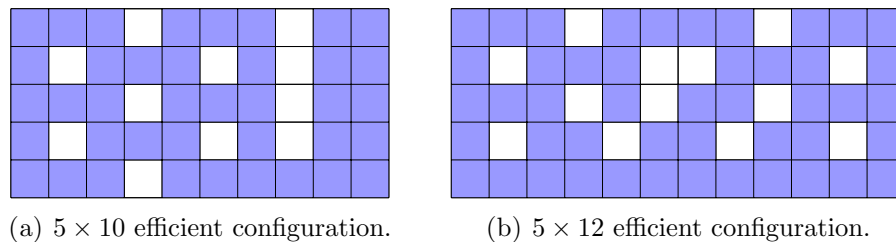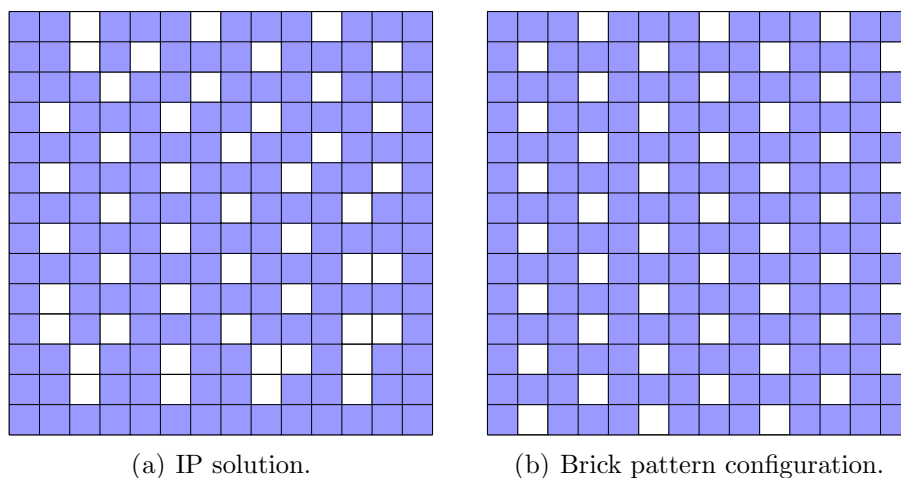
(a) 5×10 efficient configuration. (b) 5×12 efficient configuration.

FIGURE 16. Examples of efficient configurations.

(a) IP solution. (b) Brick pattern configuration.

FIGURE 17. The IP solution to the problem of finding efficient configurations and the corresponding brick pattern configuration on 14×14 grid. Note that both have the maximum occupancy of 148.

Note that the maximum occupancy is often attained by a number of different configurations. In some cases, the IP solutions resemble configurations with one of the special patterns introduced before, such as the brick or comb patterns, see Figure 17.

It is worth noting that a large portion of calculated occupancies of the efficient configurations could not be obtained by either brick nor comb patterns. Some of them are obtained by a certain combination of the two, see Figure 16(a). Some are obtained by the configurations with yet unexplained underlying patterns, see Figure 16(b).

We leave the following question open.

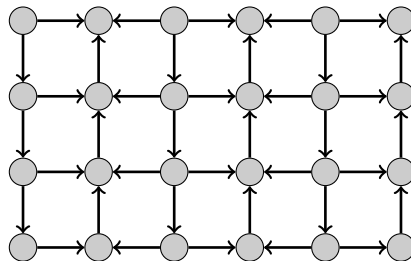**Question.** *What is the explicit expression for $E_{m,n}$?*

FIGURE 18. The graph representation of the 4×6 tract of land.

6. ALTERNATIVE FORMULATIONS OF THE PROBLEM

In this section we give alternative formulations for the problem of finding efficient configurations in the hope of getting more people interested in answering Question 1.

6.1. **Tilings.** The problem of finding *efficient* configurations can be formulated as a tiling problem with overlaps and protrusions allowed, and rotations forbidden as follows. Let C be a maximal configuration on an $m \times n$ grid. Now cover each empty lot by a $\perp$-tetromino which covers that lot as well as the neighboring lots to the east, west and north. Note that some tetrominoes may protrude from the grid, and some may overlap. The permissibility of C guarantees that the $(m-1) \times (n-2)$ subgrid $\{1, 2 \ldots, m-1\} \times \{2, 3, \ldots, n-1\}$ is completely covered by (at least one) $\perp$-tetromino. Conversely, each such a tiling guarantees the permissibility of the corresponding configuration. However, such a tiling does not guarantee that the corresponding configuration is maximal, even in the case when the tiling itself is maximal[2]. Nevertheless, efficient configurations correspond exactly to tilings with the fewest number of tiles. Therefore $E_{m,n} = mn - k$ where k is the number of tiles in an optimal tiling of the $(m-1) \times (n-2)$ grid.

This formulation of the tiling problem does not seem to be very common in literature. For some classical results about T-tetromino tilings see [3, 11, 5, 10].

6.2. **Forbidden induced subgraph problem.** A tract of land can be represented as a rectangular lattice graph where each lot is represented by a vertex. In order to distinguish the north from the south we orient the edges as in Figure 18. The problem of finding efficient configurations is equivalent to finding the largest set of vertices for which the induced graph does not contain either of the subgraphs in Figure 19. The literature on the problem of forbidden induced subgraphs is extensive, see [15, 13].

---

[2]The tiling is maximal if removing any tile leaves some lot in $\{1, 2 \ldots, m-1\} \times \{2, 3, \ldots, n-1\}$ uncovered.
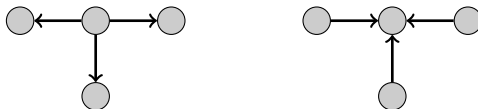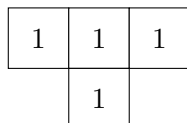
FIGURE 19. The forbidden subgraphs.

FIGURE 20. The forbidden pattern.

6.3. **Subshift on** $\mathbb{Z}^2$**.** Note that permissible configurations on $\mathbb{Z}^2$ can be interpreted as elements of the $\mathbb{Z}^2$-shift of finite type with the alphabet $\mathcal{A} = \{0,1\}$ where the set of forbidden patterns consists of a single pattern in Figure 20, see [4]. An empty lot is represented by 0, and occupied by 1. It is interesting to note that the question of whether the shift space defined by a set of forbidden patterns is empty, in general, is undecidable [1, 9].

6.4. **Dominating set.** In graph theory, a dominating set for a graph $G = (V, E)$ is a subset D of V such that every vertex not in D is adjacent to at least one member of D. The domination number $\gamma(G)$ is the number of vertices in a smallest dominating set for G, see [12]. This problem is usually studied in the setting of an undirected graph, but if we introduce direction to the edges, as in Figure 21, then our problem can be equivalently formulated as the problem of finding the domination number. For domination on digraphs see [7].

In Remark 1.1 we observed that for any tract of land there exists an efficient configuration with all lots along the southern, eastern and western borders occupied. Therefore, the problem of finding efficient configurations can be restricted to finding efficient configurations with all the lots along the southern, eastern and western borders occupied. The empty lots in any such maximal configuration with border lots occupied can be seen as the dominating set for the corresponding directed graph, see Figure 21. Conversely, any dominating set in the digraph can be translated to a maximal configuration with occupied lots on the southern, eastern and western borders. Therefore, the occupancy of the efficient configuration on the $m \times n$ tract of land is exactly $mn - \gamma(G)$ where $\gamma(G)$ is the domination number. Note that the directed graph only represents the inner $(m - 1) \times (n - 2)$ region of the $m \times n$ tract of land without the southern, eastern and western borders.
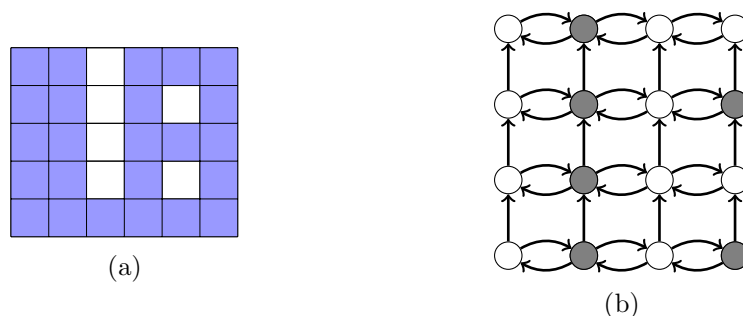
(a)

(b)

FIGURE 21. The efficient configuration on the 5×6 tract of land and the corresponding digraph representation with the dominating set shaded.

7. Concluding remarks

In this chapter, we give several possible directions for further research. One possibility is to move away from rectangular tracts of land and vary their shape; another is to change the shape of lots and the underlying structure of how they fit together. One could consider hexagonal lots where each lot has six neighbors and a house can get sunlight from southeast, south and southwest (but not from north-east, north and north-west). Yet another possibility is to consider multi-story buildings which block the sunlight to more than just one neighboring lot. Note that in this situation it would be natural to assume that the sunlight to different stories of a building could arrive from different sides, which makes this variation of the problem much more complex.

In the accompanying paper [8] we consider a dynamical version of the problem where the houses are built one-by-one. This setup resembles the Random sequential adsorption model which was studied extensively. The main question in these kinds of models is to determine the saturation coverage, i.e. the expected proportion of the land used, when houses are built in random order, respecting the underlying rule. In the aforementioned paper we tackle this problem only with the aid of computer simulations, but it would be even more interesting to tackle the problem in a formal way and obtain theoretical results. The simulations in that paper point to the saturation coverage of approximately 0.636 when the tract of land is large. Since obtaining exact results seems quite involved, one could first consider a toy-model with a rectangular tract of land of a small height m. If one takes a fully built southern-most row, the toy-model reduces to a one-dimensional problem with slightly modified rules.

In the dynamical context, it would be also interesting to study a version of the problem with selfish agents which do not wish to build on a shaded lot, but do not care about others and are allowed to build a house on a lot even if that house is going to block the sunlight to an already built house.

Alternatively, one can consider altruistic agents which do not mind building on a shaded lot, but are not allowed to block sunlight to existing houses. Here, again, one would be interested in the saturation coverage.

Acknowledgments

We wish to thank Juraj Božić who introduced us to this problem that he came up with during his studies at the Faculty of Architecture, University of Zagreb.

We additionally want to thank our colleagues Petar Bakić, Matija Bašić and Stipe Vidak thank to whom one particular instance of this problem ended up in the 10th Middle European Mathematical Olympiad in Vöklabruck, Austria (see [6]).

We also wish to thank Professor Tomislav Došlić for fruitful and stimulating discussions.

Lastly, we wish to thank the anonymous reviewer for their careful reading of our manuscript, their insightful comments and suggestions of possible new perspectives on this problem that are now included in the final section of the paper.

References

1. R. Berger, *The undecidability of the domino problem*, Mem. Amer. Math. Soc. **66** (1966), 72. MR 216954
2. IBM ILOG CPLEX, *V12. 1: User's manual for cplex*, International Business Machines Corporation **46** (2009), no. 53, 157.
3. M. Korn and I. Pak, *Tilings of rectangles with T-tetrominoes*, Theoret. Comput. Sci. **319** (2004), no. 1-3, 3–27. MR 2074945
4. S. A. B. Lemp, *Shift spaces on groups: computability and dynamics*, Ph.D. thesis, Université de Lyon, 2017.
5. C. Merino, *On the number of tilings of the rectangular board with T-tetrominoes*, Australas. J. Combin. **41** (2008), 107–114. MR 2416969
6. Tenth Middle European Mathematical Olympiad, *Contest problems with solutions*, https://www.math.aau.at/MEMO2016/wp-content/uploads/2021/04/MEMO2016_Solutions-8.pdf, 2016.
7. Chaoyi Pang, Rui Zhang, Qing Zhang, and Junhu Wang, *Dominating sets in directed graphs*, Inform. Sci. **180** (2010), no. 19, 3647–3652. MR 2670343
8. M. Puljiz, S. Šebek, and J. Žubrinić, *Packing density of combinatorial settlement planing models*, https://arxiv.org/abs/2107.09417, 2021, Accepted for publication in The American Mathematical Monthly.
9. R. M. Robinson, *Undecidability and nonperiodicity for tilings of the plane*, Invent. Math. **12** (1971), 177–209. MR 297572
10. W. Steurer and S. Deloudi, *Tilings and coverings*, Crystallography of Quasicrystals, Springer, 2009, pp. 7–47.
11. D. W. Walkup, *Covering a rectangle with T-tetrominoes*, Amer. Math. Monthly **72** (1965), 986–988. MR 184108
12. Wikipedia, *Dominating set*, https://en.wikipedia.org/wiki/Dominating_set.
13. _____, *Forbidden graph characterization*, https://en.wikipedia.org/wiki/Forbidden_graph_characterization.
14. H. P. Williams, *Model building in mathematical programming. 5th ed*, 5th ed. ed., Hoboken, NJ: John Wiley & Sons, 2013.

15. T. Zaslavsky, *Forbidden induced subgraphs*, International Conference on Current Trends in Graph Theory and Computation, Electron. Notes Discrete Math., vol. 63, Elsevier Sci. B. V., Amsterdam, 2017, pp. 3–10. MR 3754785

(Mate Puljiz) DEPARTMENT OF APPLIED MATHEMATICS
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING
UNIVERSITY OF ZAGREB
ZAGREB
CROATIA
E-mail address: `mate.puljiz@fer.hr`

(Stjepan Šebek) DEPARTMENT OF APPLIED MATHEMATICS
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING
UNIVERSITY OF ZAGREB
ZAGREB
CROATIA
E-mail address: `stjepan.sebek@fer.hr`

(Josip Žubrinić) DEPARTMENT OF APPLIED MATHEMATICS
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING
UNIVERSITY OF ZAGREB
ZAGREB
CROATIA
E-mail address: `josip.zubrinic@fer.hr`