The Unspeakable Words Puzzle

Croix Gyurek

October 28, 2022

This is a rough draft. I intend to improve the writeup significantly before attempting to publish.

1 Introduction

In the spring of 2021, I noticed a problem, which appeared on the website Puzzling Stack Exchange six years prior, [2], by a user named "Cirdec" about prisoners attempting to survive a warped "trial": A group of p political prisoners have been arrested by a corrupt government. To give the appearance of fairness, the court requires the prisoners to enter a "statement", which must be a single word from the court's official dictionary; each statement will then be revealed to the other prisoner, and then the prisoners must each give "testimony", which is also a single word, and the testimonies must match (or else they are convicted of perjury!). If the testimonies do match, the court, having no genuine evidence, will be forced to release the defendants.

The complication is that, before the "trial", the court gives each prisoner a list of $b = 10$ banned words, which that prisoner is not allowed to use as testimony. The lists may be different for each prisoner.[1] The prisoners are allowed to strategize before receiving the banned word list (and are given copies of the official dictionary), but the court will listen in and will try to prevent the prisoners from escaping. The question is whether or not the prisoners, given the dictionary D (specified in the problem as having at least 170000 words), can form a strategy that guarantees their escape regardless of the court's actions.

The accepted answer on Puzzling Stack Exchange [2] solves this by partitioning the dictionary: one prisoner finds a letter α such that they can legally say any word beginning with α; the other finds a number n such that they can say the n-th word beginning with *any* letter (their statement could be the n-th word starting with "a"). However, since the dictionary is given ahead of time, the actual words do not really matter; the prisoners can simply pretend the dictionary is $D = \{0, 1, \ldots, 11^2 - 1\}$ and use the base-11 representation

[1] There is nothing wrong with prisoner i using a word from prisoner j's banned list as the *statement*, but since the *testimonies* must be identical, the final testimony must be on no one's banned list, or else someone will be held in contempt of court.

of $w \in D$; these numbers can be mapped to the first 121 words of the dictionary to give the actual statements and testimony. Hence, we can think of the dictionary D has a set of natural numbers, or indeed any abstract set of a given size.

This problem can be easily generalized to any number of prisoners p and any number of banned words b; I assume that when $p > 2$, all prisoners will be given all statements. Since the prisoners only have one chance to communicate (their statements), they need to communicate a subset of their non-banned words such that, when all statements are revealed, these "safe sets" have at least one intersection (if there are multiple mutually safe words, the prisoners can agree to choose the alphabetically earliest).

Problem 1.1. *Given $p, b, N \in \mathbb{Z}^+$, let D be a set of size N. Does there exist a collection of p safe-set functions $(h_j : D \to \mathcal{P}(D))_{j=1}^p$ such that:*

1. *Each h_j can avoid any set of size b, that is, for every $j \in \{1, \ldots, p\}$ and $B \subseteq S$ where $|B| = b$, there exists a "safe" $h_j(s)$ such that $h_j(s) \cap B = \emptyset$, and*
2. *For every $P = (s_1, \ldots, s_p)$ where $s_j \in S$, $\bigcap_{j=1}^p h_j(s_j) \neq \emptyset$?*

If so, we call the ordered collection $(h_j)_{j=1}^p$ a solution.

To simplify the problem, we can require the prisoners to adopt the same strategy. This gives rise to the following special case:

Problem 1.2. *Given positive integers p, b, N, let D be a set (the "dictionary") of size N. Does there exist a function $h : D \to \mathcal{P}(D)$ such that:*

1. *Every ban list of size b can be avoided, that is, for every $B \subseteq S$ where $|B| = b$, there exists a "safe" $h(w)$ such that $h(w) \cap B = \emptyset$, and*
2. *Any p safe sets intersect, that is, for every $P = (s_1, \ldots, s_p)$ where $s_j \in D$, $\bigcap_{j=1}^p h(s_j) \neq \emptyset$?*

We call such functions h anonymous solutions.

1.1 Connections to Other Areas of Mathematics

Although the most general form of the problem is a statement about finite sets, it may be helpful to interpret the problem in terms of other areas of mathematics. For example, when restricted to anonymous solutions, the problem can be represented in the language of graph theory as follows:

Problem 1.3. *Given p and b, find a directed graph X with the minimum possible number of vertices, where loops are allowed, such that (i) for every b vertices $\{b_i : 1 \leq i \leq b\}$, there exists a vertex v such that $v \not\to b_i$ for each i, and (ii) for every p vertices $\{v_i : 1 \leq i \leq p\}$, there exists a vertex u such that $v_i \to u$ for each i.*

This does bear some similarity to the notion of *n-existentially closed* graphs, defined in [3] as follows:

Definition 1.4 (Hai, Phuc, and Vinh)**.** *For a positive integer n, a graph (V, E) is* existentially closed *if for every pair of disjoint subsets $A, B \subseteq V$ with $|A| + |B| = n$, there exists a vertex $z \notin A \cup B$ adjacent to every vertex in A and no vertex in B.*

The problem in this paper differs in several respects from the above definition, since the graph X is directed, and the conditions on A and B are separated.

I did not find the graph construction useful, but most of my current solutions involve the use of a group to form function h. To simplify the notation for this case I refer to one of the output sets as itself a "solution":

Definition 1.5. *Given a group $(G, \cdot)$, an* anonymous solution for (p, b, G) *is a subset $M \subseteq G$ such that the function $h : G \to \mathcal{P}(G)$ given by $h(x) = x \cdot M = \{x \cdot m : m \in M\}$ satisfies b-dodging and p-intersection. (M need not be a subgroup, only a subset.)*

Restricting the problem to groups does limit the potential solution sets somewhat (for example, it excludes solutions where the outputs of h have different sizes), but all of my current known solutions use groups, so I will present the sub-problem as follows:

Problem 1.6. *Given p and b, find a group G with minimum order such that there exists $M \subset G$ where (i) for any $\{b_i : i \leq i \leq b\}$, there exists g such that $gM \cap \{b_i\} = \emptyset$, and (ii) $\bigcap_{i=1}^{p} g_i M \neq \emptyset$ for any $\{g_i : 1 \leq i \leq p\} \subseteq G$.*

It will be helpful for analysis to define the two criteria for solutions separately. In the general case:

Definition 1.7. *A function $h : S \to \mathcal{P}(S)$* dodges *b, or* satisfies *b-*dodging*, if and only if for any $B \subset S$, if $|B| \leq b$, then $h(s) \cap B = \emptyset$ for some $s \in S$.*

Definition 1.8. *A function $h : S \to \mathcal{P}(S)$* satisfies *p-*intersection *if and only if for any $s_1, \ldots, s_p \in S$, $h(s_1) \cap \cdots \cap h(s_p)$ is nonempty.*

The 2-intersection property has been discussed in the literature under the name "intersecting families", e.g. [1]. However, usually the term "t-intersecting family" refers to the minimum *size* of the intersection of *two* sets, $|S_i \cap S_j| \geq t$, rather than the *number of sets* which must intersect, $S_{i_1} \cap \cdots \cap S_{i_p} \neq \emptyset$, as in this problem. The "intersecting families" literature also usually does not have $|\mathcal{S}| = N$.

In the graph-theoretic context, $G = (V, E)$ satisfies b-dodging if for every B there exists a vertex $u \in V$ such that $u \not\to B$, and it satisfies p-intersection if for every $P \subseteq V$, where $|P| = p$, there exists a vertex v such that $u \to v$ for *every* $u \in P$.

This allows us to restate the goal as follows:

Definition 1.9. *An* anonymous solution for (p, b, N) *is a function h on a set of cardinality N, such that h satisfies b-dodging and p-intersection. A graph X with N vertices, or a subset M of a group G, with the associated properties can also be termed an* anonymous solution.

2 Naïve Solutions

The accepted answer to the Puzzling Stack Exchange question works as follows: one prisoner finds a letter such that they are allowed to say the first 11 words beginning with that letter; the other finds a number n such that they can say the nth word starting with any letter. This strategy depends on there being 11 or more letters that have 11 or more words, although the answerer noted that the prisoners see the dictionary in the planning phase, so they could adapt the strategy for any dictionary by just using the nth word in the dictionary to represent the number n, and using the digits of $n-1$ in base 11 instead of the word's actual spelling.

In effect, the solution involves making an 11×11 grid out of specific words; one player chooses a "safe" row and the other chooses a "safe" column. In general, for b banned words, two prisoners can win if $N \geq (b+1)^2$. This strategy easily generalizes to an arbitrary number of prisoners, making a p-dimensional grid of size $b+1$, with each prisoner choosing a $(p-1)$-dimensional safe slice and the testimony being the intersection. At this point,

The solution can be made anonymous by letting each statement indicate *both* a safe row and column that intersect at the statement word (or, for $p > 2$, the intersection of one $(p-1)$-dimensional slice in each direction). That is, $S = \{(s_1, \ldots, s_p) : s_i \in \{0, \ldots, b\} \text{ for all } 1 \leq i \leq p\}$, and $h(s) = \{(x_1, \ldots, x_p) : x_i = s_i \text{ for some } 1 \leq i \leq p\}$.

The puzzle's author hinted that there was a second class of solutions for $N \geq (p+1)^b$ instead. One such solution involves constructing a b-dimensional grid of size $(p+1)$. If a prisoner's banned list is $\{w_i : 1 \leq i \leq b\}$, then the prisoner can *remove* one slice in each direction based on the ban list; for instance, let $c = (w_{11}, w_{22}, \ldots, w_{bb})$; then the message set can be $\{(x_1, \ldots, x_p) : x_i \neq c_i \text{ for all } 1 \leq i \leq b\}$. This time, the "center" c might be itself banned, but the prisoners can use (s_i), where $s_i = (c_i + 1) \mod b$, as the statement.

These two solutions are very similar. Switching p and b in the first anonymous solution leads to a grid of the same size as the second, where the message sets of the two are complements of each other. In fact, the criteria of intersection and dodging are complements of one another, as Theorem 2.1 shows:

Theorem 2.1. *Suppose that $h : S \to \mathcal{P}(S)$ is an anonymous solution for (p, b, N). Then (b, p, N) has an anonymous solution as well.*

Proof. Let $h^* : S \to \mathcal{P}(S)$ be defined by $h^*(s) = \{s^* \in S : s \notin h(s^*)\}$. Then we must prove that h^* satisfies b-intersection and p-dodging.

For p-dodging, let $W = \{w_1, \ldots, w_p\} \subset S$. Since h satisfies p-intersection, there is some s so that $s \in h(w_i)$ for all $1 \leq i \leq p$. Then, for any $m \in h^*(s)$, $s \notin h(m)$, and for any $w \in W$, $s \in h(w)$. Therefore, $h^*(s)$ and the ban list W are disjoint. Since W was arbitrary, h^* satisfies p-dodging.

For b-intersection, the proof is similar: let $A = \{s_1, \ldots, s_b\} \subset S$; it suffices to prove intersection when $s_1, \ldots, s_b$ are distinct. Since h satisfies b-dodging, there exists $m \in S$

such that $h(m) \cap A = \emptyset$. Then for each i, $s_i \notin h(m)$, so $s_i \in h^*(m)$ by definition of h^*. Therefore, $m \in h(s_1) \cap \cdots \cap h(s_b)$. Since the s_i were arbitrary, h^* satisfies b-intersection, and since it also satisfies p-dodging, it must be a solution for (b, p, N). □

In particular, when the solution is based on an Abelian group $(G, \cdot)$ with identity e, $M^* = \{s^* \in G : e \notin s^* \cdot M\}$ is equal to the complement of M^{-1}.

3 Improved Bounds

The above solution uses as N the smaller of $(b+1)^p$ and $(p+1)^b$. A natural question is whether or not a smaller N suffices.

For $p = 1$, the naïve solution gives $N \leq b + 1$ or $N \leq 2^b$. Clearly the former is better; indeed, it must be optimal because if $N \leq b$, there would be no acceptable words to use as the prisoner's statement.

3.1 $p = 2$ or $b = 2$

For $p = 2$, the naïve solution gives $N \leq (b+1)^2 = b^2 + 2b + 1$. For $b = 10$ this gives $N = 121$, a number mentioned as minimal for the accepted answer strategy. However, I have found a better solution that uses just over half of that N.

Theorem 3.1. $\left(2, b, \frac{(b+1)(b+2)}{N}\right)$ *and* $\left(p, 2, \frac{(p+1)(p+2)}{N}\right)$ *have anonymous solutions.*

Proof. We proceed in cases depending on whether b is even or odd.

If $b = 2k$, let $G = (\mathbb{Z}/(k+1)\mathbb{Z}) \times (\mathbb{Z}/(2k+1)\mathbb{Z})$ and $M = \{(i, 0) : 0 \leq i < k+1\} \cup \{(0, i) : 1 \leq i \leq k\}$. If $b = 2k - 1$, let $G = (\mathbb{Z}/k\mathbb{Z}) \times (\mathbb{Z}/(2k+1)\mathbb{Z})$ and $M = \{(i, 0) : 0 \leq i < k\} \cup \{(0, i) : 1 \leq i \leq k\}$. (See figure 1 below to visualize.)

It must be shown that these solutions satisfy 2-intersection and b-dodging. The former is easy: given (x_1, y_1) and (x_2, y_2), let $d_1 = (x_1 - x_2) \mod 2k + 1$, and $d_2 = (x_2 - x_1) \mod 2k + 1$. Clearly, $d_1 + d_2 = 2k + 1$, so one of them must be at most k. If it is d_1 (the d_2 case is symmetric), then the point (x_2, y_1) will be in both cosets: it is in M_1 because $(x_2 - x_1 \mod 2k + 1) \leq k$, and it is in M_2 because the first coordinate matches (x_2, y_2).

To prove b-dodging, consider first the case of $b = 2k$, and let $B = \{(x_i, y_i), 1 \leq i \leq 2k\}$, be the ban list. Let C be the set of empty columns, i.e. $C = \{x \in \mathbb{Z}/(2k+1)\mathbb{Z} : x \neq x_i \text{ for all } 1 \leq i \leq b\}$. Because there are $b + 1$ total columns and at most b are removed, C is nonempty.

I now show that there exists $c \in C$ such that at most k bans appear in the next k columns. Call such a column *suitable*. The proof is by induction on $|C|$. If C has one element c, then since only one column is empty, b columns must be occupied, and since there are only b bans, each column apart from c must contain exactly one ban. So there is one ban in each of the k columns after c, for a total of k bans in that region. Therefore, c is suitable.

Suppose now that we can find a suitable empty column whenever $|C| = n$. If $|C| = n+1$, where $n \geq 1$ then there must be at least one column with two or more bans. Choose an arbitrary such column x, relabel B so that b_1 is in column x, and move b_1 to the left, one column at a time, until it reaches an empty column x'. Now, in the new ban list B', there are only n empty columns, and so for some column c, there are at most k elements of B' in the next k columns after c. Now, B is obtained by moving b'_1 from column x' to x. If x is among the next k columns after c, i.e. $(x - c \mod 2k+1) \leq k$, then x' must have been as well, since x' is the nearest empty column left of x, and c remained empty, so $(x' - c \mod 2k+1) < (x - c \mod 2k+1)$. Thus, b_1 is not an additional ban in range of c, so c is still suitable.

Because c is suitable, within the next k columns there are at most k bans. That region, however, has $k+1$ rows, so at least one must be empty. Let r be one such row. Then (c, r) (or, rather, its coset of M) dodges the b bans, because none of the latter are in the same column (because column c is empty), and none of them intersect the rightward arm either (as per above). Thus, the solution satisfies b-dodging.

If $b = 2k-1$, then the proof is similar except that now a suitable column needs to have at most $k-1$ bans within k columns, because the grid only has k rows this time. However, because the grid width is now $b+2$, there must be at least two empty columns. If $|C| = 2$, then there is one ban in each non-empty column. Let c_1, c_2 be the empty columns. Either $(c_2 - c_1 \mod 2k+1)$ or $(c_1 - c_2 \mod 2k+1)$ must be at most k, so one of the two empty columns must be in range of the other. Suppose this is c_2 in range of c_1. Then of the k columns right of c_1, one is empty (c_2), and the other $k-1$ each have one ban. So, c_1 is suitable. If $|C| > 2$, then a very similar induction proof shows that there still exists a suitable column c_1. In either case, there is still at least one more row than bans within range of c, so (c, r) will again create a coset that dodges the b bans.

In both the even and odd cases, a solution has been constructed that satisfies 2-intersection and b-dodging.

□

Figure 1 shows a graphical view of these solutions for $b = 5$ and $b = 6$ (represented as green squares). I have not been able to find any solutions with smaller values of N. Computer searches have ruled out the possibility of cyclic group solutions $(2, b, \mathbb{Z}/N\mathbb{Z})$ for $N < (b+1)(b+2)/2$ when $b \leq 6$. Note that since $k+1$ and k are coprime to $2k+1$, the groups used in the above solutions are actually cyclic.

3.2 $p \geq 3$ and $b \geq 3$

The fact that $(1, b, b+1)$ and $(2, b, \frac{(b+1)(b+2)}{2})$ are solvable suggests that a solution may exist for $(3, b, \frac{(b+1)(b+2)(b+3)}{3!}) = (3, b, \binom{b+3}{3})$, and that more generally, (p, b, N) can be solved for $N = \binom{b+p}{p} = \binom{b+p}{b}$. However, so far, I have not found a general solution for this value of

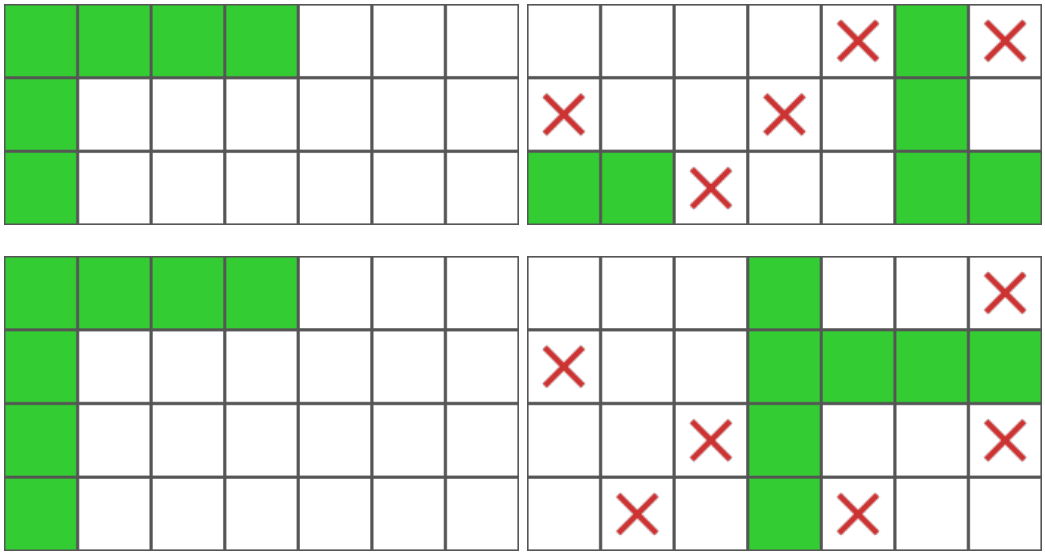

Figure 1: Representations of the improved solutions for two values of b. The left two images show the base sets and the right two images show it moved to dodge b banned tiles (represented by red X's).

N, though I did confirm via computer search that $(3, b, \binom{b+3}{3})$ have cyclic-group solutions for $b \leq 6$.

Conjecture 3.2. *For all $b \geq 1$ and $p \geq 1$, there exists an anonymous solution for $(b, p, \binom{b+p}{p})$.*

However, there is another generalization of the pattern of $p \leq 2$ that works for all values of p and b. Recall the solutions for $p = 2$, which consisted of one column that was full, and half of the remaining columns had one element of M each. In effect, those columns were solutions to $(1, k, k+1)$ or $(1, k-1, k)$ stretched horizontally, and the coset was selected so that only k or $k-1$ bans would reside in that region.

For larger values of p we can proceed in a similar way. Take $p = 3, b = 6$. We can use the same trick as before: our three-dimensional grid will have 7 layers. When p was 2, making less than half of the columns empty allowed for 2-intersection; here we need to make less than one-third (in this case, 2) of the columns empty. Then there will be 1 full layer, 4 partially filled planes, and 2 that are empty.

Therefore, we need the pattern in the partial layers to dodge 4 bans (two can always be moved into empty layers). Also, it needs to satisfy 2-intersection: there will always be a full layer that intersects both of the other two's partial layers (as will be shown in the proof), but the other two partial layers will still need to intersect.

But from Theorem 3.1, we have a solution for 2-intersection and 4-dodging. Each layer, then, would be a 3×5 grid (represented below as an equivalent row of 15). See Figure 2.

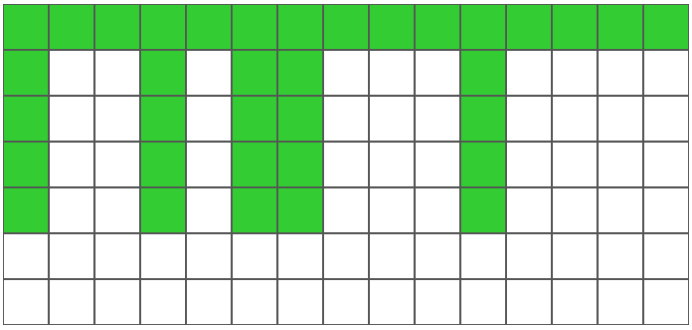

Figure 2: Representation of a solution for $p = 3$, $b = 6$.

This method even generalizes to $p = 4$ and higher. While the above example has the property that $b \equiv 1 \mod p$, the method can be adapted to any b. Before stating the main result, I first present the improved value of N. Let

$$N^*(p, b) = \prod_{i=1}^{p} \left(1 + i \cdot \left\lfloor \frac{i + b - 1}{p} \right\rfloor\right)$$

For small values of p, b the function behaves as follows:

				b			
p	1	2	3	4	5	6	7
2	$1 \cdot 3$	$2 \cdot 3$	$2 \cdot 5$	$3 \cdot 5$	$3 \cdot 7$	$4 \cdot 7$	$4 \cdot 9$
	3	6	10	15	21	28	36
3	$1 \cdot 1 \cdot 4$	$1 \cdot 3 \cdot 4$	$2 \cdot 3 \cdot 4$	$2 \cdot 3 \cdot 7$	$2 \cdot 5 \cdot 7$	$3 \cdot 5 \cdot 7$	$3 \cdot 5 \cdot 10$
	4	12	24	42	70	105	150
4	$1 \cdot 1 \cdot 1 \cdot 5$	$1 \cdot 1 \cdot 4 \cdot 5$	$1 \cdot 3 \cdot 4 \cdot 5$	$2 \cdot 3 \cdot 4 \cdot 5$	$2 \cdot 3 \cdot 4 \cdot 9$	$2 \cdot 3 \cdot 7 \cdot 9$	$2 \cdot 5 \cdot 7 \cdot 9$
	5	20	60	120	216	378	630

Also, note that when $p = b$, $N^*(p, p) = p!$. This is smaller than the original solution's $N = (p + 1)^p$ but larger than the conjectured $\binom{2p}{p} < 4^p$. In general, if p is constant and b increases, $N^*(p, b) \approx \prod_{i=1}^{p} \frac{ib}{p} = b^p \frac{p!}{p^p}$.

In the table, note that the products in each row contain copies of those in previous rows; for example, $N^*(4, 7) = 2 \cdot 5 \cdot 7 \cdot 9$, and $N^*(3, 5) = 2 \cdot 5 \cdot 7$. This is because, as the following theorem shows, we can construct a solution for $(4, 7)$ by using that of $(3, 5)$:

Theorem 3.3. *There exists an anonymous Abelian group solution for $(p, b, N^*(p, b))$, where N^* is defined above.*

Proof. We proceed by induction on p. The case $p = 1$ is trivial: $N^*(1, b) = 1 + (1 + b - 1) = b + 1$. The solution consists of giving as statement and testimony the one word

that is not banned. Suppose now that for a specific $p > 1$, there exists a solution to $(p-1, b, N^*(p-1, b))$ for all $b \geq 1$. I will now prove $(p, b, N^*(p-1, b))$ is solvable for arbitrary $b \geq 1$.

Let $b = kp + r$, where $1 \leq r \leq p$ (note that $r = p$ is possible here). Then let $h = (k+1)p+1$. The solution will use the group $G = H \times \mathbb{Z}/h\mathbb{Z}$, where H will be specified shortly. The set $M \subset G$ will consist of one full layer, $(k+1)(p-1)$ layers containing a solution in H, and $(k+1)$ layers that are empty, arranged in that order. (Let "up" correspond to the direction that moves from the full layer to the partial layers and then to the empty ones.) The solution in H will satisfy $(p-1)$-intersection and $(b-k-1)$-dodging, which must exist by the inductive hypothesis.

Now, with b bans there will be at least $(h-b) = p+1-r$ ban-free layers. Apply the down-slide algorithm to produce a new ban set B', and let its ban-free layers be $e_0 < e_1 < \cdots < e_{p-r}$. Now, $(e_1 - e_0) + (e_2 - e_1) + \cdots + (e_0 - e_{p-r} \mod h) = h = (k+1)p+1$. Therefore, at least one of the terms on the left side must be greater than $\frac{(k+1)p}{p+1-r}$, which is at least $k+1$. So there must be some i such that $(e_{i+1} - e_i \mod h) > k+1$. If we now place the filled layer of our proposed solution at layer e_{i+1}, then all of the other $p-r$ ban-free rows will appear in the partial rows, so that only $(k+1)(p-1) - (p-r) = kp+p-k-1-p+r = kp+r-k-1 = b-k-1$ bans appear there. But since each of those layers contains a copy of a $(b-k-1)$-dodging pattern, there exists a translation of that layer that dodges those bans. Since all other bans are in the empty rows of the pattern, it follows that all b bans have been dodged.

Proving p-intersection is simpler. Given any p cosets, let $f_0 < f_1 < \cdots < f_{p-1}$ be the indices of their full layers. Again, $(f_1 - f_0) + (f_2 - f_1) + \cdots + (f_0 - f_{p-1} \mod h)$ is a sum of p terms adding to $h = (k+1)p+1$, and since $r > 0$, at least one of those differences $(f_{j+1} - f_j)$ must be strictly larger than $k+1$. Then layer f_{j+1} will contain at least one full layer, and all of the other cosets will have either the full layer or a partial layer on f_{j+1}. These are $p-1$ instances of a solution with $(p-1)$-intersection, so they must intersect. Thus, my solution satisfies p-intersection.

This recursive construction guarantees a solution using a group G for all (p, b). It remains to show that $|G| = N^*(p, b)$. By the inductive hypothesis, $|H| = N^*(p-1, b-k-1) = \prod_{i=1}^{p-1} u_i$, where

$$u_i = \left(1 + i \cdot \left\lfloor \frac{i+b-k-2}{p-1} \right\rfloor\right), 1 \leq i \leq p-1$$

Also, $N^*(p, b) = \prod_{i=1}^{p} v_i$, where

$$v_i = \left(1 + i \cdot \left\lfloor \frac{i+b-1}{p} \right\rfloor\right), 1 \leq i \leq p$$

We know that $|G| = h \cdot N^*(p-1, b-k-1)$, so it suffices to show that $u_i = v_i$ for $1 \le i \le p-1$, and that $v_p = h$. The latter is simple:

$$v_p = 1 + p\left\lfloor \frac{p+b-1}{p} \right\rfloor = 1 + p\left\lfloor \frac{p+kp+r-1}{p} \right\rfloor = 1 + p(k+1) + \left\lfloor \frac{r-1}{p} \right\rfloor = h$$

because $r - 1 < p$. Now, if $1 \le i < p$, then we must show that

$$\left\lfloor \frac{i+b-k-2}{p-1} \right\rfloor = \left\lfloor \frac{i+b-1}{p} \right\rfloor$$

Expanding the b in the numerators turns them into $k(p-1)+i+r-2$ and $kp+i+r-1$ respectively. Therefore we must show that

$$\begin{aligned} k + \left\lfloor \frac{i+r-2}{p-1} \right\rfloor &= k + \left\lfloor \frac{i+r-1}{p} \right\rfloor \\ \left\lfloor \frac{i+r-2}{p-1} \right\rfloor &= \left\lfloor \frac{i+r-1}{p} \right\rfloor \end{aligned}$$

Since i and r are strictly less than p, both quantities must be 0 or 1. Furthermore, $i+r-2 < p-1$ if and only if $i+r-1 < p$, so the floored fractions must either both be 0 or both be 1. In either case they are equal.

Therefore, $|G| = h \cdot |H| = h \cdot N^*(p-1)(b-k-1) = N^*(p,b)$. Since b and $p > 1$ were arbitrary, we conclude by induction on p that there exists a $(p, b, N^*(p,b))$ solution for all p and b. □

4 Lower Bounds

My current research has focused almost entirely on *upper* bounds for N. Lower bounds seem to be more difficult, but I was able to prove a lower bound on $p = 2$. We start with some simple cases:

Theorem 4.1. *There is no solution, anonymous or not, for* $(2, 2, 5)$.

Proof. Suppose a solution (h_1, h_2) exists. Consider prisoner 1's function h_1 where $h_1(i) \subseteq \{1,2,3,4,5\}$ for all $i \in \{1,2,3,4,5\}$. Since h_1 satisfies 2-dodging, $|h_1(i)| \le 3$ for all i; however, to satisfy 2-intersection we must have $|h_1(i)| \ge 3$ for all i (if $|h_1(i)| \le 2$ for some i, then B_2 may cover $h_1(i)$, but for some j, $h_2(j) \cap B_2 = h_2(j) \cap h_1(i) = \emptyset$, contradicting 2-intersection). However, there are 10 possible ban lists B_1, each of which allows only one safe set, $\{1,2,3,4,5\} \setminus B_1$. Since the range of h_1 contains only 5 elements, it is impossible for h_1 to satisfy 2-dodging. This is a contradiction, so no such h_1 and hence no such solution (h_1, h_2) exists. □

Theorem 4.2. *There is no solution, anonymous or not, for* $(2, 3, 8)$.

Proof. As before, assume to the contrary that a solution (h_1, h_2) exists. Because $b = 3$ we must have $4 \leq |h_1(i)| \leq 5$. There are $\binom{8}{3} = 56$ ways to choose B_1. Each B_1 allows 5 safe sets of size four, and 1 of size five. (Actually, the set of size five contains the sets of size four, and thus it does not give an extra option.) We would need at least $56/5 > 11$ safe sets, but there are 8 elements in the range of h_1, so we once again have a contradiction. □

The general rule for $p = 2$ is that $(2, b, N)$ is guaranteed to be unsolvable if

$$\binom{N}{b} > N\binom{N-b}{b+1}$$

This can be simplified to

$$\frac{(N-1)!}{(N-b)!}(b+1) > \frac{(N-b)!}{(N-2b-1)!} \tag{1}$$

which can be proven exactly analogously to Theorem 4.2. This requires that $N > 2b$, but if $N \leq 2b$ then the problem is trivially unsolvable because the ban sets can cover the dictionary. We can use (1) to prove an asymptotic lower bound for $p = 2$:

Theorem 4.3. *For every $c > 0$ and $\varepsilon > 0$, if $N \leq cb^{2-\epsilon}$, then for sufficiently large b, there is no solution for $(2, b, N)$.*

Proof. Using an error bound of Stirling's approximation due to Robbins [4], that $n! = \sqrt{2\pi n}(n/e)^n e^{r_n}$, where $\frac{1}{12n+1} < r_n < \frac{1}{12n}$, we can write the following equation, which will imply (1):

$$\frac{\sqrt{(N-1)}\left(\frac{N-1}{e}\right)^{N-1} e^{\frac{1}{12N-12+1}}}{\sqrt{(N-b)}\left(\frac{N-b}{e}\right)^{N-b} e^{\frac{1}{12N-12b}}}(b+1) > \frac{\sqrt{(N-b)}\left(\frac{N-b}{e}\right)^{N-b} e^{\frac{1}{12N-12b}}}{\sqrt{(N-2b-1)}\left(\frac{N-2b-1}{e}\right)^{N-2b-1} e^{\frac{1}{12N-24b-12+1}}}$$

Taking logarithms of both sides and rearranging slightly, we get

$$\begin{aligned}&\ln(b+1) + \tfrac{1}{2}\ln\left(\tfrac{N-1}{N-b}\right) + (N-1)\ln(N-1) + \tfrac{1}{12N-11} - 2(N-b)\ln(N-b) - \tfrac{2}{12N-12b} \\ &> \tfrac{1}{2}\ln\left(\tfrac{N-b}{N-2b-1}\right) + (N-2b-1)\ln(N-2b-1) + \tfrac{1}{12N-24b-11} - 2\end{aligned}$$

where the final -2 term comes from the e^{N-1}, e^{N-b}, e^{N-b}, and e^{N-2b-1} factors.

This can be further simplified to

$$\begin{aligned}&\left(N-b+\frac{1}{2}\right)\ln\left(1-\frac{b^2+2N-2b-1}{N^2-2bN+b^2}\right) + b\ln\left(1+\frac{2b}{N-2b-1}\right) \\ &-\ln\left(\frac{(N-1)(N-2b-1)}{b+1}\right) + 2 + \frac{264N+288b^2-264b-242}{12^3(N-b)\left(N-\frac{11}{12}\right)\left(N-2b-\frac{11}{12}\right)} > 0\end{aligned} \tag{2}$$

Now, suppose that $N = cb^{2-\varepsilon}$. Then the above expression becomes (note that the first term is negative)

$$\left(cb^{2-\varepsilon} - b + \frac{1}{2}\right)\ln\left(1 - \frac{b^2 + 2cb^{2-\varepsilon} - 2b - 1}{c^2b^{4-2\varepsilon} - 2cb^{3-\varepsilon} + b^2}\right) + b\ln\left(1 + \frac{2b}{cb^{2-\varepsilon} - 2b - 1}\right)$$
$$- \ln\left(\frac{(2cb^{2-\varepsilon} - 1)(2cb^{2-\varepsilon} - 2b - 1)}{b+1}\right) + 2 + \frac{264N + 288b^2 - 264b - 242}{12^3\,(N-b)\left(N - \frac{11}{12}\right)\left(N - 2b - \frac{11}{12}\right)}$$
$$> cb^{2-\varepsilon}\ln\left(1 - \frac{b^2 + 2cb^{2-\varepsilon}}{c^2b^{4-2\varepsilon} - 2cb^{3-\varepsilon}}\right) + b\ln\left(1 + \frac{2}{cb^{1-\varepsilon}}\right) - \ln\left(c^2b^{3-2\varepsilon}\right)$$
$$= cb^{2-\varepsilon}\left(-\frac{b^\varepsilon + 2c}{c^2b^{2-\varepsilon} - 2cb} - \frac{(\text{same})^2}{2(1-\xi_1)^2}\right) + b\left(2(cb^{1-\varepsilon})^{-1} - \frac{2(cb^{1-\varepsilon})^{-2}}{2(1+\xi_2)^2}\right) - \ln(c^2b^{3-2\varepsilon})$$

where $0 < \xi_1 < \frac{b^\varepsilon + 2c}{c^2b^{2-\varepsilon} - 2cb}$ and $0 < \xi_2 < \frac{2}{cb^{1-\varepsilon}}$ by Taylor's Theorem. As $b \to \infty$, the only terms that matter are the first term, which is asymptotically $-\frac{b^\varepsilon}{c}$, and the second, which is asymptotically $\frac{2b^\varepsilon}{c}$. For sufficiently large b, this will behave like $\frac{b^\varepsilon}{c}$ and thus be positive, so for such b, there is no $(2, b, N)$ solution.

Technically, the above only works when N is *equal* to $cb^{2-\varepsilon}$. However, if there is no solution for $(2, b, N)$ then there cannot be a solution for $(2, b, N')$ where $N' < N$, since a solution for $(2, b, N')$ can be trivially extended to $(2, b, N)$ by assigning duplicate sets (say, $h(s_1)$) to $h(s_j)$ where $j > N'$.

Therefore, the given b is unsolvable for all $N \le cb^{2-\epsilon}$.

□

Figure 4 shows the largest b such that (1) **does not hold** when N is a power of 2, and hence these b values are *upper* bounds on b for these N:

N	8	16	32	64	128	256	512	1024	2048	4096	2^{13}	2^{14}	2^{15}	2^{16}
b	2	5	9	15	24	39	60	92	139	208	310	458	675	992

Figure 3: Smallest b such that $(2, b, N)$ is guaranteed to have no solution.

Going the other way, we can create the following sequence:

b	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	2	5	8	11	14	18	22	26	31	35	40	45	50	56	61

Figure 4: Largest N such that $(2, b, N)$ is guaranteed to have no solution.

Some admittedly crude numerical experiments appear to suggest an asymptotic approximation $N \approx \frac{b^2}{2\ln(b)}$ for the threshold (the ratio of the two sides would approach 1 as $b \to \infty$).

5 Discussion

The fact that $N^*(p,b) \neq N^*(b,p)$ for many pairs (p,b) suggests that, like the naive $(b+1)^p$ solutions, the N^* solutions are not optimal. Indeed, through various computer searches I have found solutions for $(3,3,\mathbb{Z}/20\mathbb{Z})$, $(3,4,\mathbb{Z}/35\mathbb{Z})$, $(3,5,\mathbb{Z}/56\mathbb{Z})$, $(3,6,\mathbb{Z}/84\mathbb{Z})$, and $(4,4,\mathbb{Z}/70\mathbb{Z})$ (they are presented in Appendix A). However, attempts to generalize the binomial pattern even to $(3,6k,\binom{6k+3}{3})$ have failed. The ideas I have been trying are based on the following:

- A grid of size $(2k+1)\times(3k+1)\times(6k+1)$ gives the right value of N, with a height of $b+1$.
- This means we can use one full layer, like in the N^* solution, and have $4k$ partial layers.
- So far, my ideas have all used three blocks of $2k$ layers each: a "heavy" block adjacent to the filled layer and a "light" block adjacent to an empty block and the heavy block.
- For this to satisfy $6k$-dodging, at the very least, the light and heavy layers need to be able to co-dodge $2k$ bans each ("co-dodge" meaning that given $2k$ bans in each block, there is a *single* translation that allows both blocks to dodge their $2k$ bans simultaneously).
- Additionally, the light block must be able to dodge $4k-1$ bans. (Originally, I believed the required number was $4k$, but if $4k$ bans are in the light block, the other $2k$ must be in the empty block. It may be possible to move the whole pattern $2k$ layers up and put the $4k$ bans in the heavy section. This seems to make the problem worse, but if the light and heavy patterns are constructed so that efficiently blocking the light pattern leaves gaps for the heavy pattern, this may work. In fact, it does work for $(3,6,84)$.)
- Finally, to satisfy 3-intersection, the heavy pattern must be guaranteed to intersect itself and also the light pattern.

In fact, even the binomial coefficients are not always optimal. There is, for instance, a solution for $(3,3,\mathbb{Z}/19\mathbb{Z})$, which is unique for that group up to rotation, reflection, and complement (this was found by simple brute force search). Solutions also exist for $(3,4,\mathbb{Z}/34\mathbb{Z})$ (there appear to be between 8 and 16 distinct solutions when rotations and reflections are ignored). One example of each appears in Appendix A.

Also, the induction step of Theorem 3.3 only requires the existence of an Abelian group solution for $(p-1,b-k-1)$; the construction of the (p,b) solution is independent of the smaller solution. Hence it is, for instance, possible to solve $(5,5,420)$ by applying the construction to the $(4,4,70)$ solution shown in Appendix A.6, even though $N^*(5,5)=720$. The results of applying this method to all $1 \leq p,b \leq 7$ are given in Appendix B.

As for the lower bounds, no obvious sequence derived from Figure 4 ("obvious" meaning the sequence of b, $b+1$, or the inverse sequence of N values needed to reach a certain b) appear to be in the OEIS as of June 3, 2022. (Searching the OEIS for sequences derived from Table B is likely useless because the values themselves are conjectural.)

A natural extension would be to extend the lower bound to any number of prisoners:

Conjecture 5.1. *For any $c > 0$ and $\varepsilon > 0$, there exists N_0 such that for any $N > N_0$, if $b > cN^{1/p+\varepsilon}$ or $p > cN^{1/b+\varepsilon}$, there is no solution, anonymous or not, for (b, p, N).*

A Specific Results

Here are some specific solutions I have found in small cyclic groups. The set M is represented by 1's, so the string "110100" represents the set $\{0, 1, 3\} \subset \mathbb{Z}/6\mathbb{Z}$. For the first four groups I have exhaustively checked all $2^{|G|}$ possible patterns (although in the case of $(3, 4)$ I did limit my search space to between 12 and 17 elements in M). For the others I used a more restrictive search space. I have not searched non-cyclic groups because cyclic groups can be represented using bit vectors for great efficiency.

The "up to rotation" figures do not account for reflection or (when $p = b$) complement. They also do not account for the possibility of automorphisms of the group, because it may be possible for two distinct transformations to return the same output for a given input (indeed, that must be the case for the $(3, 3, \mathbb{Z}/19\mathbb{Z})$ solution and its mirror image).

A.1 $(3, 3, \mathbb{Z}/19\mathbb{Z})$

76 solutions, 4 solutions up to rotation. One example: 0000101011110010011

(This solution is unique up to reflection and complement, at least among group solutions.)

A.2 $(3, 3, \mathbb{Z}/20\mathbb{Z})$

6400 solutions; 320 solutions up to rotation. One example: 00010010100110011111

As a 5×4 grid:

```
01110
00100
01010
11011
```

A.3 $(3, 4, \mathbb{Z}/34\mathbb{Z})$

544 solutions, 16 up to rotation. One example: 0000001011110100100111001101000001

A.4 $(3, 4, \mathbb{Z}/35\mathbb{Z})$

157920 solutions, presumably 4512 up to rotation. One example:
0000100101010110001101001000110011 1

As a 7×5 grid:

```
0000000
1000100
1000100
1000111
1111111
```

A.5 $(3, 5, \mathbb{Z}/56\mathbb{Z})$

I stopped looking after finding 3 solutions. One example:
00010111001001010000100100010001001000110000011100001111

The example is based on this grid pattern (the other two were based on automorphisms of the layers):

```
1111111
0001111
0001111
0001001
0001001
0001001
0000000
0000000
```

There may be solutions in $(\mathbb{Z}/7\mathbb{Z}) \times (\mathbb{Z}/4\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ but that group is not cyclic.

A.6 $(4, 4, \mathbb{Z}/70\mathbb{Z})$

I have found only one solution.
0011100001001110001101011011010100101111010110001100101000010011100111

My search pattern was extremely restrictive; I searched for subsets $H, M, L \subset \mathbb{Z}/14\mathbb{Z}$ such that any translations of the three, as well as translations of their complements, are guaranteed to intersect. The answer as a grid looked like this:

```
11111111111111
01011011011011
00100010011111
00010001100011
00000000000000
```

There may have been an error in my search pattern because automorphisms of the layers were not printed. For instance, the second layer can be changed into 00000111111111 with an affine transformation (if 0 is on the left, $n \to 5n + 4$ works). Alternatively, my assumption that shifting individual layers preserves solutions may have been incorrect.

B Table of Best Known N for $1 \leq p, b \leq 7$

This table has been obtained by taking the known solutions from Appendix A for $p \leq 3, b \leq 3$, and $(p, b) = (4, 4)$, using the recursive formula to fill the upper triangle, and then using Theorem 2.1 to fill $p > b$.

	$b=1$	$b=2$	$b=3$	$b=4$	$b=5$	$b=6$	$b=7$
$p=1$	2	3	4	5	6	7	8
$p=2$	3	6	10	15	21	28	36
$p=3$	4	10	19	34	56	84	150
$p=4$	5	15	34	70	171	306	504
$p=5$	6	21	56	171	420	770	1881
$p=6$	7	28	84	306	770	2940	5460
$p=7$	8	36	150	504	1881	5460	23520

Figure 5: My best known values of N for $1 \leq p, b \leq 7$. The pattern of $N \leq \binom{b+3}{3}$ when $p = 3$ is not known to persist beyond $b = 6$. $\binom{3+7}{3} = 120$, which is less than the entry 150 at $(3, 7)$ and $(7, 3)$.

References

[1] Mengyu Cao, Benjian Lv, and Kaishun Wang. "The structure of large non-trivial t-intersecting families of finite sets". In: *European Journal of Combinatorics* 97 (2021), p. 103373. ISSN: 0195-6698. DOI: https://doi.org/10.1016/j.ejc.2021.103373. URL: https://www.sciencedirect.com/science/article/pii/S0195669821000652.

[2] Cirdec. *Ten unspeakable words*. URL: https://puzzling.stackexchange.com/q/9021. (accessed: 06.02.2022).

[3] Nguyen Minh Hai, Tran Dang Phuc, and Le Anh Vinh. "Existentially closed graphs via permutation polynomials over finite fields". In: *Discrete Applied Mathematics* 214 (2016), pp. 116–125. ISSN: 0166-218X. DOI: https://doi.org/10.1016/j.dam.2016.05.017. URL: https://www.sciencedirect.com/science/article/pii/S0166218X16302426.

[4] Herbert Robbins. "A Remark on Stirling's Formula". In: *The American Mathematical Monthly* 62.1 (1955), pp. 26–29. ISSN: 00029890, 19300972. URL: http://www.jstor.org/stable/2308012 (visited on 05/25/2022).