The Minnesota Journal of Undergraduate Mathematics

I've got your number: a multiple choice guessing game

Ryan Cushman[1] and Adam Hammett[2]

[1]Western Michigan University
[2]Cedarville University

The Minnesota Journal of Undergraduate Mathematics

Volume 3 (2017-2018 Academic Year)

Sponsored by
School of Mathematics
University of Minnesota
Minneapolis, MN 55455

The Minnesota Journal of Undergraduate Mathematics

Volume 3 (2017-2018 Academic Year)

I've got your number: a multiple choice guessing game

Ryan Cushman* [1] and Adam Hammett[2]

[1]Western Michigan University
[2]Cedarville University

Abstract. We consider the following guessing game: fix positive integers k, m, and n. Player A ("Ann") chooses a (uniformly) random integer α from the set $\{1,2,3,\ldots,n\}$, but does not reveal α to Player B ("Gus"). Gus then presents Ann with a k-option multiple choice question about which number she chose, to which Ann responds truthfully. After m such questions have been asked, Gus must attempt to guess the number chosen by Ann. Gus wins if he guesses α. The purpose of this note is to find all "canonical" m-question algorithms which maximize the probability of Gus winning the game. An analysis of a natural extension of this game is also presented.

1. Introduction

Suppose I tell you that I'm thinking of a number between 1 and 3000, and your goal is to correctly guess my number. If I provide you with no additional information, you have a $\frac{1}{3000}$ probability of guessing correctly, which does not allow for much strategy. Thus, to make things more interesting, I will allow you to present me with one 3-option multiple choice question about my number, to which I will respond truthfully. In order for me to be able to answer your question though, I require one (and only one) of your options to be true regardless of the number I chose.

Next, consider the following potential questions you could ask me:

- Q_1: (a) Your number is 1; (b) Your number is 2; (c) Your number is between 3 and 3000.
- Q_2: (a) Your number is 1 or 2; (b) Your number is 3, 4, or 5; (c) Your number is between 6 and 3000.
- Q_3: (a) Your number is 1, 2, or 3; (b) Your number is 4, 5, or 6; (c) Your number is between 7 and 3000.
- Q_4: (a) Your number is between 1 and 1000; (b) Your number is between 1001 and 2000; (c) Your number is between 2001 and 3000.

* *Corresponding author*

Here and below, "between" is understood to be in the inclusive sense. Which of these questions, if any, would (upon my answering the question) yield the greatest probability of you guessing my number? Would you be surprised to discover that each question delivers the *same probability* that you'll guess my number? Indeed, the probabilities that you will guess my number correctly, assuming you guess rationally after I've delivered my response to any of these questions, are all equal to $\frac{3}{3000}$. For example, the probability of guessing correctly if you decided to use Q_1 is given by

$$\frac{1}{3000} \cdot 1 + \frac{1}{3000} \cdot 1 + \frac{2998}{3000} \cdot \frac{1}{2998} = \frac{3}{3000}.$$

Compare this with the probability of winning if you decide to use Q_4, which is likely the question you were inclined to ask:

$$\frac{1000}{3000} \cdot \frac{1}{1000} + \frac{1000}{3000} \cdot \frac{1}{1000} + \frac{1000}{3000} \cdot \frac{1}{1000} = \frac{3}{3000}.$$

Thus Q_4 equips you with the same probability of guessing correctly as Q_1! Should you choose Q_2 or Q_3, the same probability occurs. Also, note that by asking any of the 3-option questions presented above, your probability of winning improves by a factor of 3 over guessing blindly.

Things change considerably if I allow you to present me a second 3-option question after my answer to your initial query. Let us suppose you asked me Q_1. For the readers' convenience, we restate it below:

Q_1: (a) Your number is 1; (b) Your number is 2; (c) Your number is between 3 and 3000.

If I happen to choose "1," then a second question doesn't provide any further information since you already know which number I picked. The same is true if I chose "2." However, if my response comes back "(c)," then a suitably-chosen second question would be profitable. Let us consider asking the following second question:

Q_1': (a) Your number is 3; (b) Your number is 4; (c) Your number is between 5 and 3000.

Now the probability you will win has increased to

$$\frac{1}{3000} \cdot 1 + \frac{1}{3000} \cdot 1 + \frac{2998}{3000}\left(\frac{1}{2998} \cdot 1 + \frac{1}{2998} \cdot 1 + \frac{2996}{2998} \cdot \frac{1}{2996}\right) = \frac{5}{3000}.$$

We remark that this sequence of two 3-option multiple choice questions does not maximize your probability of correctly guessing my number. In fact, it is not hard to find a sequence of two 3-option multiple choice questions which gives you a probability of $\frac{8}{3000}$ of guessing correctly and another such sequence which yields a $\frac{9}{3000}$ probability.

The purpose of this note is to determine the maximum probability of winning the more general version of the game detailed in the abstract, and then to delineate *all strategies* for achieving this maximum probability, in analogy to the binary version of the game considered by Hammett and Oman in [3]. We also provide a natural extension of this game, and give similar results in this setting as well.

2. Preliminaries

The goal of this section is to formalize the sketch of the game presented in the introduction. We begin by fixing some notation. Throughout, we will let $\mathbb{Z}^+$ denote the set $\{1,2,3,\ldots\}$ of positive integers. In addition, for $n \in \mathbb{Z}^+$, the set $\{1,2,3,\ldots,n\}$ will be denoted by $[n]$.

Next, let $k,m,n \in \mathbb{Z}^+$ and let $X \subseteq \mathbb{Z}^+$ have cardinality n. The set X will be presented to both the answerer, "Ann," and the guesser, "Gus." Ann will then pick a number $\alpha \in X$ uniformly at random (known only to Ann). Gus presents m "questions" sequentially, with each question consisting of k "options," *exactly one* of which is correct. Ann responds to each of these m questions in sequence by indicating which of the k options is correct. Both m and k are known by Ann and Gus before the game commences. After Gus presents all m questions and receives Ann's subsequent answers, he must then attempt to guess the number chosen by Ann. Gus wins if he guesses α; otherwise, Ann wins. This game shall be denoted by $\mathcal{G}(k,m,X)$. When $X = [n]$, however, we shall simply write $\mathcal{G}(k,m,n)$.

We now attempt to rigorously define how Gus is allowed to structure his questions, with our primary concern being to ensure that the game eventually ends in a winner and models the rules described in the introduction. To kick things off, consider the following scenario.

Example 2.1. Suppose that Gus presents the following question to Ann in the game $\mathcal{G}(2,1,1)$:

Q: (a) $P = NP$; (b) $P \neq NP$.

Of course, whether or not $P = NP$ is currently one of the biggest unsolved problems in computer science. So Ann has no choice but to guess what the answer to the question is, violating the spirit of the game.

In light of this example, we idealize Ann as follows:

Assumption 1. Ann has perfect knowledge, that is, Ann knows the truth value of every proposition.

To understand a few other possible difficulties, consider the example below.

Example 2.2. In the game $\mathcal{G}(3,1,1000)$, Gus considers presenting one of the following questions to Ann:

Q_1: (a) Your number is between 1 and 9; (b) Your number is between 10 and 19; (c) Your number is between 20 and 30.

Q_2: (a) Your number is between 1 and 400; (b) Your number is between 300 and 600; (c) Your number is between 500 and 1000.

Q_3: (a) The capital of Colorado is Denver; (b) 4 is a prime number; (c) The Earth is flat.

Each of Q_1 or Q_2, if chosen by Gus to present, will cause a dilemma for Ann. In fact, as we will see below in a more thorough discussion, no matter how Ann responds to either of these questions, her response will either provide Gus an extra option per question (so that he's effectively playing the game $\mathcal{G}(4,1,1000)$), or cause the game to end in a stalemate. Thus, we want to rule out questions analogous to these two, on the grounds that the game is negatively or unfairly influenced.

Next, although Q_3 is not very helpful to Gus, it does not cause the same sort of trouble as Q_1 or Q_2 since there are no overlapping references between options and exactly one of them is true. These seemingly unrelated statements can be translated into options that reflect the fact that Ann chose a unique number in [1000]:

Q_3': (a) Your number is a member of [1000]; (b) Your number is a member of the empty set; (c) Your number is a member of the empty set.

Notice that, in regards to Gus's knowledge about Ann's choice, presenting Q_3' rather than Q_3 puts Gus in the same position. Both Q_3 and Q_3' effectively ask "Have you picked a number between 1 and 1000?" We will see below that it is always possible to transform a question Q in the game $\mathcal{G}(k,m,X)$ into an equivalent question Q' of k logical disjunctives on the set X. Ann's choosing a unique element of X is a critical ingredient of this equivalence.

We need to carefully consider Example 2.2, as this will form the basis for our remaining assumptions about how the game is played. First, we offer some further explanation concerning the ill-conceived questions Q_1 and Q_2.

The question Q_1 is problematic because the interval $(30,1000]$ is not included, and so the options are inexhaustive. Indeed, if Ann's choice $\alpha > 30$ then she will be forced into a predicament. She may choose to reveal that none of the options in Q_1 are true, but this would effectively indicate Ann's affirmation of a hidden fourth option: "Your number is between 31 and 1000." So this would permit Gus to play as if he were in the game $\mathcal{G}(4,1,1000)$, and not the game we intended! On the other hand, should Ann refuse to answer, then the game ends in a stalemate. Either way the game has been compromised, and consequently "inexhaustive options" like those in Q_1 ought to be disallowed.

A similar problem arises should Gus present Q_2 to Ann. Indeed, if there are multiple options correct then Ann will once again face a dilemma. If Gus presents Q_2 and Ann's choice $\alpha \in [300,400]$, she could either indicate that both (a) and (b) are true, once again affirming a hidden fourth option ("Your number is between 300 and 400."), or she could refuse to answer due to ambiguity, and the game will end in a stalemate. Either outcome is bad, so "overlapping options" like those in Q_2 should be disallowed as well. Thus, we arrive at our second assumption:

Assumption 2. All "questions" presented by Gus consist of "options" that are propositions, precisely one of which is true no matter which $\alpha \in X$ Ann chose.

At its core Assumption 2 prohibits Gus from presenting questions like Q_1 and Q_2, thus preserving the integrity of the game. Henceforth we shall refer to questions that adhere to Assumption 2 as "allowable." Before turning to the problems implicit in question Q_3, we need to take a brief detour into the propositional logic. The reason for the foray into logic that follows is to rigorously delineate the sorts of options that we may assume Gus will present in his questions, and we will show that disjunctive propositions like "Your number is 1 or 2 or 3." are without loss of generality the best Gus can do in terms of extracting information from Ann. In other words, outlandish, esoteric options within an allowable (Assumption 2) question are always equivalent to an allowable question whose options are simple disjunctives!

Suppose we are in the game $\mathcal{G}(k,m,X)$, let $X = \{x_1,\ldots,x_n\}$ and let's imagine that Gus presents as a question option a proposition P in the literals $\ell_1,\ldots,\ell_n \in \{0,1\}$. Ann then responds to the option P by setting $\ell_i = 1$ (indicating "true") and the other $\ell_j = 0$ (indicating "false") precisely if her choice $\alpha = x_i$. She then returns the truth value of P to Gus under her literal assignment. Importantly, since Ann chose a unique $\alpha \in X$, we have $\alpha = x_i$ if and only if $\ell_i = 1$ and $\ell_j = 0$ for all $j \neq i$.

A bit more formally, for each $i \in [n]$ letting $\mathbf{e}_i \in \{0,1\}^n$ denote the n-dimensional binary vector with ith entry 1 and 0s elsewhere, we have $\boldsymbol{\ell} := (\ell_1,\ldots,\ell_n) = \mathbf{e}_i$ if and only if $\alpha = x_i$. Letting $\mathcal{E} := \{\mathbf{e}_i : i \in [n]\}$, note that in this setting two propositions are logically equivalent precisely if they agree for all n special literal assignments belonging to $\mathcal{E}$; let's denote this special logical equivalence by "$\equiv_{\mathcal{E}}$". Let $G := \{i \in [n] : P \text{ is true if } \boldsymbol{\ell} = \mathbf{e}_i\}$. Then it is straightforward to see that the disjunctive proposition $P' := \vee_{i\in G}\ell_i$ is such that $P \equiv_{\mathcal{E}} P'$. Indeed, P' as we've defined it is true exactly when $\boldsymbol{\ell} = \mathbf{e}_i$ with $i \in G$, and these are *by definition* exactly the literal assignments in $\mathcal{E}$ for which P is true. An illustration of this equivalence where $n = 5$ is provided in Table 1, and we implore the interested reader to work through it carefully.

$\boldsymbol{\ell}$	$P := ((\ell_1 \wedge \neg\ell_4) \vee \ell_5) \Rightarrow (\neg\ell_2 \wedge \ell_3)$	$P' := \ell_2 \vee \ell_3 \vee \ell_4$
$\mathbf{e}_1$	0	0
$\mathbf{e}_2$	1	1
$\mathbf{e}_3$	1	1
$\mathbf{e}_4$	1	1
$\mathbf{e}_5$	0	0

Table 1. Truth table showing that $P \equiv_{\mathcal{E}} P'$ when $n = 5$.

Importantly, the logical equivalence of P and P' is *only guaranteed* valid for literal assignments belonging to $\mathcal{E}$. In our $n = 5$ example, for instance, the assignment $\ell_1 = \ell_2 = \ell_5 = 1$ and $\ell_3 = \ell_4 = 0$ shows that the equivalence $P \equiv_{\mathcal{E}} P'$ does not necessarily extend to *all possible* literal assignments. Ponder for a moment what we have just shown in this example: Gus presenting the complicated question option "If your number is 1 and not 4, or your number is 5, then your number is not 2 and your number is 3." is $\mathcal{E}$-equivalent to his simply presenting "Your number is 2 or 3 or 4."

Why all the fuss over logical formalities here? In essence, this discourse rigorously justifies something that our intuition says *must* be the case: any option presented by Gus in a question may as well be a proposition that is a logical disjunctive in the literals $\ell_1,\ldots,\ell_n$. Moreover (Assumption 2), in aggregate these propositional disjunctives must be exclusive (no two options can be simultaneously true) and exhaustive (precisely one option is true for any given choice $\alpha \in X$ by Ann). Indeed, even if an option is initially presented as a seemingly unrelated statement like in Q_3 above, this option can be transformed into an equivalent proposition P in the literals $\ell_1,\ldots,\ell_n$, which in turn can be transformed into an $\mathcal{E}$-equivalent proposition P' of the disjunctive type just described. This was at the heart of our Example 2.2 reasoning that Gus could have presented question Q_3' instead, and gleaned the same information from Ann. We formalize this in the next assumption:

Assumption 3. All "options" presented by Gus are propositions of the form "Your number is a member of G." for some $G \subseteq X$.

Notice that we allow for the possibilities that $G = \varnothing$ or $G = X$ in Assumption 3. In fact, this occurred in our question Q_3' of Example 2.2. Clearly, Gus's presenting the proposition "Your number is a member of G." to Ann for her evaluation is equivalent to Gus presenting her with the logical disjunctive $\vee_{i \in G} \ell_i$, to which Ann returns the truth value.

Assumption 3 should not be misunderstood to mean that every manifestation of the game will involve a highbrow presentation of symbolic disjunctives by Gus. Rather, Assumptions 2–3 are meant to delineate the sorts of options that we may assume Gus will present in his questions, and we have managed to show that disjunctive propositions like "Your number is 1 or 2 or 3." are without loss of generality the best Gus can do for the purposes of extracting information from Ann.

3. Setting the Stage

With most of the technical preliminaries out of the way, we now set up notation which will move us toward the formalized game. First, we establish a probabilistic paradigm. Recall that the game $\mathcal{G}(k, m, X)$ begins with Ann choosing, uniformly at random, an integer $\alpha \in X$. The game concludes with Gus making a guess $\gamma \in X$ based upon the answers he receives from Ann to his m questions, and Gus wins the game if and only if $\gamma = \alpha$. Throughout the remainder of this paper, we shall denote Ann's choice by α and Gus's guess by γ. In this setting, we regard X as a probability space endowed with the uniform distribution

$$\mathbb{P}(x) = \frac{1}{|X|} \text{ for all } x \in X.$$

We formalize Gus's presentation of questions with the following definition.

Definition 3.1. Let Q_r denote the rth question presented by Gus, $1 \le r \le m$, and let $G_{r,i}$ denote the event "Ann selects option i in Q_r" on the probability space X, $1 \le i \le k$.

Notice that by Assumption 3, any option in Q_r must be a proposition of the form "Your number is a member of G." for some $G \subseteq X$. Thus, it follows that the events $G_{r,i}$ as we've defined them are precisely the subsets of X such that the ith option of Q_r is "Your number is a member of $G_{r,i}$." The following example further clarifies this relationship.

Example 3.2. Consider the game $\mathcal{G}(3, 1, 5)$ in which Gus asks the following question:

Q_1: (a) Your number is 1 or 5; (b) Your number is 2 or 4; (c) Your number is 3.

The options of Gus's question correspond to the respective events $G_{1,1} = \{1, 5\}$, $G_{1,2} = \{2, 4\}$, and $G_{1,3} = \{3\}$. So, if $\alpha = 4$, $G_{1,2}$ would occur. Furthermore, we can say that $\mathbb{P}(G_{1,1}) = \frac{2}{5}$, $\mathbb{P}(G_{1,2}) = \frac{2}{5}$, and $\mathbb{P}(G_{1,3}) = \frac{1}{5}$.

In light of Assumptions 1–3 above, we can now establish an upper bound on the probability of Gus winning $\mathcal{G}(k,m,X)$. This is accomplished by using Assumption 2 to ensure the k^m events that could result in a winning game are mutually exclusive. Then, we use conditional probability to show that the likelihood of any such winning event is at most $1/|X|$.

Theorem 3.3. *The probability that Gus wins the game $\mathcal{G}(k,m,X)$ is at most $\frac{\min(|X|,k^m)}{|X|}$.*

Proof. Gus will win $\mathcal{G}(k,m,X)$ with probability at most $1 = \frac{|X|}{|X|}$. Thus it suffices only to prove that the probability of Gus winning is at most $\frac{k^m}{|X|}$. Let's assume that Ann has chosen $\alpha \in X$ uniformly at random, and fix an arbitrary m-tuple $(Q_1,\ldots,Q_m)$ of questions to be presented by Gus (in this order). Let W denote the event "Gus's guess γ is equal to α." In other words, W is the event "Gus wins the game."

Recall from Definition 3.1 above that for any question Q_r ($r \in [m]$), we let $G_{r,i}$ ($i \in [k]$) denote the event "Ann selects option i in Q_r" on the probability space X. Thus, W occurs if and only if $G_{1,i_1} \cap G_{2,i_2} \cap \cdots \cap G_{m,i_m} \cap W$ occurs for some $(i_1,i_2,\ldots,i_m) \in [k]^m$. Since Assumption 2 guarantees a unique event $G_{r,i}$ occurs for each Q_r, it follows that for $(i_1,\ldots,i_m) \neq (j_1,\ldots,j_m)$ the events $G_{1,i_1} \cap \cdots \cap G_{m,i_m} \cap W$ and $G_{1,j_1} \cap \cdots \cap G_{m,j_m} \cap W$ are mutually exclusive. Hence, we obtain

$$\mathbb{P}(W) = \sum_{(i_1,\ldots,i_m)\in[k]^m} \mathbb{P}(i_1,\ldots,i_m,W), \tag{1}$$

where $(i_1,\ldots,i_m,W)$ is the vector naming the event $G_{1,i_1} \cap \cdots \cap G_{m,i_m} \cap W$. Since there are k^m such vectors, it suffices to show that $\mathbb{P}(\mathbf{v}) \leq \frac{1}{|X|}$ for any such vector $\mathbf{v}$. Thus define $G \subseteq X$ as the "guessing set" from which Gus guesses γ uniformly at random after presenting all m questions. Then let $\mathbf{v} := (i_1,\ldots,i_m,W)$ be arbitrary. Observe that

$$\mathbf{v} \text{ occurs if and only if } \alpha \in A := G_{1,i_1} \cap \cdots \cap G_{m,i_m} \text{ and Gus's guess } \gamma \text{ equals } \alpha. \tag{2}$$

Thus if $A = \emptyset$, then $\mathbf{v}$ cannot occur, and $\mathbb{P}(\mathbf{v}) = 0 < \frac{1}{|X|}$. Assume now that $A \neq \emptyset$. Then

$$\begin{aligned} \mathbb{P}(\mathbf{v}) &= \mathbb{P}(\alpha \in A) \cdot \mathbb{P}_G(\gamma = \alpha \mid \alpha \in A) \\ &= \frac{|A|}{|X|} \cdot \frac{|A \cap G|}{|A| \cdot |G|} = \frac{|A \cap G|}{|G|} \cdot \frac{1}{|X|} \leq \frac{1}{|X|}; \end{aligned} \tag{3}$$

here, $\mathbb{P}_G(\gamma = \alpha \mid \alpha \in A)$ denotes the conditional probability that $\gamma = \alpha$ *given* that $\alpha \in A$ and Gus's guess is uniform on G. To conclude the proof, we justify why $\mathbb{P}_G(\gamma = \alpha \mid \alpha \in A) = \frac{|A\cap G|}{|A|\cdot|G|}$. Since Ann's choice lies in A, there are a total of $|A| \cdot |G|$ equally likely pairs (α,γ) of possible choices by Ann and Gus, respectively. Gus will win provided the pair (α,γ) satisfies $\alpha = \gamma$, and there are $|A \cap G|$ such pairs. The proof is now complete. ☐

Of course, the numerator k^m appearing in the upper bound on $\mathbb{P}(W)$ above is no coincidence, as the "obvious" strategy is for Gus to continuously divide up the set he knows contains α into k roughly equal parts. In this way, Gus can effectively improve his probability of winning by a factor of k with each of his m questions. This has clear conceptual parallels with the binary search algorithm in computer science – but instead of bisecting at each stage, we "k-sect" the set containing the element in question. For example, consider searching for the number 2 in the array of numbers $(1,2,3,4,5,6,7,8,9)$. If we consider the $k=3$ variation, we can divide this up into 3 contiguous subarrays of length $\frac{9}{3}=3$. When we get back $(1,2,3)$, we again divide this up into 3 separate length $\frac{3}{3}=1$ subarrays and get back the subarray (2), and we have found our number.

Notice that from Theorem 3.3 the upper bound on the probability of winning with 2 queries if we are allowed to trisect at each query is $\frac{\min(9,3^2)}{9}=1$, so it is no surprise that only 2 stages are necessary to locate one of nine numbers if trisections are permitted. The purpose of the next section is to show that even with much less stringent assumptions (recall that Gus is free to ask Ann *any* allowable question (consisting of propositions); the question need not even obviously relate to the game being played), Gus still cannot do any better than he can by adopting a natural set of additional rules for game play. We shall shortly introduce such a set of rules, and then find all strategies which maximize Gus's probability of winning in this modified setting. Our justification for Assumption 3 was the first major step in this direction, as it validates that this new set of rules elicits no loss of generality.

We conclude this section with an interesting application of Theorem 3.3. Suppose we are in the game $\mathcal{G}(3,1,4)$. Then in accordance with Theorem 3.3, the probability that Gus will win the game is at most $\frac{\min(4,3^1)}{4}=\frac{3}{4}$. In other words, no matter how clever Gus is with his 3-option question, he cannot guarantee ahead of time that he will be able to correctly guess Ann's number. Let's see why this is reasonable in the setting of propositional logic. Putting Assumption 3 aside for the moment, suppose by way of contradiction that there is, in fact, an allowable question consisting of three (potentially very complicated) propositions P_1, P_2, and P_3 that will ensure Gus's winning regardless of the number Ann chose. Let S be the statement "Ann will choose a uniformly random number $\alpha \in [4]$" (observe that Gus knows S). For each $i \in [3]$, let $S_i := P_i \wedge S$. Then since we are assuming Gus will be victorious, we must have $S_1 \Rightarrow$ "Ann chose a," $S_2 \Rightarrow$ "Ann chose b," and $S_3 \Rightarrow$ "Ann chose c" for some subset $\{a,b,c\} \subseteq [4]$. It follows that $S_1 \vee S_2 \vee S_3 \Rightarrow$ "Ann chose a or b or c." But by Assumption 2, $S_1 \vee S_2 \vee S_3$ is true. So by the modus ponens rule of inference, it follows that Ann's choice is either a or b or c. But as $|[4]|=4$, this clearly need not be so!

A completely analogous argument to the above can be made in the game $\mathcal{G}(n-1,1,n)$ for any $n \geq 2$. We refer the interested reader to [1], [2], and [4] for further reading on probability, logic, and algorithms.

4. Main Results

We are now ready to introduce the canonical version of the game, with Definition 3.1 above governing our notation. **Henceforth, we shall assume $\mathcal{G}(k,m,X)$ to be as defined in Definition 4.1 below unless stated otherwise.**

Definition 4.1 (The game $\mathcal{G}(k,m,X)$, canonical version)**.** A finite, nonempty set $X \subseteq \mathbb{Z}^+$ is presented to Ann and Gus. Ann randomly chooses a number $\alpha \in X$ known only to her. Further, Gus is given positive integers m and k.

For each $r \in [m]$, Gus will present to Ann an ordered k-tuple of disjoint subsets of X, which we denote by $Q_r := (G_{r,1}, G_{r,2}, \ldots, G_{r,k})$. Ann then responds by returning $A_r := G_{r,a_r}$ if and only if $a_r \in [k]$ is such that $\alpha \in G_{r,a_r}$. To be clear, this all transpires sequentially in the sense that Gus first presents Q_1 to Ann, to which she responds with A_1, at which point Gus then presents Q_2 to Ann, to which she responds with A_2, and so on. After Gus has presented all m "questions" $(Q_1, \ldots, Q_m)$ to Ann, and subsequently received all m responses $(A_1, \ldots, A_m)$ from Ann, Gus then attempts to guess the number Ann chose. Now set $A_0 := X$. We further require the following:

(1) Gus's question $Q_r = (G_{r,1}, \ldots, G_{r,k})$ satisfies $\bigsqcup_{1 \le i \le k} G_{r,i} = A_{r-1}$ for all $r \in [m]$, and

(2) Gus's guess γ is a random member of A_m (that is, A_m is Gus's "guessing set," as in the proof of Theorem 3.3).

Gus wins if his guess $\gamma = \alpha$.

Schematically, here is an effective way to visualize the canonical game:

$$\alpha \in X = A_0 \overset{\xrightarrow{Q_1}}{\supseteq} A_1 \overset{\xrightarrow{Q_2}}{\supseteq} A_2 \overset{\xrightarrow{Q_3}}{\supseteq} \cdots \overset{\xrightarrow{Q_m}}{\supseteq} A_m \ni \alpha, \gamma. \tag{4}$$

Indeed, (1) guarantees that $A_0 \supseteq A_1 \supseteq \cdots \supseteq A_m \ni \alpha$, and (2) that Gus's guess $\gamma \in A_m$.

A few words of justification are in order, as we want it to be evident that the rules of the canonical game really engender no loss of generality. First, we have already seen (Assumptions 2–3) that any option within a given question is a proposition, and that this proposition is equivalent to "Your number is a member of G." for some $G \subseteq X$. Thus, it is clear that we can codify any question Q_r ($r \in [m]$) by its corresponding subsets of X, namely $G_{r,1} \subseteq X$ corresponds to the first option, $G_{r,2} \subseteq X$ to the second, and so on. We have formalized this in the canonical game by writing each question as a k-tuple $Q_r = (G_{r,1}, G_{r,2}, \ldots, G_{r,k})$, as these subsets of X are really all that matter.

Next, by Assumption 2 we see that for each $r \in [m]$ question Q_r must have precisely one option that is true. This means that there is one (and only one) integer $a_r \in [k]$ such that Ann answers "true" to the proposition "Your number is a member of G_{r,a_r}." Specifically, this demands that these subsets of X be disjoint, and that Ann's choice α belong to one of them. We have codified this in Ann's returning $A_r = G_{r,a_r}$ to Gus, the unique subset in this rth question such that her choice $\alpha \in G_{r,a_r}$.

Finally, we justify the additional rules (1) and (2). To see why (1) is reasonable (and does not endanger generality), assume that Gus is about to present question Q_r to Ann, for some $r \in [m]$. This means that Gus *knows* that Ann's choice $\alpha \in A_{r-1}$. So how ought he structure his question $Q_r = (G_{r,1}, \dots, G_{r,k})$? In order to satisfy Assumption 2, it follows that the disjoint union $\bigsqcup_{1 \le i \le k} G_{r,i} \supseteq A_{r-1}$. Of course, Gus *could* in theory structure Q_r so that this containment is proper, but why? He would not glean any more information by doing so, since such a question would include elements that Gus *already knows* are not equal to α. Recalling that our aim is to *maximize* Gus's probability of winning, we thus impose (1) without apology.

And at the heart of our requirement (2) is the inequality (3) of Theorem 3.3. This inequality makes it clear that Gus cannot improve his probability of winning by including elements in his guessing set that he knows do not equal α. In fact, (3) shows that doing so causes Gus's probability of winning to *decrease*. Thus, to make certain that Gus *maximizes* this probability, we require (2).

At long last we have a rigorous, general, relatively simple framework in the canonical game to aid our analysis. The primary goal of this section will be to utilize this framework, exclusively, to deliver all possible algorithms Gus may deploy to maximize his probability of winning the game. Moreover, we can rest easy knowing that there aren't *other* mysterious, probability-maximizing algorithms lurking beyond our analysis. Indeed, even though this framework appears very rigid and specific, we have in fact shown that if our goal is to maximize Gus's probability of winning, then any manifestation of the game doing so has an equivalent canonical game counterpart! Let's leave the abstraction for a moment to see the canonical game in action.

Example 4.2. Gus and Ann are playing the game $\mathcal{G}(3,3,300)$, and Ann has chosen $\alpha = 36$. Before starting, note that $A_0 = X = [300]$.

- Gus gives his first question $Q_1 = (G_{1,1}, G_{1,2}, G_{1,3}) = ([1,42],[43,246],[247,300])$. Ann then returns $A_1 = G_{1,1} = [1,42]$ since $\alpha = 36 \in [1,42]$.
- Gus's next question must follow (1) of Definition 4.1, and $A_1 = [1,42]$, so he devises $Q_2 = ([1,7],[8,27],[28,42])$, to which Ann returns $A_2 = G_{2,3} = [28,42]$.
- Gus crafts and presents his final question $Q_3 = ([28,35],[36,37],[38,42])$, to which Ann returns $A_3 = G_{3,2} = [36,37]$.

Now that (Q_1,Q_2,Q_3) have been submitted and (A_1,A_2,A_3) returned, Gus must guess a random member of A_3 (as he is bound by (2) of Definition 4.1) – either 36 or 37. Gus decides to guess $\gamma = 36$. Since $\gamma = \alpha$, he wins.

As the disjoint union in (1) of Definition 4.1 is a rather cumbersome expression, for each $r \in [m]$ and question $Q_r = (G_{r,1}, \dots, G_{r,k})$ we introduce the notation

$$\bigsqcup Q_r := \bigsqcup_{1 \le i \le k} G_{r,i}.$$

Observe that requirement (1) can now be written simply as $\bigsqcup Q_r = A_{r-1}$. We need one more definition before proceeding.

Definition 4.3. Let X be a finite, nonempty subset of $\mathbb{Z}^+$, and let $m,k \in \mathbb{Z}^+$. For each $r \in [m]$, suppose $Q_r := (G_{r,1}, \dots, G_{r,k})$ is a k-tuple of disjoint subsets of X. Lastly, let $\alpha, \gamma \in X$. Then we call the sequence $\mathbf{g} := (\alpha, Q_1, \dots, Q_m, \gamma)$ a **game vector** of the game $\mathcal{G}(k,m,X)$. Further, we say that $\mathbf{g}$ is **allowable** in the game $\mathcal{G}(k,m,X)$ provided that every Q_r satisfies

(1) of Definition 4.1 relative to α, and γ satisfies (2) of Definition 4.1. Lastly, **g** is **winning** if **g** is allowable and $\gamma = \alpha$.

The definition of a game vector should remind the reader of the visual schematic (4) above in a more compact form. We can immediately tell if a game vector $\mathbf{g} = (\alpha, Q_1, \ldots, Q_m, \gamma)$ is "allowable" and/or "winning" in $\mathcal{G}(k,m,X)$. We summarize how in the following proposition.

Proposition 4.4. *Let* $\mathbf{g} := (\alpha, Q_1, \ldots, Q_m, \gamma)$ *be a game vector of the game* $\mathcal{G}(k,m,X)$, *with* $Q_r := (G_{r,1}, \ldots, G_{r,k})$ *for each* $r \in [m]$. *Then* **g** *is "allowable" if and only if there exists* $(a_1, \ldots, a_m) \in [k]^m$ *such that*

(a) $\bigsqcup Q_1 = X$,

(b) $\bigsqcup Q_r = G_{r-1,a_{r-1}}$ *for* $1 < r \leq m$, *and*

(c) $\alpha, \gamma \in G_{m,a_m}$.

Moreover, **g** *is "winning" if in addition we have* $\gamma = \alpha$.

Proof. Clearly, if **g** is "allowable," then we have (a), (b), and (c) from Definition 4.1. For the other direction, observe that (a) is the $r = 1$ requirement of (1) in Definition 4.1. Further, note that $G_{r,a_r} = A_r$, Ann's rth response, for each $1 < r \leq m$, and hence (b) fulfills the rest of (1). And lastly, (c) is equivalent to (2). Note well that

$$\alpha, \gamma \in G_{m,a_m} \subseteq G_{m-1,a_{m-1}} \subseteq \cdots \subseteq G_{1,a_1} \subseteq X,$$

which is reminiscent of (4) above. And since any allowable game vector with $\alpha = \gamma$ is winning by Definition 4.3, we have the last statement. $\square$

Proposition 4.4 enables us to correspond with an allowable game vector $(\alpha, Q_1, \ldots, Q_m, \gamma)$ the options $(a_1, \ldots, a_m)$ that were selected by Ann as "true" in the game $\mathcal{G}(k,m,X)$. Let's refer to this vector of option-selections $(a_1, \ldots, a_m)$ as Ann's "option vector." To help the reader intuit these matters, let's pause for another example.

Example 4.5. Consider the game $\mathcal{G}(2,3,9)$; that is, Gus is allotted three 2-option questions before guessing. Suppose that Ann chooses $\alpha = 1 \in [9]$. Set $Q_1 := (\{2,4,8\}, \{1,3,5,6,7,9\})$, $Q_2 := (\{1,3,5,9\}, \{6,7\})$, $Q_2' := (\{1,2,5,9\}, \{3,6,7\})$, and $Q_3 := (\{1,5\}, \{3,9\})$. Then the game vector $\mathbf{g}_1 := (1, Q_1, Q_2, Q_3, 5)$ is allowable and $\mathbf{g}_2 := (1, Q_1, Q_2, Q_3, 1)$ is winning, and each has option vector $(a_1, a_2, a_3) = (2,1,1)$. However, the game vector $\mathbf{g}_3 := (1, Q_1, Q_2, Q_3, 6)$ is not allowable since $\gamma = 6 \notin \bigsqcup Q_3$, nor is $\mathbf{g}_4 := (1, Q_1, Q_2, Q_3, 3)$ allowable since $\alpha = 1 \in \{1,5\}$ and $\gamma = 3 \notin \{1,5\}$. Finally, $\mathbf{g}_5 := (1, Q_1, Q_2', Q_3, 1)$ is not allowable either, since $\bigsqcup Q_2' \neq \{2,4,8\}$ and $\bigsqcup Q_2' \neq \{1,3,5,6,7,9\}$.

We now establish a proposition which will be heavily utilized throughout the remainder of the paper.

Proposition 4.6. *Let* X *be a finite, nonempty subset of* $\mathbb{Z}^+$, *and let* $m, k \in \mathbb{Z}^+$ *with* $m > 1$. *For each* $r \in [m]$, *suppose* $Q_r := (G_{r,1}, \ldots, G_{r,k})$ *is a* k*-tuple of disjoint subsets of* X. *Finally, let* $\alpha, \gamma \in X$. *Then the following hold:*

(a) *Let A_1 be Ann's response to Q_1 in the game $\mathcal{G}(k,m,X)$. Then A_1 is a finite subset of $\mathbb{Z}^+$ containing α. Thus the game $\mathcal{G}(k,m-1,A_1)$ is well-defined.*

(b) *$(\alpha,Q_1,\ldots,Q_m,\gamma)$ is an allowable game vector of the game $\mathcal{G}(k,m,X)$ if and only if $(\alpha,Q_2,\ldots,Q_m,\gamma)$ is an allowable game vector of the game $\mathcal{G}(k,m-1,A_1)$.*

(c) *$(\alpha,Q_1,\ldots,Q_m,\gamma)$ is a winning game vector of the game $\mathcal{G}(k,m,X)$ if and only if $(\alpha,Q_2,\ldots,Q_m,\gamma)$ is a winning game vector of the game $\mathcal{G}(k,m-1,A_1)$.*

Proof. Assume that X is a finite, nonempty subset of $\mathbb{Z}^+$, that $m,k \in \mathbb{Z}^+$ with $m>1$, and that $\alpha,\gamma \in X$. Further assume that for each $r \in [m]$, $Q_r := (G_{r,1},\ldots,G_{r,k})$ is a k-tuple of subsets of X. Finally, assume that $\alpha \in X$ is Ann's choice, and that $\gamma \in X$ also.

(a) By definition, $\alpha \in A_1 \subseteq X$, so A_1 is a nonempty subset of $\mathbb{Z}^+$. Also, since $m>1$, $m-1 \in \mathbb{Z}^+$. Therefore, the game $\mathcal{G}(k,m-1,A_1)$ is well-defined.

(b) Assume that $(\alpha,Q_1,\ldots,Q_m,\gamma)$ is an allowable game vector of the game $\mathcal{G}(k,m,X)$. Then by Proposition 4.4, we ascertain an option vector $(a_1,\ldots,a_m) \in [k]^m$ such that:

(i) $\bigsqcup Q_1 = X$;

(ii) $\bigsqcup Q_r = G_{r-1,a_{r-1}} := A_{r-1} \ni \alpha$ for $1<r\leq m$;

(iii) $\alpha,\gamma \in G_{m,a_m} := A_m$.

Now consider the game vector $(\alpha,Q_2,\ldots,Q_m,\gamma)$ in the game $\mathcal{G}(k,m-1,A_1)$. By (ii) we have $\bigsqcup Q_2 = A_1 \ni \alpha$, and by (ii)–(iii) we have that $(a_2,\ldots,a_m) \in [k]^{m-1}$ is an option vector satisfying the last two requirements of Proposition 4.4. Thus, $(\alpha,Q_2,\ldots,Q_m,\gamma)$ is an allowable game vector of $\mathcal{G}(k,m-1,A_1)$.

Conversely, assume that $(\alpha,Q_2,\ldots,Q_m,\gamma)$ is an allowable game vector of $\mathcal{G}(k,m-1,A_1)$. The definition of A_1 (Definition 4.1) implies $\bigsqcup Q_1 = X$ and that there exists $a_1 \in [k]$ with $A_1 = G_{1,a_1}$. Moreover, $(\alpha,Q_2,\ldots,Q_m,\gamma)$ allowable means that there exists an option vector $(a_2,\ldots,a_m) \in [k]^{m-1}$ with:

(iv) $\bigsqcup Q_2 = A_1 \ni \alpha$;

(v) $\bigsqcup Q_r = G_{r-1,a_{r-1}} := A_{r-1} \ni \alpha$ for $2<r\leq m$;

(vi) $\alpha,\gamma \in G_{m,a_m} := A_m$.

Conditions (iv)–(vi) together with $\bigsqcup Q_1 = X$ and $A_1 = G_{1,a_1}$ imply $(a_1,a_2,\ldots,a_m) \in [k]^m$ is the desired option vector of Proposition 4.4. So $(\alpha,Q_1,Q_2,\ldots,Q_m,\gamma)$ is an allowable game vector of $\mathcal{G}(k,m,X)$.

(c) This follows immediately from (b). □

At long last, our main result draws nigh! The proof will be by induction on m, the number of questions allotted to Gus. Since the base case is of independent interest, and provides the main idea for the general case, we present it first as a separate result. Recall that by Theorem 3.3, the probability that Gus wins the game $\mathcal{G}(k,m,X)$ cannot exceed $\frac{\min(|X|,k^m)}{|X|}$. Moreover, as we are about to state probabilistic results it is necessary that our language become even more precise. Recall that we are imagining Ann's choice α and

Gus's guess γ as both being uniformly at random on the respective sets X and $A_m \subseteq X$ in the game $\mathcal{G}(k,m,X)$. Thus, henceforth we shall also refer to *random* allowable game vectors $\mathbf{g} = \mathbf{g}(\alpha,\gamma) = (\alpha, Q_1, \ldots, Q_m, \gamma)$ in the game $\mathcal{G}(k,m,X)$. By this, we simply mean that $\mathbf{g} = \mathbf{g}(\alpha,\gamma)$ is an allowable game vector, and that both Ann's choice α and Gus's ultimate guess γ are *random* in the sense just described. Note well that once α has been randomly selected by Ann, the intermediate portion of the game where Gus presents his questions $(Q_1, \ldots, Q_m)$ to Ann is *non-random* in the sense that Gus constructs his questions based solely on the current information he has from Ann. Of course, Gus may choose different *ways* to construct his questions, but we do not assume the construction itself to be random. Gus's guess γ is a random concluding feature of the game, and so α and γ are the *only random* features of the game $\mathcal{G}(k,m,X)$.

Proposition 4.7. *Let X be a finite, nonempty subset of $\mathbb{Z}^+$, and let $k \in \mathbb{Z}^+$. Now let $\mathbf{g} = \mathbf{g}(\alpha,\gamma) := (\alpha, Q_1, \gamma)$ be a random allowable game vector of the game $\mathcal{G}(k,1,X)$, with $Q_1 := (G_{1,1}, \ldots, G_{1,k})$ satisfying $\bigsqcup Q_1 = X$. Finally, let $W_\mathbf{g}$ be the event, "$\mathbf{g}$ is a winning vector of the game $\mathcal{G}(k,1,X)$." Then*

(a) $\mathbb{P}(W_\mathbf{g}) = \frac{\min(|X|,k)}{|X|}$ *if and only if*

(b) $\min(1, |X| - k + 1) \le |G_{1,i}| \le \max(1, |X| - k + 1)$ *for each $i \in [k]$.*

Proof. To simplify matters, we introduce $g_i := |G_{1,i}|$ for each $i \in [k]$, and of course we have $|X| = n \in \mathbb{Z}^+$. So we have $\frac{\min(|X|,k)}{|X|} = \frac{\min(n,k)}{n}$ and

$$g_1 + g_2 + \cdots + g_k = n. \tag{5}$$

This last equation is interesting, because we can imagine the n integers in X as "balls," and the variables $g_1, \ldots, g_k$ as occupancy numbers for k (labeled) "bins" into which the n balls will be placed.

This framework gives us a very natural way to look at the proof. First, it's important to know if we have more "bins" than "balls," which motivates our choice of cases. Then we can look at how many of these "bins" are nonempty (we call this $|I|$ below), which translates directly into knowledge about the probability we are concerned about. Thus, we consider the following two cases.

Case 1: $n < k$**.** Let's first show that (b) implies (a). In this case (b) is equivalent to the requirement that $g_i \in \{0,1\}$ for each $i \in [k]$. So if Ann returns $A_1 = G_{1,a_1}$, then $|A_1| = 1$. Thus the random game vector $\mathbf{g} = (\alpha, Q_1, \gamma)$ is certain to be winning, since $\alpha, \gamma \in A_1$, a set of cardinality 1. That is, $\mathbb{P}(W_\mathbf{g}) = 1 = \frac{n}{n} = \frac{\min(n,k)}{n}$, and the reverse implication is proved.

We prove the forward implication by contraposition. The only way that (b) is violated in this case is if there is some $i_1 \in [k]$ with $g_{i_1} > 1$. Next, introduce the set

$$I := \{i \in [k] : g_i > 0\}. \tag{6}$$

Then for $i \notin I$ we have $g_i = 0$, i.e. $G_{1,i} = \varnothing$. Thus if $i \notin I$ we cannot have $A_1 = G_{1,i} = \varnothing$ since $\alpha \in A_1$, and so for $i \notin I$ we have $\mathbb{P}(W_{\mathbf{g}} \cap \{A_1 = G_{1,i}\}) = 0$. Therefore

$$\begin{aligned}\mathbb{P}(W_{\mathbf{g}}) &= \sum_{i\in[k]} \mathbb{P}(W_{\mathbf{g}} \cap \{A_1 = G_{1,i}\}) = \sum_{i\in I} \mathbb{P}(W_{\mathbf{g}} \cap \{A_1 = G_{1,i}\}) \\ &= \sum_{i\in I} \mathbb{P}(A_1 = G_{1,i}) \cdot \mathbb{P}(W_{\mathbf{g}} | A_1 = G_{1,i}) = \sum_{i\in I} \frac{g_i}{n} \cdot \frac{1}{g_i} \\ &= \frac{|I|}{n},\end{aligned} \tag{7}$$

so that the probability Gus wins is completely determined by $|I|$. Clearly $i_1 \in I$ and since $g_{i_1} > 1$ equation (5) says that box number i_1 is occupied by multiple balls. Since there are a total of n balls to distribute, this means that strictly less than n bins will contain a ball, which is equivalent to the assertion that $|I| < n$ since $|I|$ counts the number of occupied bins. So, in summary $1 \le |I| < n < k$, and thus from (7) we get

$$\mathbb{P}(W_{\mathbf{g}}) = \frac{|I|}{n} < \frac{n}{n} = \frac{\min(n,k)}{n},$$

and the forward implication is proved. This completes the proof in the case $n < k$.

Case 2: $n \ge k$. In this case (b) is equivalent to $1 \le g_i \le n-k+1$ for each $i \in [k]$. We first show that (b) implies (a). With the set I defined as in (6), the same argument given in (7) shows that $\mathbb{P}(W_{\mathbf{g}}) = \frac{|I|}{n}$. But here we have $I = [k]$, and consequently $\mathbb{P}(W_{\mathbf{g}}) = \frac{k}{n} = \frac{\min(n,k)}{n}$.

To prove the forward implication, we once again proceed by contraposition. Here, violation of (b) means that there exists $i_1 \in [k]$ such that either $g_{i_1} = 0$ or $g_{i_1} > n-k+1$. Let's start by assuming that $g_{i_1} = 0$. This means that $i_1 \notin I$, and so $|I| < k$. Thus

$$\mathbb{P}(W_{\mathbf{g}}) = \frac{|I|}{n} < \frac{k}{n} = \frac{\min(n,k)}{n}. \tag{8}$$

It only remains to consider the case where $g_{i_1} > n-k+1$. Recall that $|I|$ is the number of $i \in [k]$ with $g_i > 0$, and so $|I| \le k$ always. If it happened that $|I| = k$ here, then we would obtain

$$g_1 + \cdots + g_{i_1} + \cdots + g_k > (k-1)\cdot 1 + (n-k+1) = n,$$

which violates (5). Contradiction! Thus, we have $|I| < k$ in this case also, and so the argument in (8) may be repeated. This completes the proof in the case $n \ge k$, and so the proposition is proved. $\square$

Let us pause to examine what we've shown, because it is quite surprising. Suppose we are in the game $\mathcal{G}(4,1,8)$. Proposition 4.7 says that Gus maximizes his probability of winning (which is $\frac{4^1}{8} = \frac{1}{2}$ by Theorem 3.3) if and only if his question $Q_1 = (G_{1,1}, \ldots, G_{1,4})$ satisfies $1 \le |G_{1,i}| \le 5$, $i \in [4]$. Of course, we must also have $|G_{1,1}| + \cdots + |G_{1,4}| = 8$ by Assumption 2, but this affords Gus a *huge* degree of freedom in constructing his question. In fact, we encourage the interested reader to show, using the principle of inclusion-exclusion to count the number of ways to allocate 8 distinguishable "balls" into 4 labeled "bins" in accordance with these restrictions, that Gus has

$$4^8 - \binom{4}{1}3^8 + \binom{4}{2}2^8 - \binom{4}{3}1^8 = 40824 \tag{9}$$

ways to maximize his probability of winning! Here, $\binom{a}{b} := \frac{a!}{b!(a-b)!}$ is the binomial coefficient for integers $a \geq b \geq 0$. So the "intuitive" idea to divide the set $[8]$ into 4 equal parts, so that $|G_{1,1}| = \cdots = |G_{1,4}| = 2$, accounts for just $\frac{8!}{(2!)^4} = 2520$ of the 40824 total ways to optimize Gus's strategy. Curiously, even if Gus chose a *random* allocation of these 8 "balls" into the 4 "bins" from among all $4^8 = 65536$ possible distributions, he would *still* have a $\frac{40824}{65536} = 62.29\%$ chance of selecting an optimal strategy!

We are just getting started. Now, for the main attraction!

Theorem 4.8. *Let X be a finite, nonempty subset of $\mathbb{Z}^+$, and let $k, m \in \mathbb{Z}^+$. Now let $\mathbf{g} = \mathbf{g}(\alpha, \gamma) := (\alpha, Q_1, \ldots, Q_m, \gamma)$ be a random allowable game vector of the game $\mathcal{G}(k, m, X)$, with $Q_r := (G_{r,1}, \ldots, G_{r,k})$ satisfying $\bigsqcup Q_r = A_{r-1}$ for each $r \in [m]$. We remind the reader that $A_0 := X$ and for each $r \in [m]$, A_r denotes Ann's response to the question Q_r presented by Gus. Finally, let $W_{\mathbf{g}}(k, m, X)$ be the event, "$\mathbf{g}$ is a winning vector of the game $\mathcal{G}(k, m, X)$." Then*

(a) $\mathbb{P}(W_{\mathbf{g}}(k, m, X)) = \frac{\min(|X|, k^m)}{|X|}$ *if and only if*

(b) $\min(k^{m-r}, |A_{r-1}| - (k-1)k^{m-r}) \leq |G_{r,i}| \leq \max(k^{m-r}, |A_{r-1}| - (k-1)k^{m-r})$ *for all $i \in [k]$ and $r \in [m]$.*

Proof. We proceed by induction on m. Thus suppose the theorem is true for all $\ell < m$. If $m = 1$, then we are done by Proposition 4.7. Also, the theorem is obvious for $k = 1$, since in this case we have $G_{1,1} = G_{2,1} = \cdots = G_{m,1} = X$ and $A_0 = A_1 = \cdots = A_m = X$, and so condition (a) reduces to $\mathbb{P}(W_{\mathbf{g}}(1, m, X)) = \frac{1}{|X|}$ and (b) to $1 \leq |G_{r,1}| \leq |X|$ for all $r \in [m]$. Therefore, we may suppose that $k > 1$ and $m > 1$.

Now set

$$\mathbf{g}' := (\alpha, Q_2, \ldots, Q_m, \gamma),$$

which is a random allowable game vector of the game $\mathcal{G}(k, m-1, A_1)$ by Proposition 4.6. As in the proof of Proposition 4.7, we introduce $g_{r,i} := |G_{r,i}|$ for all $i \in [k]$ and $r \in [m]$, and write $|X| = n \in \mathbb{Z}^+$. So by (1) of Definition 4.1 we have

$$g_{r,1} + g_{r,2} + \cdots + g_{r,k} = |A_{r-1}|, \qquad r \in [m], \tag{10}$$

and condition (b) of Theorem 4.8 becomes

$$\begin{gathered}\min(k^{m-r}, |A_{r-1}| - (k-1)k^{m-r}) \leq g_{r,i} \leq \max(k^{m-r}, |A_{r-1}| - (k-1)k^{m-r}) \\ \text{for all } i \in [k] \text{ and } r \in [m].\end{gathered} \tag{11}$$

When it is convenient for our purposes, we shall once again invoke the "balls-in-bins" interpretation of (10). Finally, also analogous to the proof of Proposition 4.7, for each $r \in [m]$ we define

$$I_r := \{i \in [k] : g_{r,i} > 0\}. \tag{12}$$

Clearly, by (10) we have $I_r \neq \emptyset$ for each $r \in [m]$. Once again, as in the proof of Proposition 4.7, observe that

$$\begin{aligned} \mathbb{P}(W_{\mathbf{g}}(k,m,X)) &= \sum_{i\in I_1} \mathbb{P}(A_1 = G_{1,i}) \cdot \mathbb{P}(W_{\mathbf{g}}(k,m,X) \,|\, A_1 = G_{1,i}) && (13)\\ &= \sum_{i\in I_1} \frac{g_{1,i}}{n} \cdot \mathbb{P}(W_{\mathbf{g}'}(k,m-1,G_{1,i})) && (14)\\ &\leq \sum_{i\in I_1} \frac{g_{1,i}}{n} \cdot \frac{\min(g_{1,i},k^{m-1})}{g_{1,i}} = \sum_{i\in I_1} \frac{\min(g_{1,i},k^{m-1})}{n} && (15)\\ &\leq \frac{\min(n,k^m)}{n}. && (16)\end{aligned}$$

Here, in the second equality (14) we have used Proposition 4.6, and in the first inequality (15) we have applied Theorem 3.3; the second inequality (16) is patent from (10) with $r = 1$. We are ready to establish the equivalence of (a) and (b). To avoid unnecessary complications, we consider two cases.

Case 1: $n \leq k^m$**.** Let's start by assuming (b), which we have recast as (11). Here we have $n \leq k^{m-1} + (k-1)k^{m-1}$, and so $n - (k-1)k^{m-1} \leq k^{m-1}$. Consequently, the $r = 1$ condition in (11) becomes

$$\max(n-(k-1)k^{m-1}, 0) \leq g_{1,i} \leq \min(n, k^{m-1}), \qquad i \in [k]; \tag{17}$$

here, the max/min on the left- and right-hand sides (resp.) come from the $r = 1$ "balls-in-bins" requirement in (10). By the inductive hypothesis, the $2 \leq r \leq m$ conditions in (11) guarantee that the inequality (15) is, in fact, an equality. Thus, to prove (a) we need only show that the second inequality (16) is also an equality, which is clearly equivalent to proving

$$\sum_{i\in[k]} g_{1,i} = \sum_{i\in I_1} g_{1,i} = n, \tag{18}$$

since $n \leq k^m$ in this case, and (17) holds. But (18) is just the $r = 1$ "balls-in-bins" requirement in (10), and so the reverse implication is proved.

Now let us assume (a), so that $\mathbb{P}(W_{\mathbf{g}}(k,m,X)) = \frac{\min(n,k^m)}{n} = \frac{n}{n} = 1$. This means that the inequalities (15) and (16) are equalities. Since (15) is now equality, we know that

$$\mathbb{P}(W_{\mathbf{g}'}(k,m-1,G_{1,i})) = \frac{\min(g_{1,i},k^{m-1})}{g_{1,i}}, \qquad i \in I_1 \neq \emptyset.$$

Thus, by the inductive hypothesis, (11) holds for $2 \leq r \leq m$. So it only remains to show that (17) holds. To this end, observe that since (16) is now equality, we have

$$\sum_{i\in[k]} \min(g_{1,i},k^{m-1}) = \sum_{i\in I_1} \min(g_{1,i},k^{m-1}) = n. \tag{19}$$

We claim that (19) implies that $g_{1,i} \leq k^{m-1}$ for each $i \in [k]$. Indeed, if we had, say, $g_{1,1} > k^{m-1}$, then we would have

$$\sum_{i\in[k]} g_{1,i} > k^{m-1} + \sum_{2\leq i\leq k} g_{1,i} \geq \sum_{i\in[k]} \min(g_{1,i},k^{m-1}) = n,$$

which contradicts (18). This, along with the fact that $g_{1,i} \leq n$, $i \in [k]$, by (18), shows that the right-hand inequality

$$g_{1,i} \leq \min(n, k^{m-1}), \qquad i \in [k], \tag{20}$$

in (17) holds.

To prove the left-hand inequality in (17), observe that we clearly have $g_{1,i} \geq 0$, $i \in [k]$. Thus, it only remains to prove that $g_{1,i} \geq n-(k-1)k^{m-1}$, $i \in [k]$. Indeed, if we had, say, $g_{1,1} < n-(k-1)k^{m-1}$, then since (20) holds we would have

$$g_{1,1} + g_{1,2} + \cdots + g_{1,k} < (n-(k-1)k^{m-1}) + (k-1)k^{m-1} = n.$$

This contradicts (18), and so the left-hand inequality in (17) holds. This concludes the proof in this case.

Case 2: $n > k^m$. Let us assume that (b) is true. Then $n > (k-1)k^{m-1} + k^{m-1}$, and the $r = 1$ condition in (11) becomes

$$k^{m-1} \leq g_{1,i} \leq n-(k-1)k^{m-1}, \qquad i \in [k]. \tag{21}$$

Again, we can conclude from our inductive hypothesis that the $2 \leq r \leq m$ conditions in (11) ensure equality for (15). Thus we need only show that the second inequality (16) is also equality, which is equivalent to showing that

$$\sum_{i \in I_1} k^{m-1} = k^m, \tag{22}$$

since $\min(g_{1,i}, k^{m-1}) = k^{m-1}$ for each $i \in [k]$ by (21). But this is the same as showing that $g_{1,i} > 0$, $i \in [k]$ (for then $I_1 = [k]$). However, since we assumed $k > 1$ and $m > 1$, the left-hand inequality in (21) guarantees that this is the case, and consequently (22) holds. Hence, the reverse implication is proved.

Now assume that (a) holds. As in the last case, the inequalities (15) and (16) become equalities. Since (15) is now equality, we know by the inductive hypothesis that (11) holds for $2 \leq r \leq m$. We will show that (21) holds as well. Notice that we have

$$\sum_{i \in [k]} \min(g_{1,i}, k^{m-1}) = \sum_{i \in I_1} \min(g_{1,i}, k^{m-1}) = k^m \tag{23}$$

since (16) is equality. If it happened that, say, $g_{1,1} < k^{m-1}$, we would have

$$\sum_{i \in [k]} \min(g_{1,i}, k^{m-1}) = g_{1,1} + \sum_{2 \leq i \leq k} \min(g_{1,i}, k^{m-1}) < \sum_{i \in [k]} k^{m-1} = k^m,$$

contradicting (23). We conclude that the left-hand inequality in (21) is proved, i.e.

$$g_{1,i} \geq k^{m-1}, \qquad i \in [k]. \tag{24}$$

To prove that the right-hand inequality in (21) holds, suppose that $g_{1,1} > n-(k-1)k^{m-1}$. Then, invoking (24) we obtain

$$g_{1,1} + g_{1,2} + \ldots + g_{1,k} > n-(k-1)k^{m-1} + (k-1)k^{m-1} = n,$$

which contradicts (18). Therefore, the right-hand inequality in (21) holds, and the proof is complete. □

Henceforth, we shall refer to Gus's probability-maximizing strategies detailed in Theorem 4.8 as the "optimal strategies." Moreover, in light of condition (b) of Theorem 4.8, we shall have occasion to reference the m vectors of cardinalities $(|G_{r,1}|,\ldots,|G_{r,k}|)$, $r \in [m]$, in the game $\mathcal{G}(k,m,X)$ and will refer to these collectively as Gus's "cardinality vectors."

5. Some Technicalities and Consequences

In this section, we take care of some technical book-keeping items, and present some immediate consequences of our results that seem especially pertinent. In Theorem 4.8 we have determined, precisely, Gus's available optimal strategies. But there is still the question of whether Gus is *always able* to satisfy the cardinality conditions (b) of Theorem 4.8 in order to achieve this optimal outcome. We dispense with this pesky technicality now, although it will take some work!

Proposition 5.1. *There is an allowable game vector* $(\alpha, Q_1, \ldots, Q_m, \gamma)$ *in the game* $\mathcal{G}(k,m,X)$ *that satisfies* (b) *of Theorem 4.8.*

Proof. Fix $r \in [m]$, and suppose that for each $j < r$ Gus has chosen $Q_j := (G_{j,1},\ldots,G_{j,k})$, then presented Q_j to Ann, and she has given her response $A_j := G_{j,a_j}$ to him. We must show that Gus can now construct his next question $Q_r = (G_{r,1},\ldots,G_{r,k})$ to satisfy (b) of Theorem 4.8. Once again writing $g_{r,i} := |G_{r,i}|$ for $i \in [k]$, Gus needs only select a solution to the "balls-in-bins" equation

$$g_{r,1} + g_{r,2} + \cdots + g_{r,k} = |A_{r-1}| \tag{25}$$

satisfying (b) of Theorem 4.8, and consequently our task becomes to show that such a solution exists. There are two cases to consider.

Case 1: $|A_{r-1}| \le k^{m-r+1}$**.** Then $|A_{r-1}| - (k-1)k^{m-r} \le k^{m-r}$, and condition (b) from Theorem 4.8 in tandem with (25) require that

$$\max(|A_{r-1}| - (k-1)k^{m-r}, 0) \le g_{r,i} \le \min(|A_{r-1}|, k^{m-r}), \qquad i \in [k];$$

this ought to remind the reader of (17) in the proof of Theorem 4.8, as it is a generalization. Under this restriction we have

$$\sum_{i\in[k]} g_{r,i} \in [k \cdot \max(|A_{r-1}| - (k-1)k^{m-r}, 0), k \cdot \min(|A_{r-1}|, k^{m-r})], \tag{26}$$

with all sums in this range clearly possible. It follows that (25) has a solution if and only if

$$|A_{r-1}| \in [k \cdot \max(|A_{r-1}| - (k-1)k^{m-r}, 0), k \cdot \min(|A_{r-1}|, k^{m-r})], \tag{27}$$

and so we need only prove (27). First, clearly $|A_{r-1}| \le k \cdot |A_{r-1}|$ and also $|A_{r-1}| \le k^{m-r+1}$ by assumption in this case. Thus $|A_{r-1}| \le k \cdot \min(|A_{r-1}|, k^{m-r})$, so it remains to show that $|A_{r-1}| \ge k \cdot \max(|A_{r-1}| - (k-1)k^{m-r}, 0)$. This inequality is obvious if the maximum on the right-hand side is 0, so let's suppose that $|A_{r-1}| - (k-1)k^{m-r} > 0$. Then we need to show

$$|A_{r-1}| \ge k(|A_{r-1}| - (k-1)k^{m-r}). \tag{28}$$

If instead we had $|A_{r-1}| < k(|A_{r-1}| - (k-1)k^{m-r})$, we would obtain

$$(k-1)k^{m-r+1} < (k-1)|A_{r-1}|,$$

from which $k > 1$ and then $|A_{r-1}| > k^{m-r+1}$ follow. This contradicts our Case 1 assumption, and therefore (28) (and hence (27)) is proved. This completes the proof in this case.

Case 2: $|A_{r-1}| > k^{m-r+1}$**.** Then $|A_{r-1}|-(k-1)k^{m-r} > k^{m-r}$, and so following the same template as last case the analog of equation (26) becomes

$$\sum_{i\in[k]} g_{r,i} \in \left[k^{m-r+1}, k(|A_{r-1}|-(k-1)k^{m-r})\right]; \tag{29}$$

no "max" or "min" are necessary in this case. Proceeding as in Case 1, we must show

$$|A_{r-1}| \in \left[k^{m-r+1}, k(|A_{r-1}|-(k-1)k^{m-r})\right]. \tag{30}$$

But clearly $|A_{r-1}| \geq k^{m-r+1}$ by our assumption in this case. And if we had

$$|A_{r-1}| > k(|A_{r-1}|-(k-1)k^{m-r})$$

then as above we would obtain $(k-1)k^{m-r+1} > (k-1)|A_{r-1}|$, from which $k > 1$ and then $k^{m-r+1} > |A_{r-1}|$ follow, which violates our Case 2 assumption. Consequently, we must have $|A_{r-1}| \leq k(|A_{r-1}|-(k-1)k^{m-r})$, and (30) follows. The proof is complete. ☐

Now for some low-hanging fruit! Next, we determine precisely when it is that Gus has a winning strategy.

Corollary 5.2. *Gus has a winning strategy in the game $\mathcal{G}(k,m,X)$ if and only if $m \geq \log_k |X|$.*

Proof. Gus has a winning strategy if and only if $\mathbb{P}(W_{\mathbf{g}}(k,m,X)) = \frac{\min(|X|,k^m)}{|X|} = 1$ if and only if $|X| \leq k^m$ if and only if $m \geq \log_k |X|$. ☐

Another technicality we may dispose of is the "min" and "max" in condition (b) of Theorem 4.8. This makes for a much cleaner-looking statement of our result there.

Corollary 5.3. *Let $\mathbf{g} = \mathbf{g}(\alpha,\gamma) := (\alpha, Q_1, \ldots, Q_m, \gamma)$ be a random allowable game vector in the game $\mathcal{G}(k,m,X)$. Then the following hold:*

(a) *If $k^m \geq |X|$, then $\mathbb{P}(W_{\mathbf{g}}(k,m,X)) = 1$ if and only if*
$|A_{r-1}|-(k-1)k^{m-r} \leq |G_{r,i}| \leq k^{m-r}$ *for all $i \in [k]$ and $r \in [m]$.*

(b) *If $k^m \leq |X|$, then $\mathbb{P}(W_{\mathbf{g}}(k,m,X)) = \frac{k^m}{|X|}$ if and only if*
$k^{m-r} \leq |G_{r,i}| \leq |A_{r-1}|-(k-1)k^{m-r}$ *for all $i \in [k]$ and $r \in [m]$.*

Proof. Once again, let $|X| = n$ and $g_{r,i} := |G_{r,i}|$ for $i \in [k]$, $r \in [m]$. We proceed by induction on m. Hence, assume the corollary holds for all $\ell < m$. If $m = 1$, Theorem 4.8 guarantees (a) and (b) hold, since under the hypothesis of (a) we have $n-(k-1)k^{m-1} \leq k^{m-1}$, and under (b) we have $k^{m-1} \leq n-(k-1)k^{m-1}$. So assume that $m > 1$. In addition, by Theorem 4.8 we only need to prove the forward implications.

We prove (a), since the proof of (b) is similar. Suppose $k^m \geq n$ and $\mathbb{P}(W_{\mathbf{g}}(k,m,X)) = 1$. Then $n-(k-1)k^{m-1} \leq k^{m-1}$, and by Theorem 4.8 we know that $n-(k-1)k^{m-1} \leq g_{1,i} \leq k^{m-1}$ for all $i \in [k]$. Now, for some $a_1 \in [k]$ we have $\alpha \in A_1 = G_{1,a_1}$ and hence $|A_1| = g_{1,a_1} \leq k^{m-1}$.

This and the inductive hypothesis applied to the game $\mathcal{G}(k, m-1, A_1)$ with random game vector $\mathbf{g}' := (\alpha, Q_2, \ldots, Q_m, \gamma)$ now imply that

$$|A_{r-1}| - (k-1)k^{m-r} \leq g_{r,i} \leq k^{m-r} \text{ for all } i \in [k] \text{ and } 2 \leq r \leq m.$$

□

By Corollary 5.2, Gus has a winning strategy in the game $\mathcal{G}(k, m, X)$ if and only if we have $m \geq \lceil \log_k |X| \rceil$. But even in the extreme case $m = \lceil \log_k |X| \rceil$, Gus may still have a good deal of freedom in choosing his cardinality vectors for a winning strategy. Indeed, consider the specific case of $m = \lceil \log_4 8 \rceil = 2$ in the game $\mathcal{G}(4, 2, 8)$. By (a) of Corollary 5.3, Gus can begin a winning strategy in the game by constructing his first cardinality vector $(|G_{1,1}|, \ldots, |G_{1,4}|)$ with coordinates in the interval $[0, 4]$. Using once again the "balls-in-bins" context, and subtracting away those ball placements with 5, 6, 7, or 8 in a box (there can only be one such "bad" box), we see that there are a total of

$$4^8 - 4\binom{8}{5}3^3 - 4\binom{8}{6}3^2 - 4\binom{8}{7}3^1 - 4\binom{8}{8}3^0 = 58380$$

ways to start that satisfy this requirement! Of course, here we are accounting for all the different ways that Gus could construct a given cardinality vector. Since there are a total of $4^8 = 65536$ ways that Gus could begin the game, this says that $\frac{58380}{65536} = 89.08\%$ of possible starts to the game are actually optimal! Moreover, Ann's response A_1 is certain to satisfy $|A_1| \leq 4$. And since Gus still has one more 4-option question left, he is certain to win so long as he follows the second cardinality vector condition in (a) of Corollary 5.3, namely that $(|G_{2,1}|, \ldots, |G_{2,4}|)$ have all coordinates belonging to $\{0, 1\}$.

Lastly, we determine exactly when it is that the optimal strategy cardinality vectors in the game $\mathcal{G}(k, m, X)$ are all uniquely determined. Perhaps unsurprisingly, the answer connects us back to our initial "intuitive" strategy for playing the game.

Corollary 5.4. *Let $\mathbf{g} = \mathbf{g}(\alpha, \gamma) := (\alpha, Q_1, \ldots, Q_m, \gamma)$ be a random allowable game vector in the game $\mathcal{G}(k, m, X)$, with maximum probability $\mathbb{P}(W_{\mathbf{g}}(\mathcal{G}(k, m, X))) = \frac{\min(|X|, k^m)}{|X|}$. Then*

(a) *the cardinality vectors $(|G_{r,1}|, |G_{r,2}|, \ldots, |G_{r,k}|)$, $r \in [m]$, are uniquely determined if and only if*

(b) $|X| = k^m$.

Proof. We let $|X| = n$ and $g_{r,i} := |G_{r,i}|$ for $i \in [k]$ and $r \in [m]$. Assume (a). By Theorem 4.8 we know that $\min(k^{m-1}, n-(k-1)k^{m-1}) \leq g_{1,1} \leq \max(k^{m-1}, n-(k-1)k^{m-1})$, and since $g_{1,1}$ is uniquely determined we get $g_{1,1} = \min(k^{m-1}, n-(k-1)k^{m-1}) = \max(k^{m-1}, n-(k-1)k^{m-1})$. But this can only occur when $k^{m-1} = n-(k-1)k^{m-1}$, so $k^m = n$.

Conversely, assume (b). By Corollary 5.3, we get that $g_{r,i} = k^{m-r}$ for all $i \in [k]$ and $r \in [m]$. □

6. An Extension

We close with an interesting extension. Let $\mathbf{k} := (k_1, \ldots, k_m) \in (\mathbb{Z}^+)^m$, define $\|\mathbf{k}\| := k_1 k_2 \cdots k_m$, and for each $r \in [m]$ introduce $\|\mathbf{k}\|_{m-r} := \frac{\|\mathbf{k}\|}{k_1 k_2 \cdots k_r}$. For instance, we have

$$\|\mathbf{k}\|_{m-1} = k_2 k_3 \cdots k_m, \quad \|\mathbf{k}\|_{m-2} = k_3 k_4 \cdots k_m \quad \text{and} \quad \|\mathbf{k}\|_0 = \|\mathbf{k}\|_{m-m} = 1.$$

Note that $\|\mathbf{k}\|_{m-r}$ consists of the last $m-r$ coordinates of $\mathbf{k}$, multiplied together. A natural extension of the canonical game $\mathcal{G}(k, m, X)$ is the following.

Definition 6.1 (The game $\mathcal{G}(\mathbf{k}, m, X)$, canonical version)**.** A finite, nonempty set $X \subseteq \mathbb{Z}^+$ is presented to Ann and Gus. Ann randomly chooses a number $\alpha \in X$ known only to her. Further, Gus is given a positive integer $m \in \mathbb{Z}^+$ and a vector of m positive integers $\mathbf{k} := (k_1, \ldots, k_m) \in (\mathbb{Z}^+)^m$.

For each $r \in [m]$, Gus will present to Ann an ordered k_r-tuple of disjoint subsets of X, which we denote by $Q_r := (G_{r,1}, G_{r,2}, \ldots, G_{r,k_r})$. Ann then responds by returning $A_r := G_{r,a_r}$ if and only if $a_r \in [k_r]$ is such that $\alpha \in G_{r,a_r}$. To be clear, this all transpires sequentially in the sense that Gus first presents Q_1 to Ann, to which she responds with A_1, at which point Gus then presents Q_2 to Ann, to which she responds with A_2, and so on. After Gus has presented all m "questions" $(Q_1, \ldots, Q_m)$ to Ann, and subsequently received all m responses $(A_1, \ldots, A_m)$ from Ann, Gus then attempts to guess the number Ann chose. Now set $A_0 := X$. We further require the following:

(1) Gus's question Q_r satisfies $\bigsqcup Q_r := \bigsqcup_{1 \le i \le k_r} G_{r,i} = A_{r-1}$ for all $r \in [m]$, and

(2) Gus's guess γ is a random member of A_m.

Gus wins if his guess $\gamma = \alpha$.

In essence, Definition 6.1 only differs from our original canonical game notion in that it permits the number of options per question to vary. Naturally, we have the following extension of Theorem 3.3.

Theorem 6.2. *The probability that Gus wins the game $\mathcal{G}(\mathbf{k}, m, X)$ is at most $\frac{\min(|X|, \|\mathbf{k}\|)}{|X|}$.*

And subsequently, we obtain a generalization of Theorem 4.8.

Theorem 6.3. *Let X be a finite, nonempty subset of $\mathbb{Z}^+$, let $m \in \mathbb{Z}^+$, and we introduce $\mathbf{k} := (k_1, \ldots, k_m) \in (\mathbb{Z}^+)^m$. Now let $\mathbf{g} = \mathbf{g}(\alpha, \gamma) := (\alpha, Q_1, \ldots, Q_m, \gamma)$ be a random allowable game vector of the game $\mathcal{G}(\mathbf{k}, m, X)$, with $Q_r := (G_{r,1}, \ldots, G_{r,k_r})$ satisfying $\bigsqcup Q_r = A_{r-1}$ for each $r \in [m]$. Finally, let $W_{\mathbf{g}}(\mathbf{k}, m, X)$ be the event, "$\mathbf{g}$ is a winning vector of the game $\mathcal{G}(\mathbf{k}, m, X)$." Then*

(a) $\mathbb{P}(W_{\mathbf{g}}(\mathbf{k}, m, X)) = \frac{\min(|X|, \|\mathbf{k}\|)}{|X|}$ *if and only if*

(b) $\min(\|\mathbf{k}\|_{m-r}, |A_{r-1}| - (k_r - 1)\|\mathbf{k}\|_{m-r}) \le |G_{r,i}| \le \max(\|\mathbf{k}\|_{m-r}, |A_{r-1}| - (k_r - 1)\|\mathbf{k}\|_{m-r})$ *for all $i \in [k_r]$ and $r \in [m]$.*

Indeed, the proofs of Theorems 6.2–6.3 run parallel to those of Theorems 3.3–4.8 with very minor modifications. Of course, Proposition 4.7 still applies. And of course, if $\mathbf{k} = (k, k, \ldots, k)$ for some $k \in \mathbb{Z}^+$, we have $\|\mathbf{k}\|_{m-r} = k^{m-r}$ for each $r \in [m]$ and $\|\mathbf{k}\| = k^m$, and we obtain Theorems 3.3–4.8 from Theorems 6.2–6.3. Moreover, Proposition 5.1 and Corollaries 5.2–5.4 also have their analogs in this setting. We leave the details of these modifications, and all the proofs, to the interested reader.

7. Acknowledgments

We would like to thank Greg Oman for his encouragement, suggestions and helpful comments. Also, we would like to extend our gratitude to an anonymous referee for truly helpful feedback that much improved the paper's readability and presentation.

References

[1] P. Billingsley, *Probability and Measure*. Wiley, Hoboken, 2012.
[2] H. Enderton, *A mathematical introduction to logic*. Academic Press, New York - London, 1972.
[3] A. Hammett & G. Oman, *I'm thinking of a number...*, Missouri J. of Math. Sci. **28(1)** (2016).
[4] K. Rosen, *Discrete Mathematics and Its Applications, Seventh Edition*. McGraw-Hill, New York, 2012.

Student biographies

Ryan Cushman: (*Corresponding author:* ryan.cushman@wmich.edu) Ryan Cushman graduated from Bethel College (Mishawaka, IN) in 2016 with a B.S. in Mathematics. Currently, he is pursuing a Ph.D. in Mathematics at Western Michigan University with research interests in Combinatorics and Graph Theory.