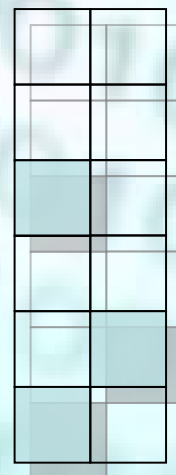




Ulam 的游戏和编码

东北育才学校 俞玮



Ulam 的游戏

- 有 n 个数字。
- 其中有一个与其他的不同，是游戏者所选定的。
- 你可以提出“这个数字在集合 X 中吗？”这样的问题。
- 你能收到的答案只有“是”或“否”。
- 可能会有一个错误的回答，但不多于一个。
- 用尽可能少的提问次数求出这个选定的数字。



原型题目及解法

- 同样是 n 个数字，选定一个。
- 问题的方式也是一样。
- 但回答中不包含错误。
- 解答：
 - 使用二分法。 $\log n$ 次提问。
 - 每次把可能为选定数字的那些数字分成尽可能相等的两个部分。
 - 以其中一个部分作为问题中的 x 。
 - 按照回答，去掉不可能为选定数字的数字。
 - 结果为最后剩下的一个数字。



二分法还是筛选法

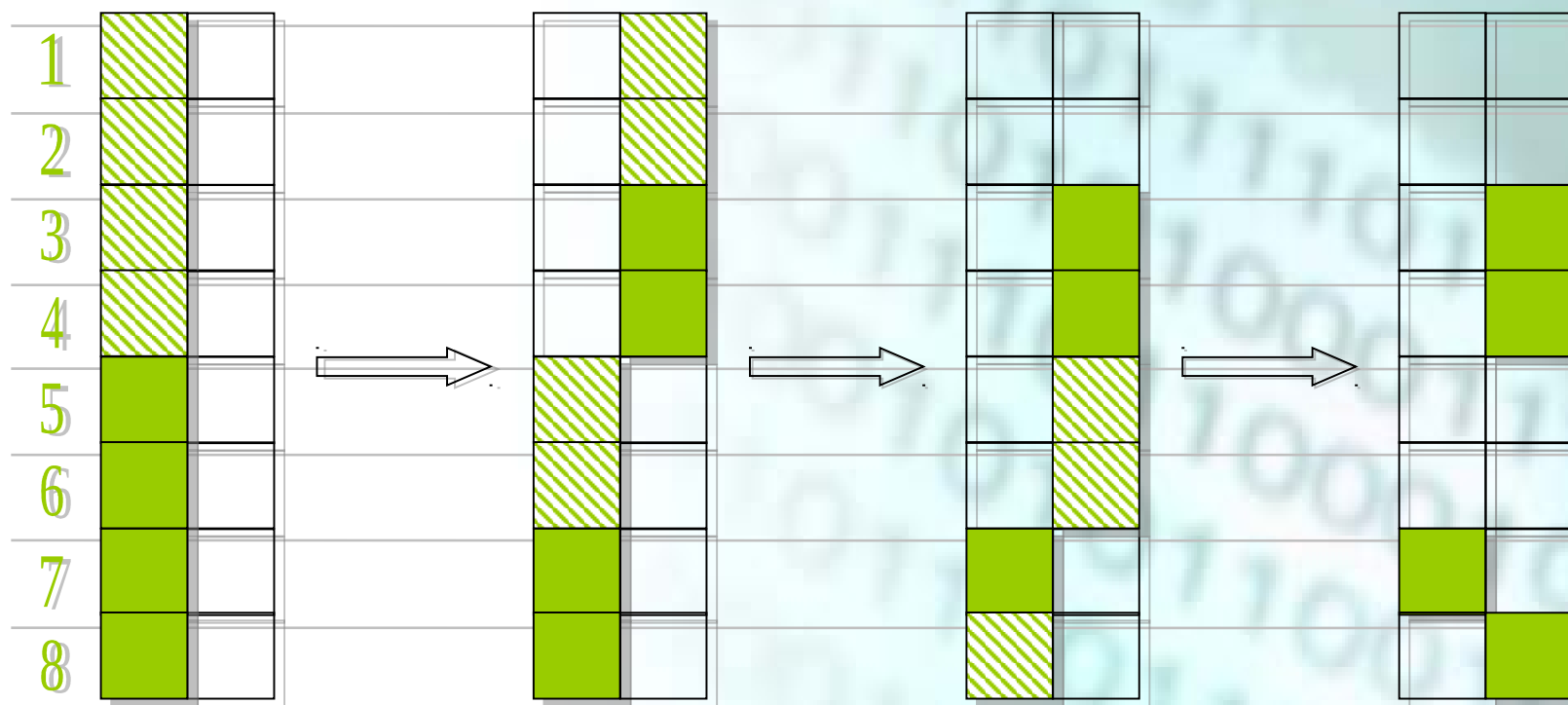


- 我们的二分法只是一种比较贪心的筛法。
- 让算法在最坏的情况下筛去的数字最多。

二分，按错误分类

- 原理：
 - 平分的不是数量。
 - 平均分的为选定数字的可能性。
 - 两个部分包含选定数字的可能性是相等的。
- 实现：
 - 对于每个数字，找出回答错误几次后，它才可能为选定数字。
 - 按照该次数，对数字进行分类。
 - 每一类中的数字可能为选定数字的可能性都是相等的。
 - 尽可能平均分每一类的数字，并各取一部分作为问题中的 X 。

图一：一个操作实例



- 图中的两列分别代表两类，所对应的列上非空白的部分属于该类。阴影部分被判定为不包含选定数字的一部分。

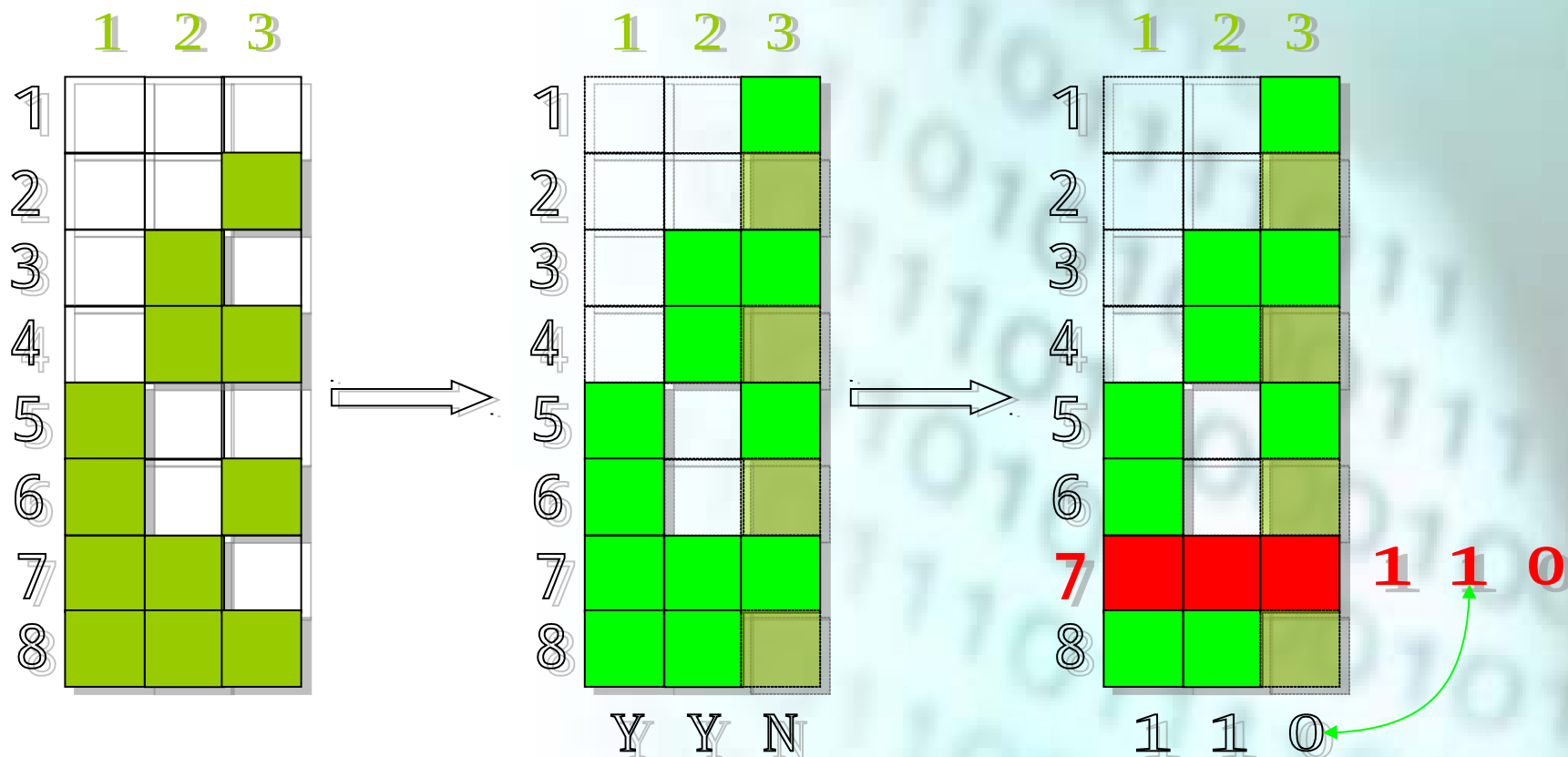
二分法的次数估计

- 第一类中的数字在 k 次提问后的数目：
 - $n/2^k$ 个方格。
- 第二类中的数字在 k 次提问后的数目：
 - 设为 $f(k)$ 。
 - 可知 $f(k)=f(k-1)/2+n/2^k$ 。
 - $f(k)=kn/2^k$ 。
- 当第一二类中的数字只有一个时，即确定了答案：
 - $f(k)+n/2^k \leq 1$ 。
 - $2^k \geq kn+n$ 。

新的方法—编码

- 筛选法的限制：
 - 结果的保存。
 - 对上一次结果的依赖。
 - 边界条件的处理。
- 新的编码方法：
 - 为每一数字赋一确定的二进制编码。
 - 第 i 个问题 \Leftrightarrow 编码第 i 位为 1。
 - 结果编码： $Y \Leftrightarrow 1$, $N \Leftrightarrow 0$ 。
 - 答案： 编码与结果编码一样的数字。

图二：编码的例子

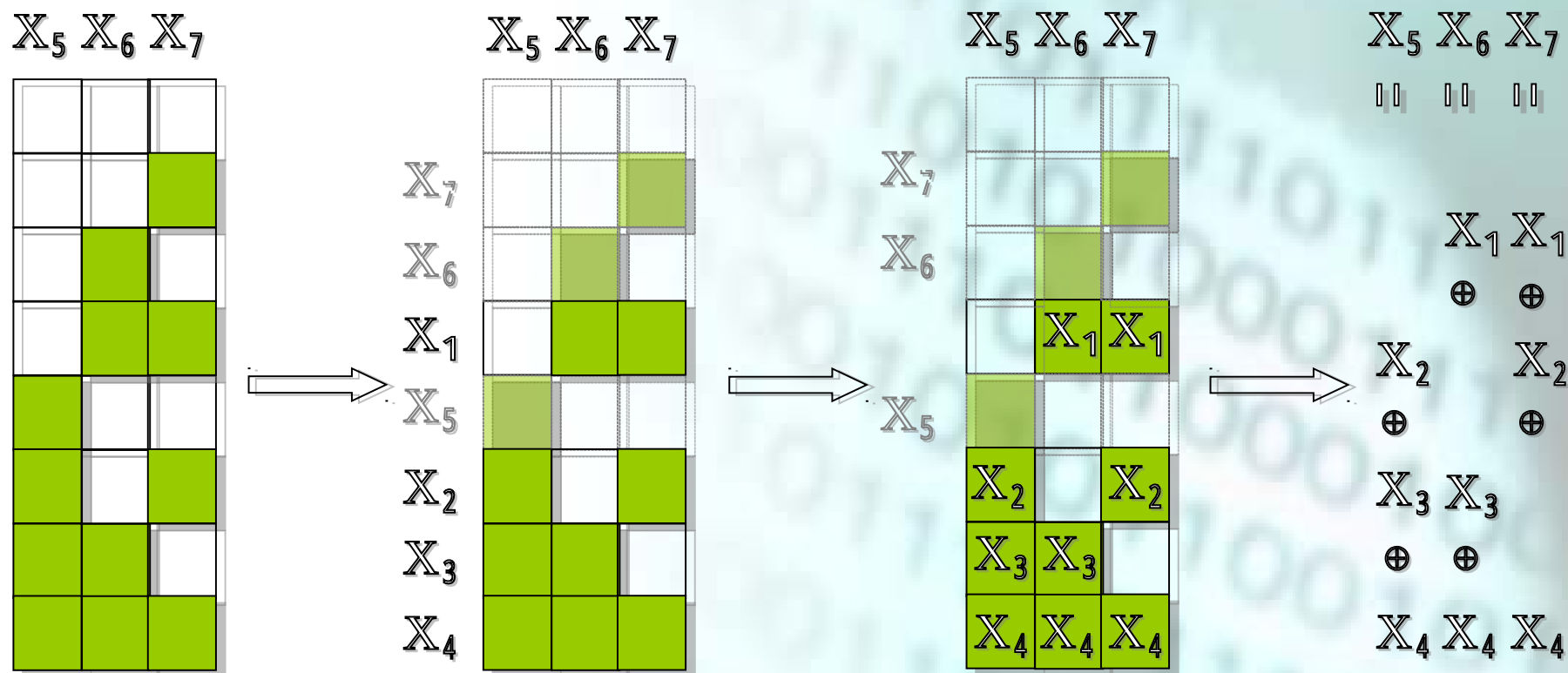


- 对于每次的回答，我们在相应的列中把对应的答案都标记为高亮。
- 一行均为高亮的数字 ⇔ 和结果编码一样的数字。

错误与纠错码

- 错误:
 - m 次回答中有一个错误 $\Leftrightarrow n$ 位编码中有一个错误。
 - 对不同的回答编码 \Leftrightarrow 对不同的错误编码。
- 纠错码的设计:
 - 使用奇偶校验码，纠错码为其他编码的二进制加法和。
 - 一位编码的错误将会导致所有相关等式的不成立。
 - 假设非纠错码位是正确的，找出“错误”的纠错码位。
 - 如果没有，那么没有错误。
 - 如果只有一个，错误的是纠错码位。
 - 如果有多个，错误的是对应的非纠错码位。

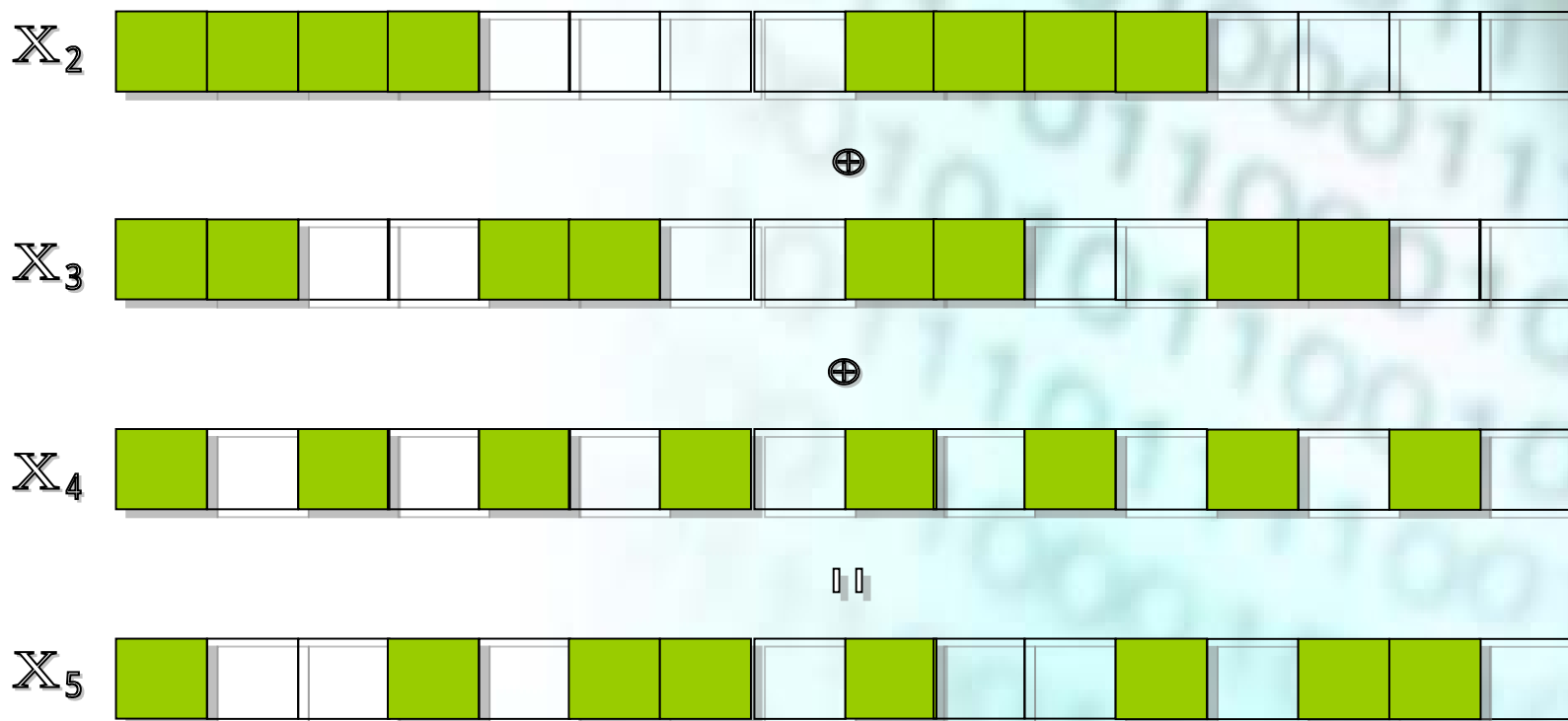
图三：纠错码的构造



- $m=\log n=4$, 所以由 $2^k \geq m+k+1$ 知 k 即纠错码的位数至少为 3 。

图四：纠错码与实际问题的对应

- 对编码的异或运算等价于对问题本身的异或。
- 以 $X_5 = X_2 \oplus X_3 \oplus X_4$ 为例。

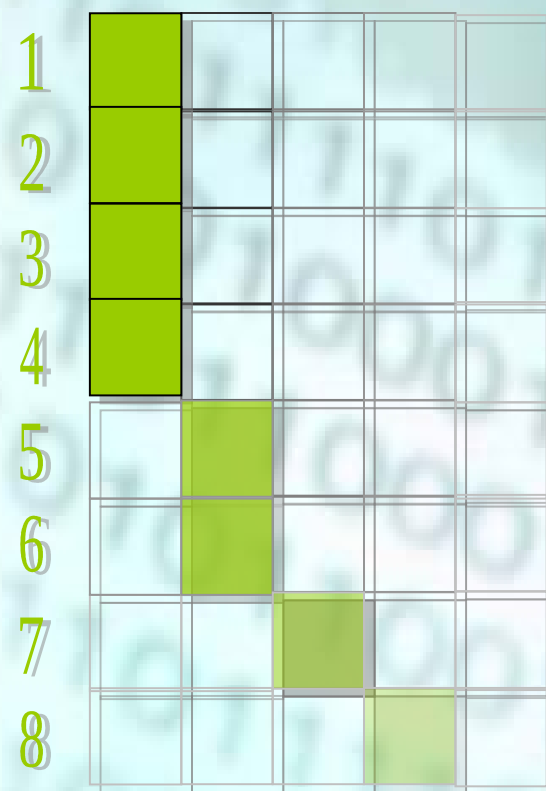


有关问题次数的一些结果

- 二分法的次数：
 - $2^k \geq kn + n$ 的最小解 k 。
- 编码方法的次数：
 - 除去必需的 m 位有效信息外，每个错误对应一个纠错码：
 - $2^{k-m} \geq k+1$ 。
 - 所有的可能性每个至少对应一个不同的编码：
 - $2^k \geq 2^m(k+1)$ 。
 - 结果：满足不等式 $2^{k-m} \geq k+1$ 的最小解 k 。
- 两者的比较：
 - $m = \log n$, $2^k \geq kn + n \Leftrightarrow 2^k \geq k2^m + 2^m \Leftrightarrow 2^{k-m} \geq k+1$ 。

扩展 1：更多错误时的解法

- 在此情况下，编码方法的扩展是很困难的。
- 对于二分法的扩展：
 - 加入更多的分类，最多 k 个错误的时候分为 $k+1$ 类。
 - 平分每一类。



扩展 2：更多的回答，称球问题

- 筛选方法需要保存更多的元素。
- 编码的方法：
 - 扩展进制数。
 - 设计多进制的纠错码。
 - 称球问题：
 - 三进制。
 - 错误方式也要编码。
 - 三进制纠错码。





Ulam 的游戏和编码

- 多重错误的筛选。
- 编码：
 - 表示方法。
 - 纠错码的设计。
- 更多的扩展。

