Randomness & independence

Ben Langmead

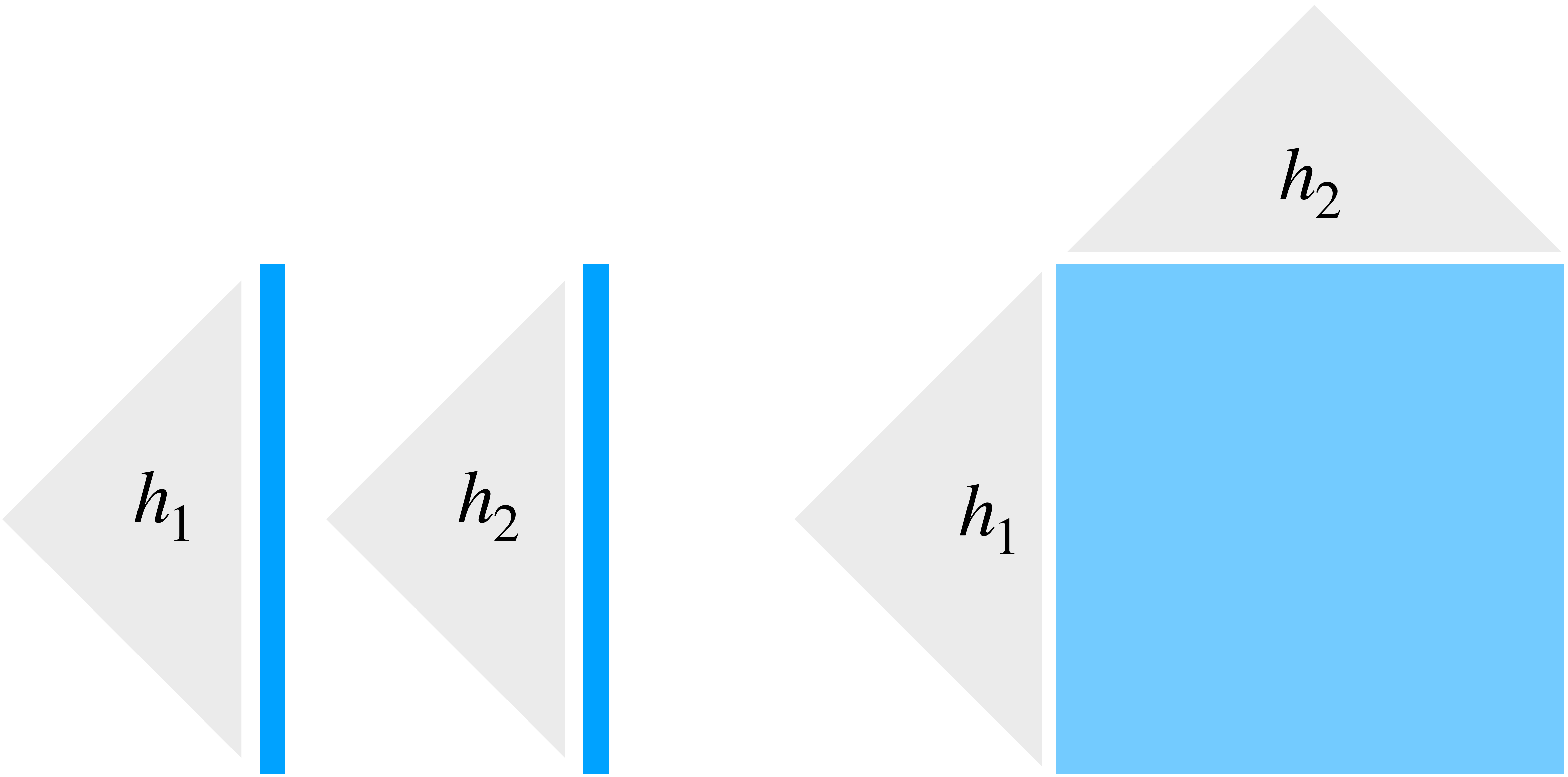
JOHNS HOPKINS
WHITING SCHOOL of ENGINEERING

Department of Computer Science

Randomness & independence

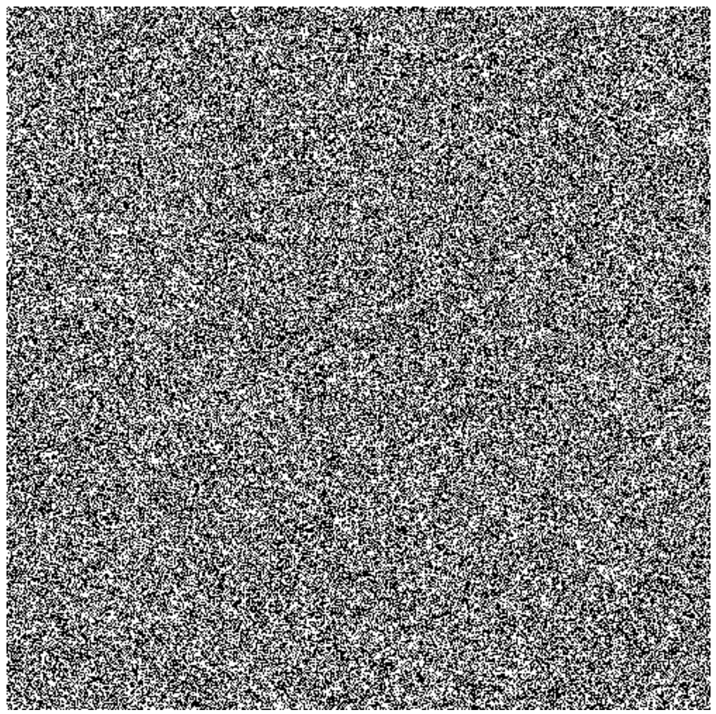
To make 2 uniform hash functions requires randomness

To make ***2 uniform & independent*** hashes requires more



Randomness

True randomness is hard to come by



73735	45963	78134	63873
02965	58303	90708	20025
98859	23851	27965	62394
33666	62570	64775	78428
81666	26440	20422	05720
15838	47174	76866	14330
89793	34378	08730	56522
78155	22466	81978	57323
16381	66207	11698	99314
75002	80827	53867	37797
99982	27601	62686	44711
84543	87442	50033	14021
77757	54043	46176	42391
80871	32792	87989	72248
30500	28220	12444	71840

Might ultimately come from environment

Can be "amplified" deterministically, e.g. pseudo-random generation

https://en.wikipedia.org/w/index.php?title=Special:CiteThisPage&page=A_Million_Random_Digits_with_100%2C000_Normal_Deviates&id=932646504

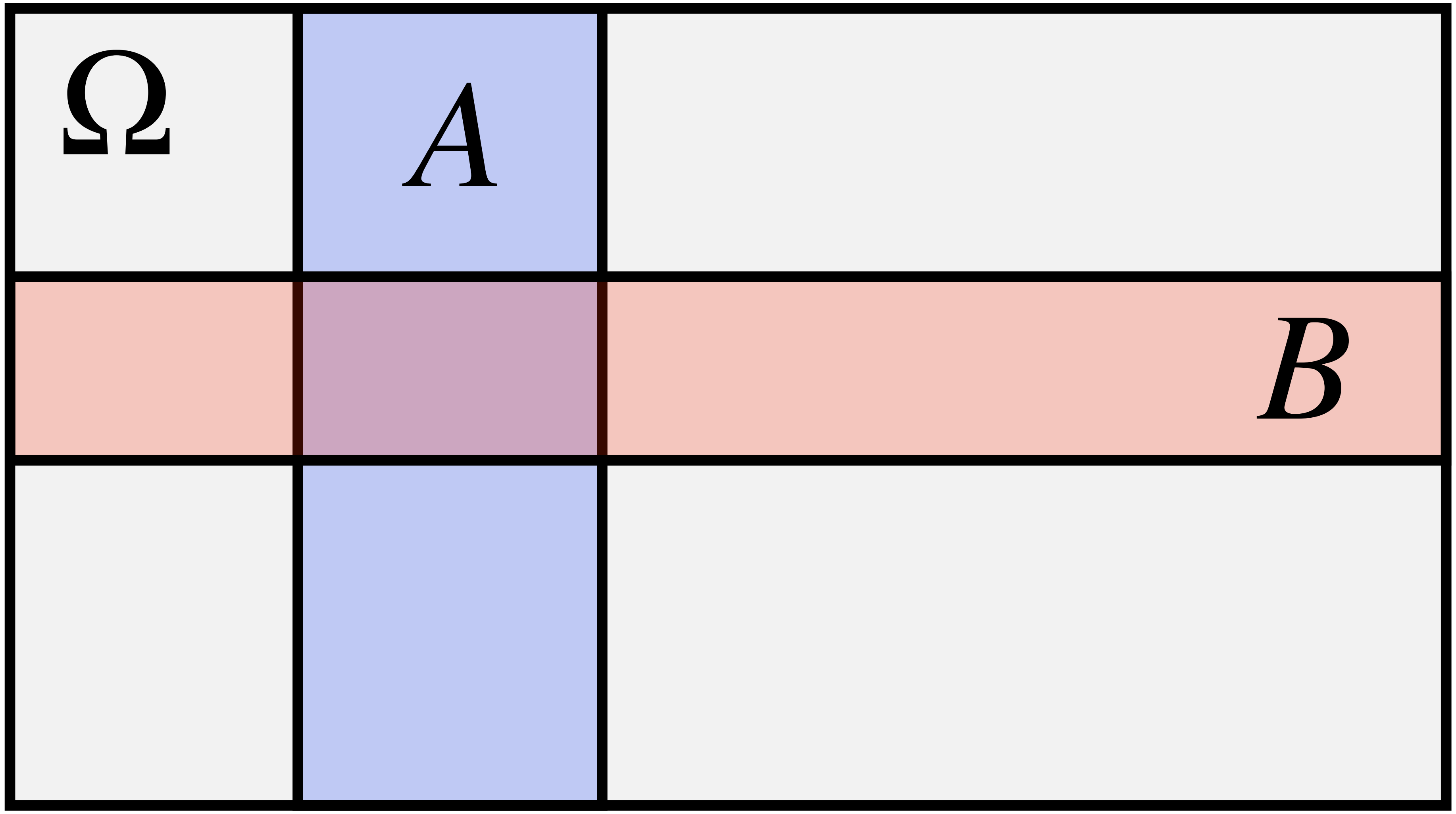
Independence

Events $E_1, E_2, \ldots, E_k$ are **mutually** independent if and only if, for any subset $I \subseteq [1, k]$

$$\Pr\left(\bigcap_{i \in I} E_i\right) = \prod_{i \in I} \Pr(E_i)$$

Independence

Independent *events*



$$P(A, B) = P(A) \cdot P(B)$$

Independence

Two r.v.s X and Y are independent when:

$$\Pr(X = x \cap Y = y) = \Pr(X = x) \cdot \Pr(Y = y)$$

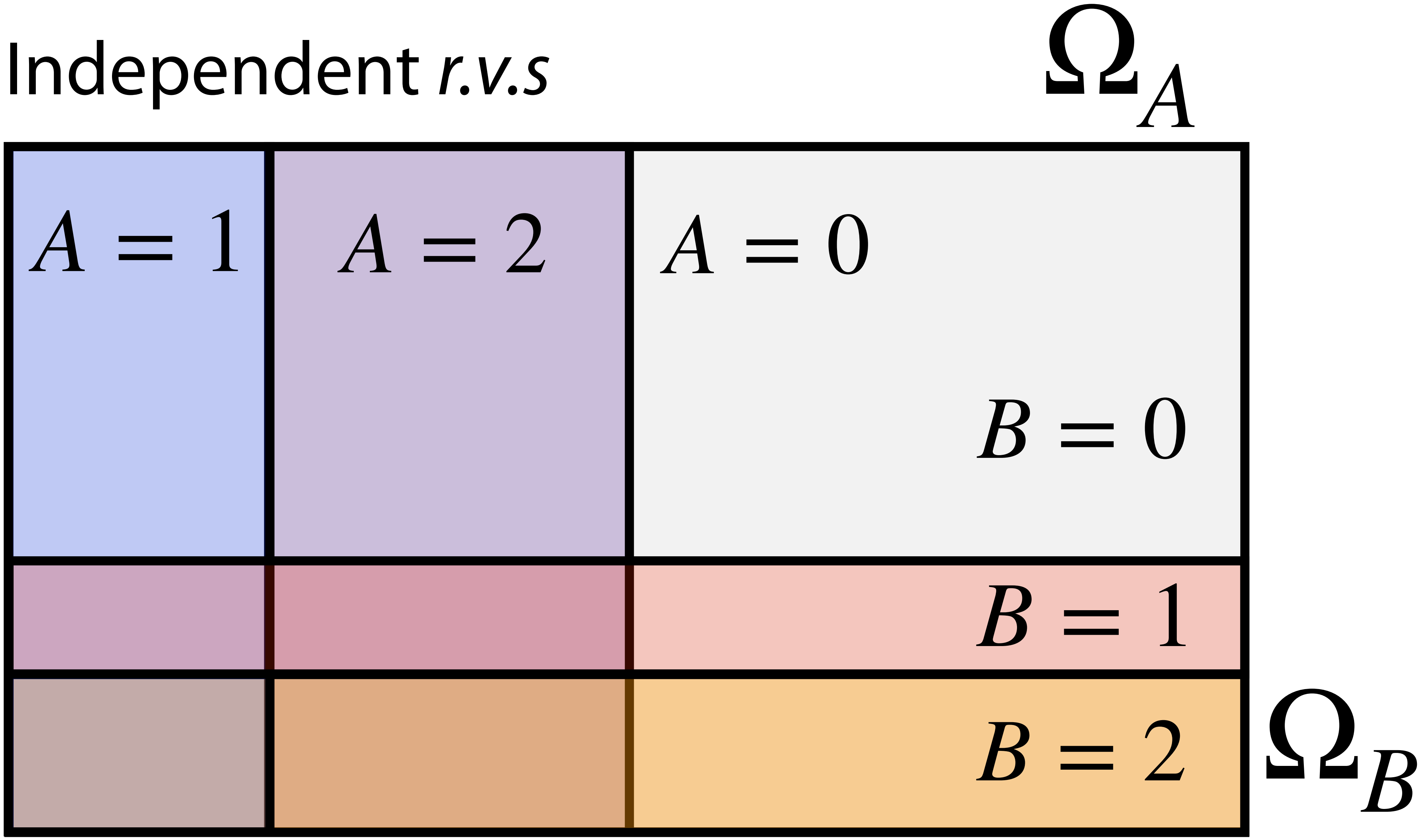
for all x, y

R.v.s $X_1, X_2, \ldots, X_k$ are **mutually** independent when, for any $I \subseteq [1, k]$ and values $x_i, i \in I$

$$\Pr\left(\bigcap_{i \in I} (X_i = x_i)\right) = \prod_{i \in I} \Pr(X_i = x_i)$$

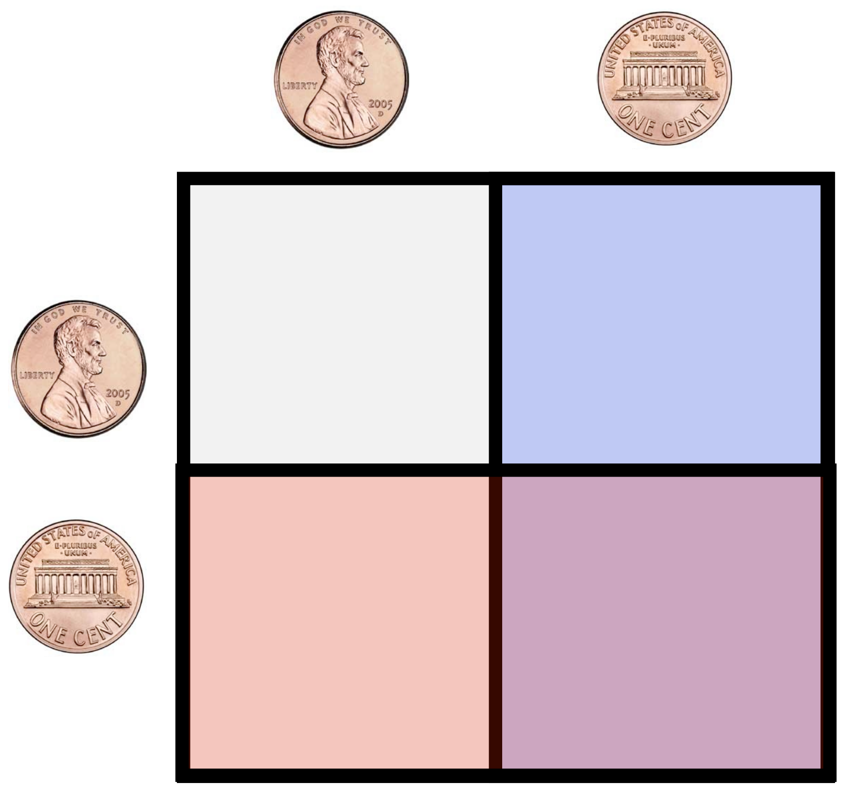
Independence

Independent *r.v.s*

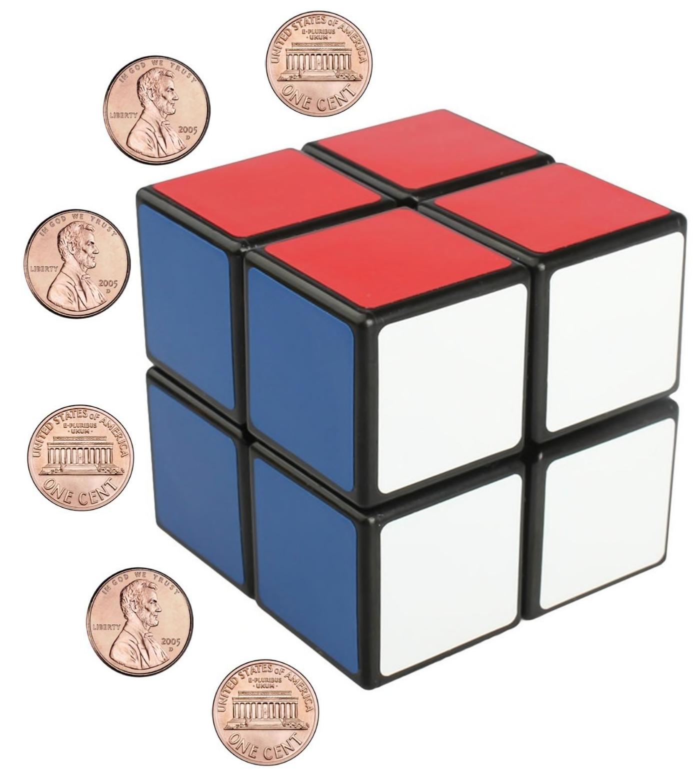


$$P(A = a, B = b) = P(A = a) \cdot P(B = b)$$

Independence



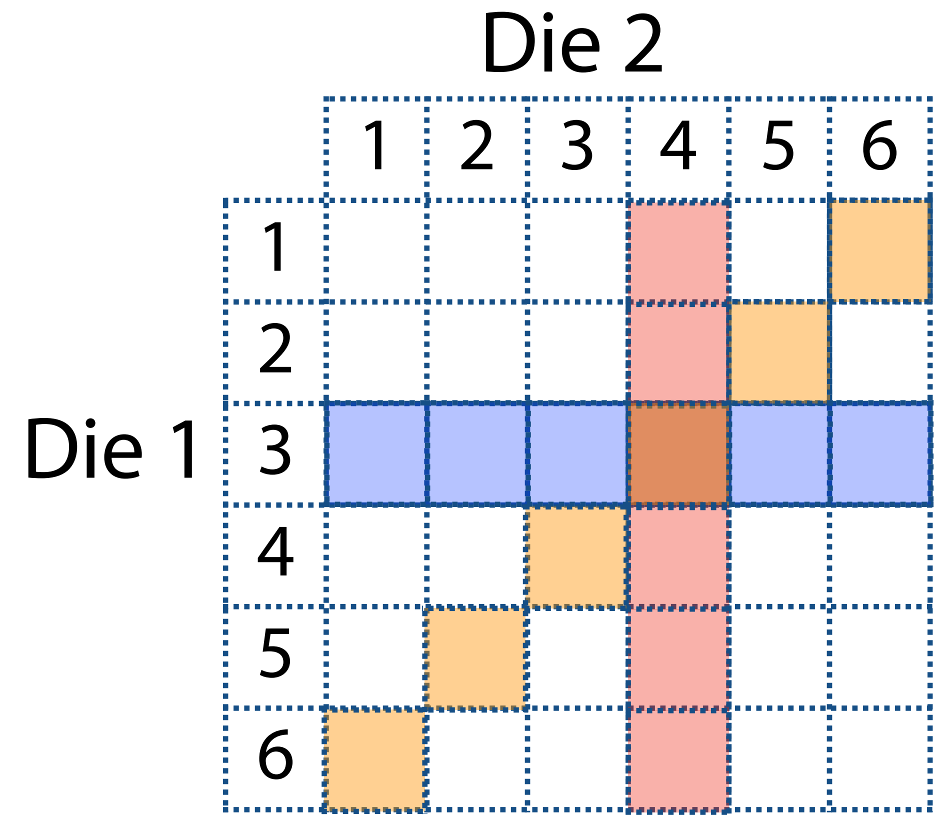
Pairwise independent



3-wise independent

When would we have pairwise independence but not 3-wise independence?

Independence



1. Event A: Die 1 = 3
2. Event B: Die 2 = 4
3. Event C: Sum of dice = 7

Pairwise independence ✅

$$\Pr(A \cap B) = \Pr(A)\Pr(B) = 1/36$$
$$\Pr(B \cap C) = \Pr(B)\Pr(C) = 1/36$$
$$\Pr(A \cap C) = \Pr(A)\Pr(C) = 1/36$$

3-wise independence ❌

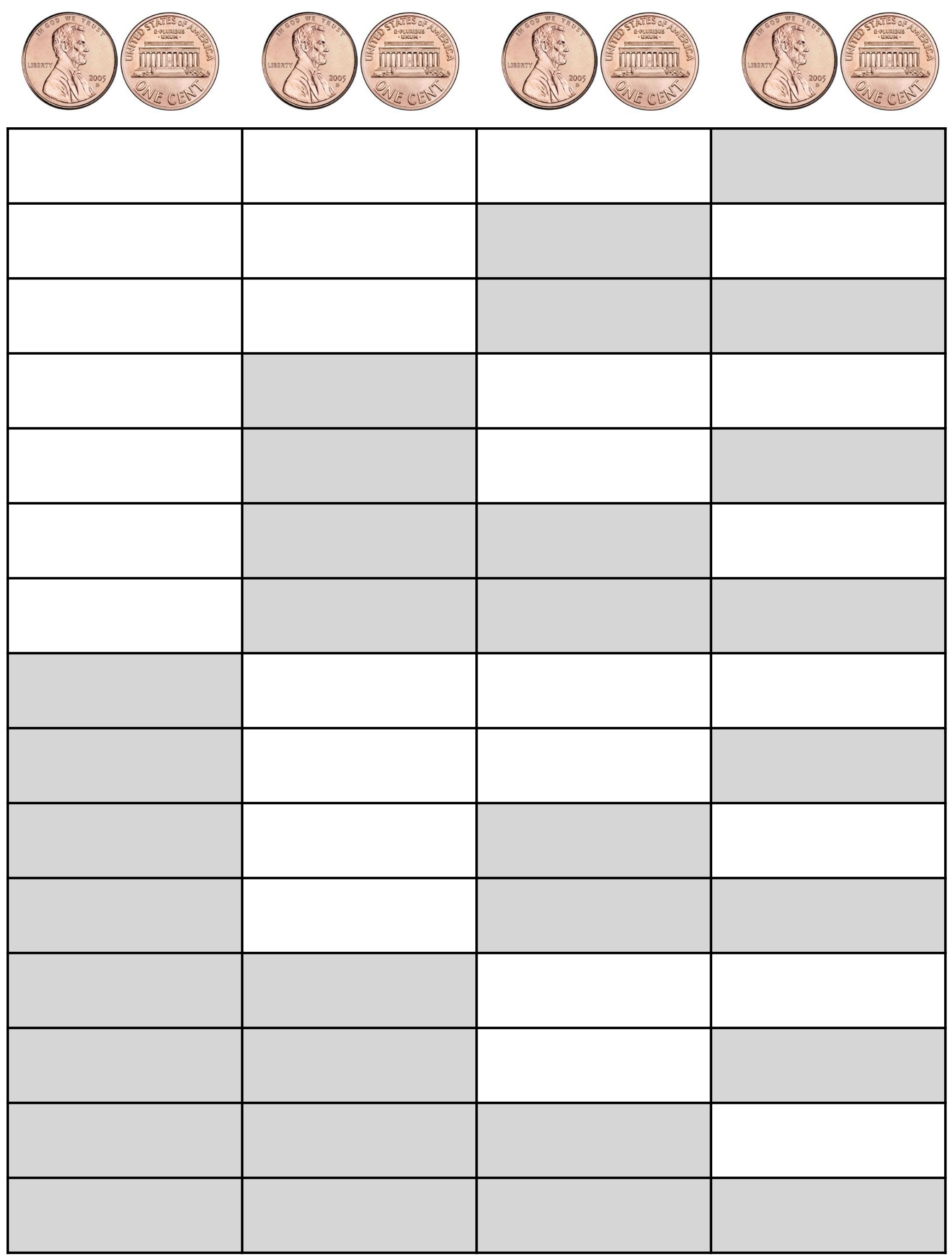
$$\Pr(A \cap B \cap C) = \frac{1}{36} \neq \frac{1}{216} = \Pr(A)\Pr(B)\Pr(C)$$

Randomness

Given a few *mutually* independent coin flips (bits), can I construct many *pairwise* independent flips?

Randomness

Mutually independent flips



Use coin flip?

yes no

Randomness

Mutually independent flips

0	**1**	**0**	**1**		
			1	$\oplus$	
		0		$\oplus$	
		0	1	$\oplus$	
	1			$\oplus$	
	1		1	$\oplus$	
	1	0		$\oplus$	
	1	0	1	$\oplus$	
0				$\oplus$	
0			1	$\oplus$	
0		0		$\oplus$	
0		0	1	$\oplus$	
0	1			$\oplus$	
0	1		1	$\oplus$	
0	1	0		$\oplus$	
0	1	0	1	$\oplus$	

Use coin flip?

yes | no

XOR

Randomness

Mutually independent flips

0	1	0	1		
			1	$\oplus$	1
		0		$\oplus$	0
		0	1	$\oplus$	1
	1			$\oplus$	1
	1		1	$\oplus$	0
	1	0		$\oplus$	1
	1	0	1	$\oplus$	0
0				$\oplus$	0
0			1	$\oplus$	1
0		0		$\oplus$	0
0		0	1	$\oplus$	1
0	1			$\oplus$	1
0	1		1	$\oplus$	0
0	1	0		$\oplus$	1
0	1	0	1	$\oplus$	0

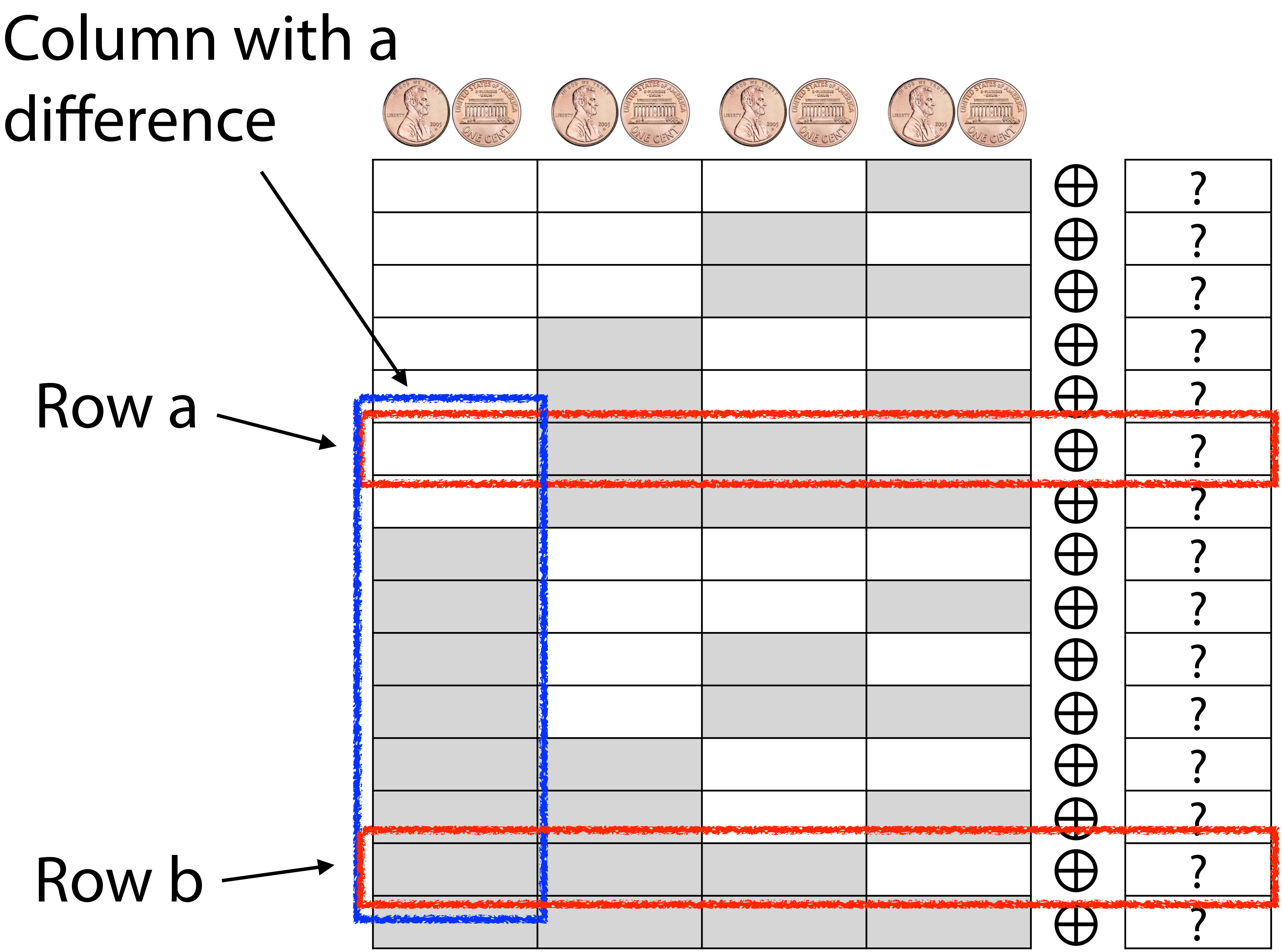
Use coin flip?

yes | no

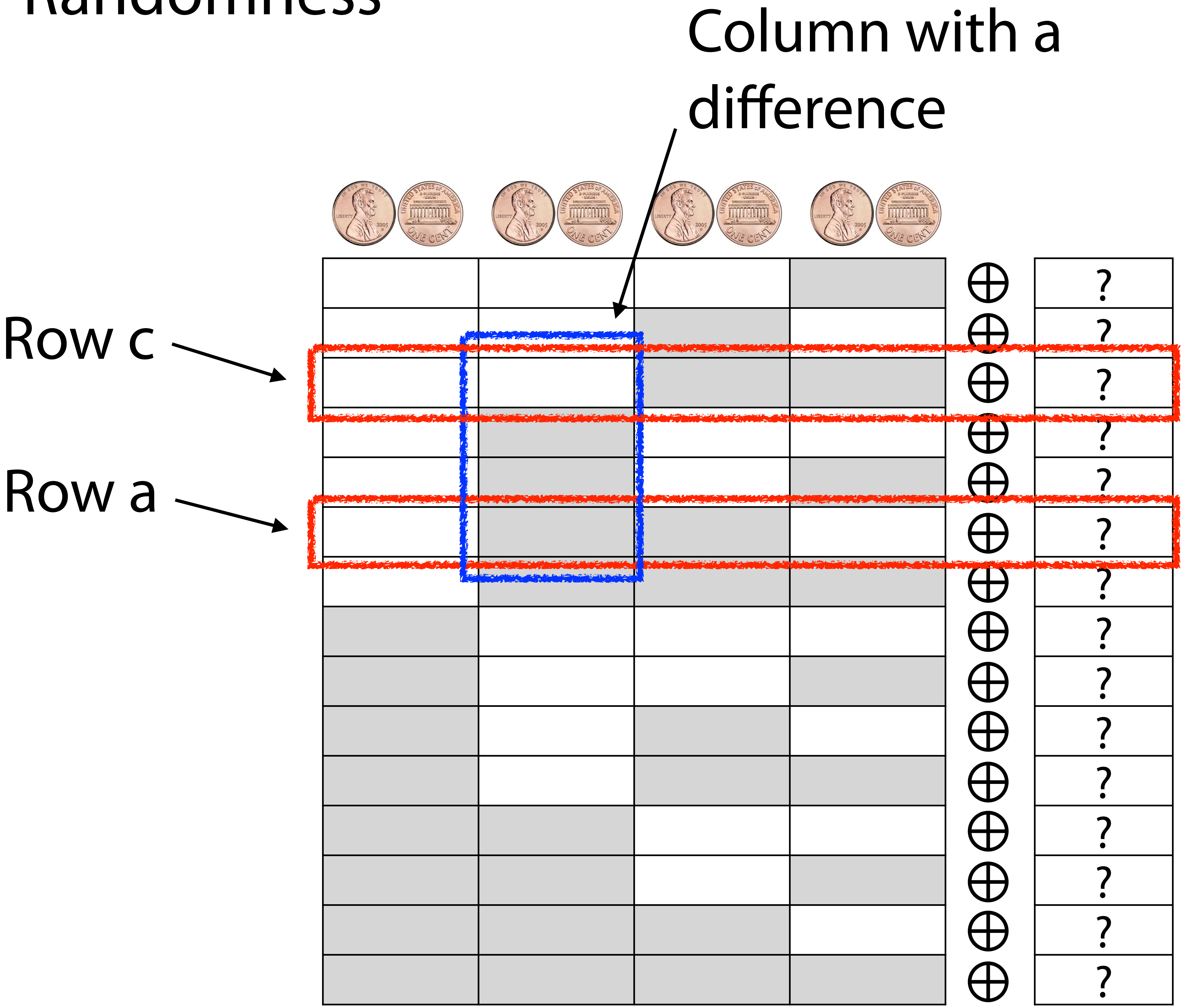
Are these pairwise independent?

XOR

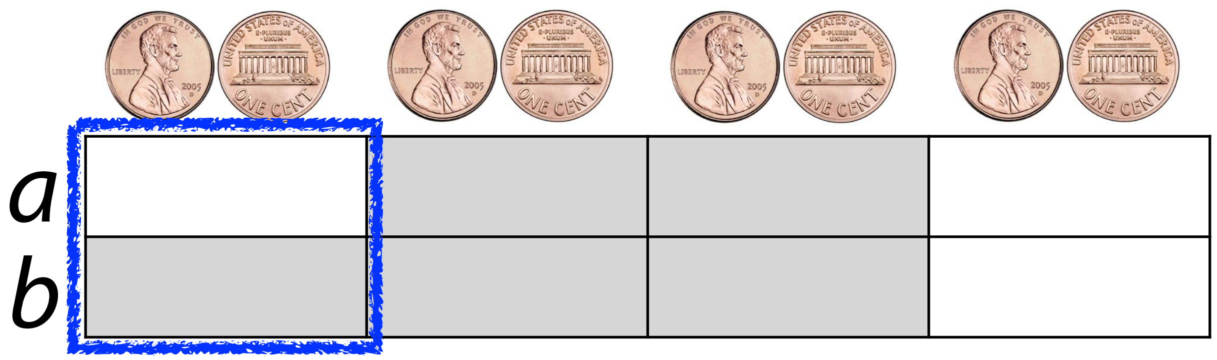
Randomness

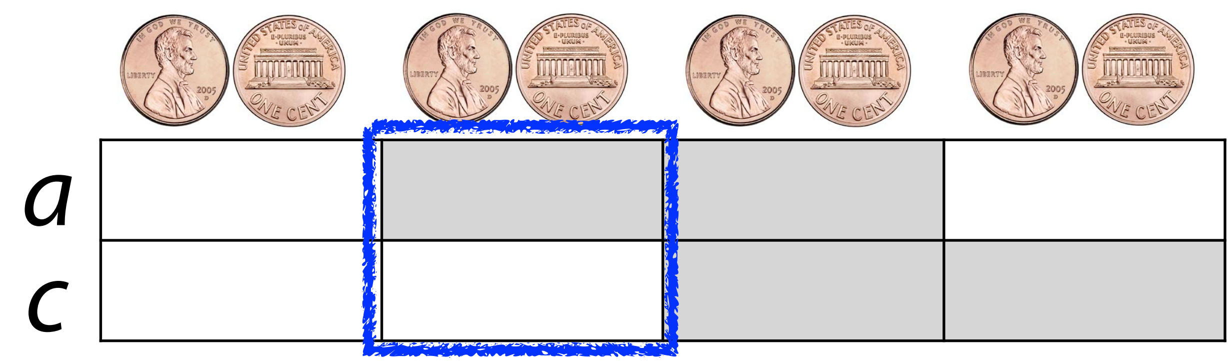


Randomness



Randomness





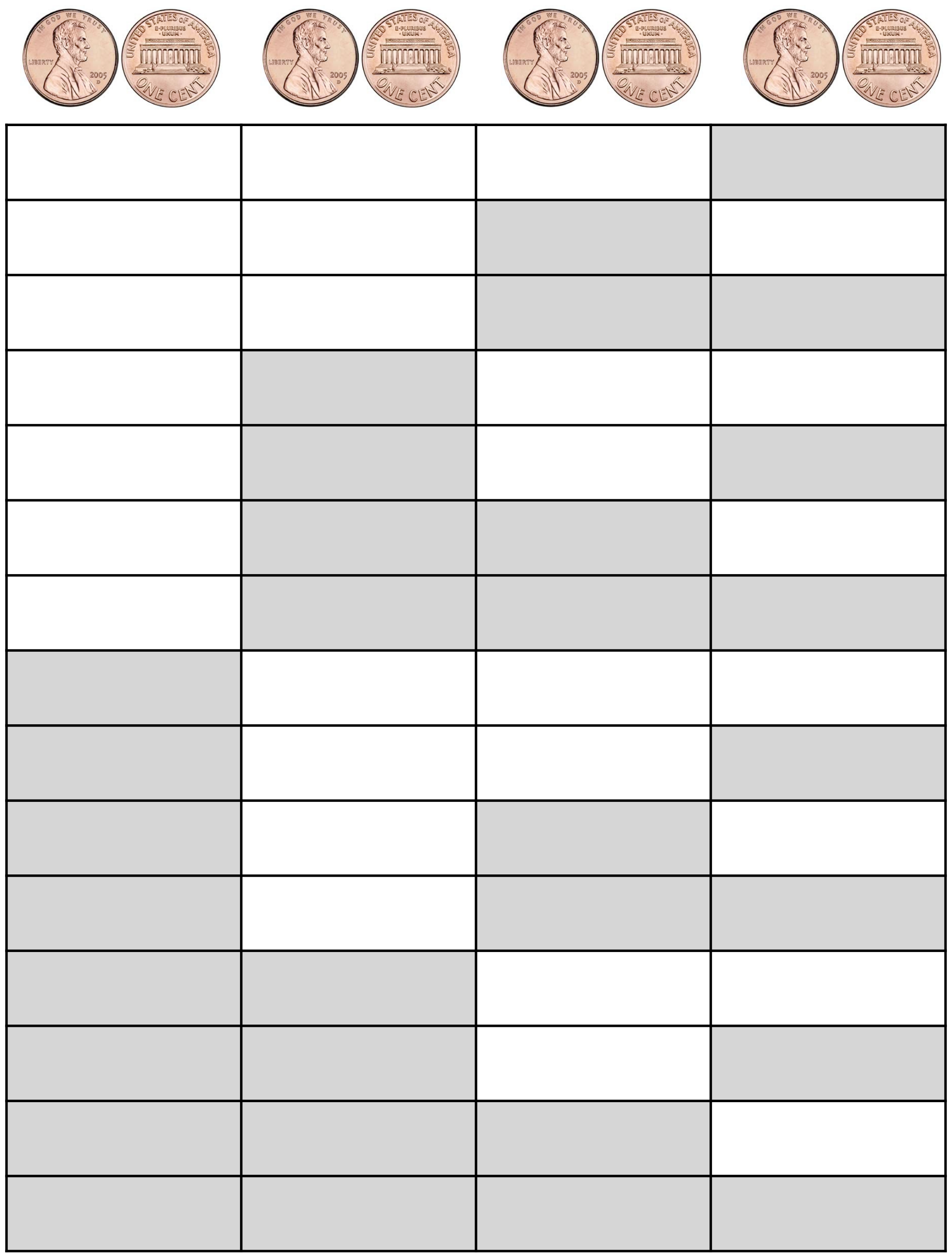
Use principle of deferred decisions: assume we do all the XOR'ing outside the blue column first

Final value for one row is determined by *new coin flip*

"New": not yet used in either row

With final flip, we have pairwise independence

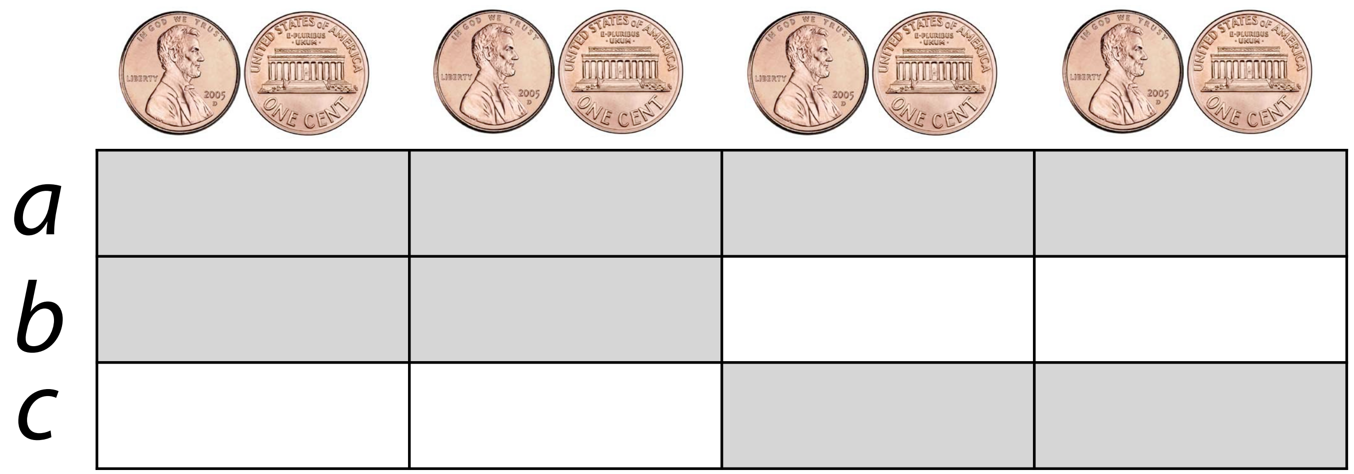
Randomness



We can find a column with a difference for any pair of rows, by construction

Do we have 3-wise independence?

Randomness



row a = row $b \oplus$ row c

No 3-wise independence among rows X