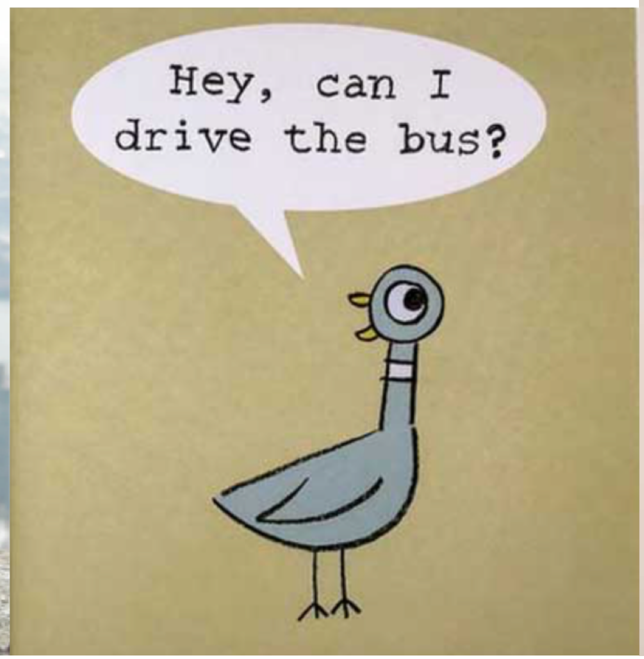
Today’s slides & handout: Canvas > Files > Lecture slides > Annotated Slides (Diaz)

Lecture 13
The Pigeonhole Principle



Announcements

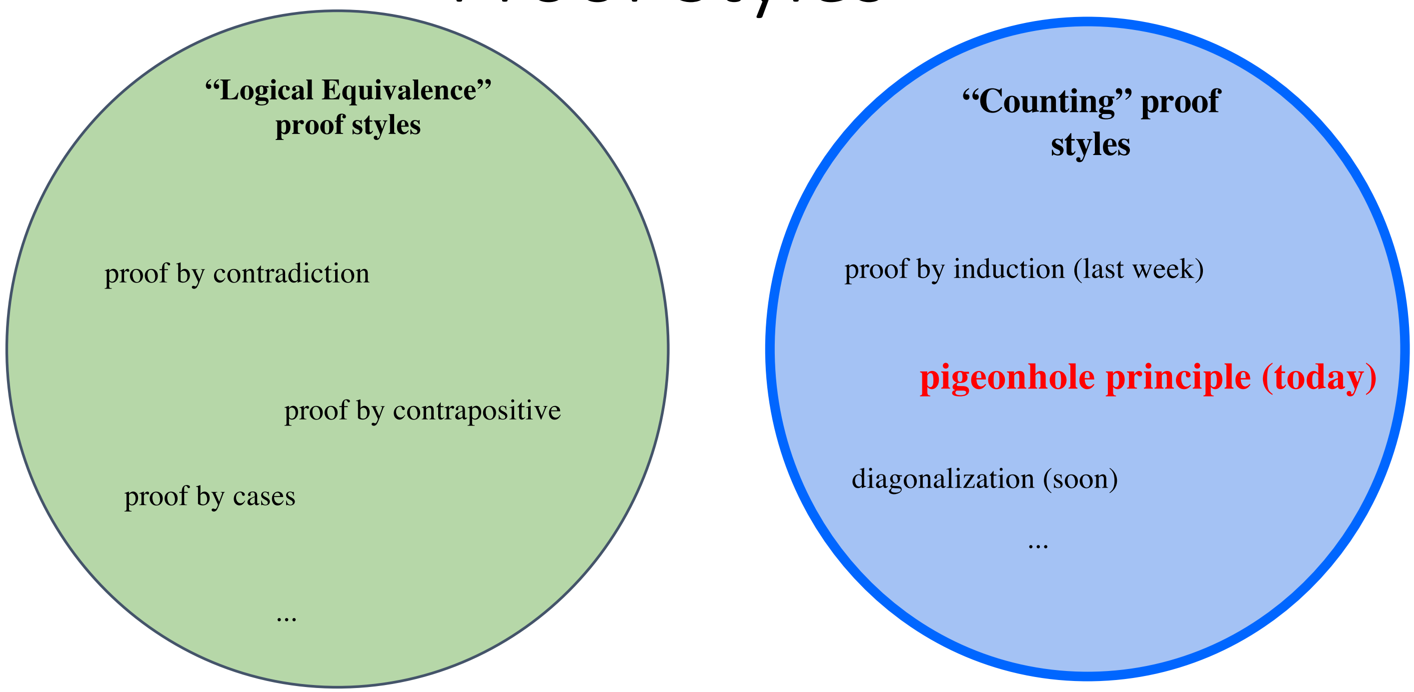
- Reminder: do not discuss the exam with anyone until we post solutions
- It usually takes about a week for us to grade all of the exams

Learning Objectives: Lec 13

After today's lecture (and the associated readings, discussion, & homework), you should be able to:

- **Know what the pigeonhole principle is.**
- **Use counting arguments like the pigeonhole principle in proofs.**

Second installment of “Counting” Proof Styles



Valid because of **logical equivalences**

Valid because of **counting ideas ...**

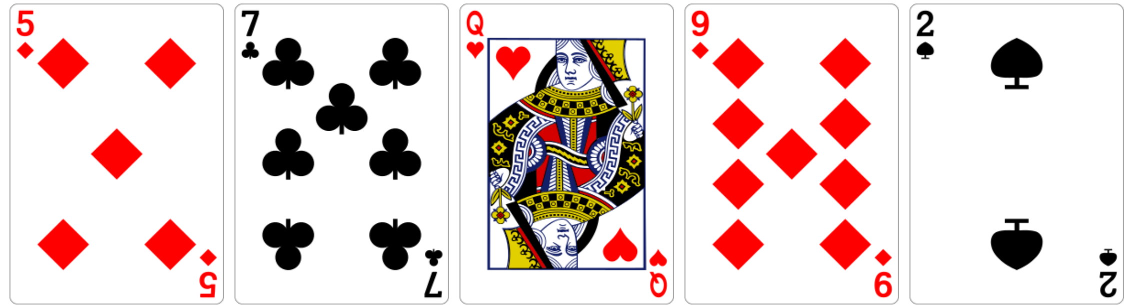
Outline

- **The card trick**
- Lunch orders
- Sets of numbers
- Handshakes
- The game of 30 questions
- The card trick revisited

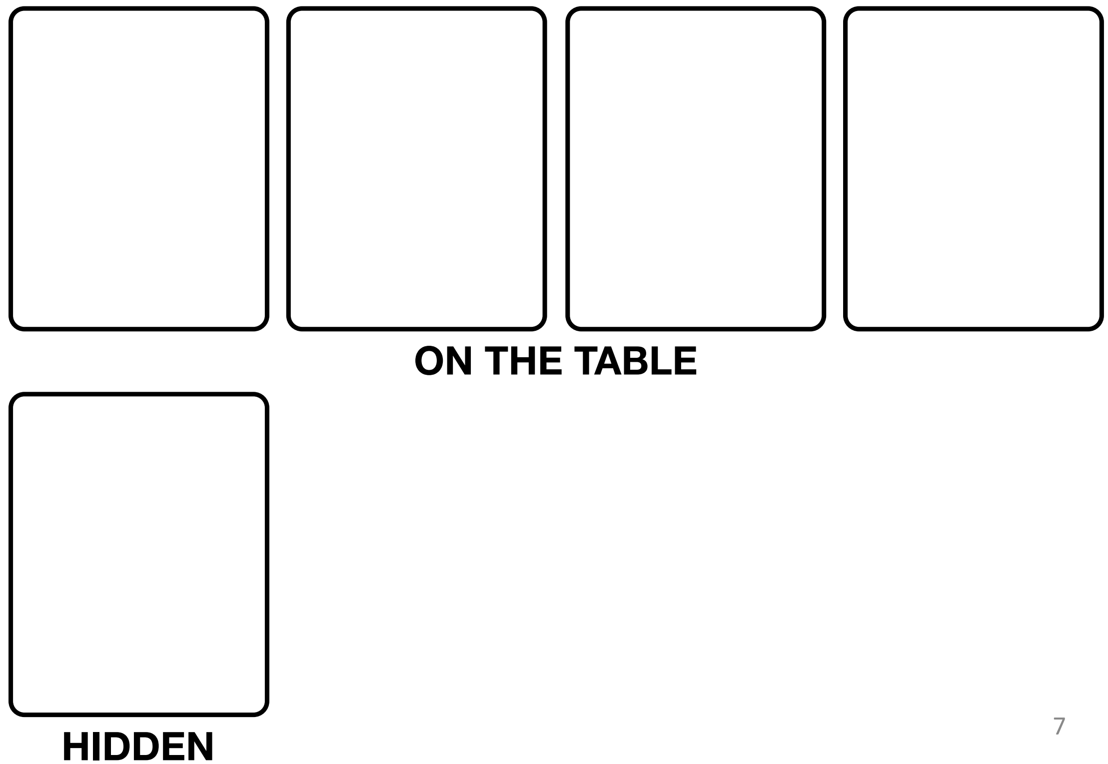
The Card Trick

- **Person A** is dealt a **hand of 5 cards**, examines it, then lays down a **sequence of 4** of them, face-up on the table. The 5th card remains hidden (face down).

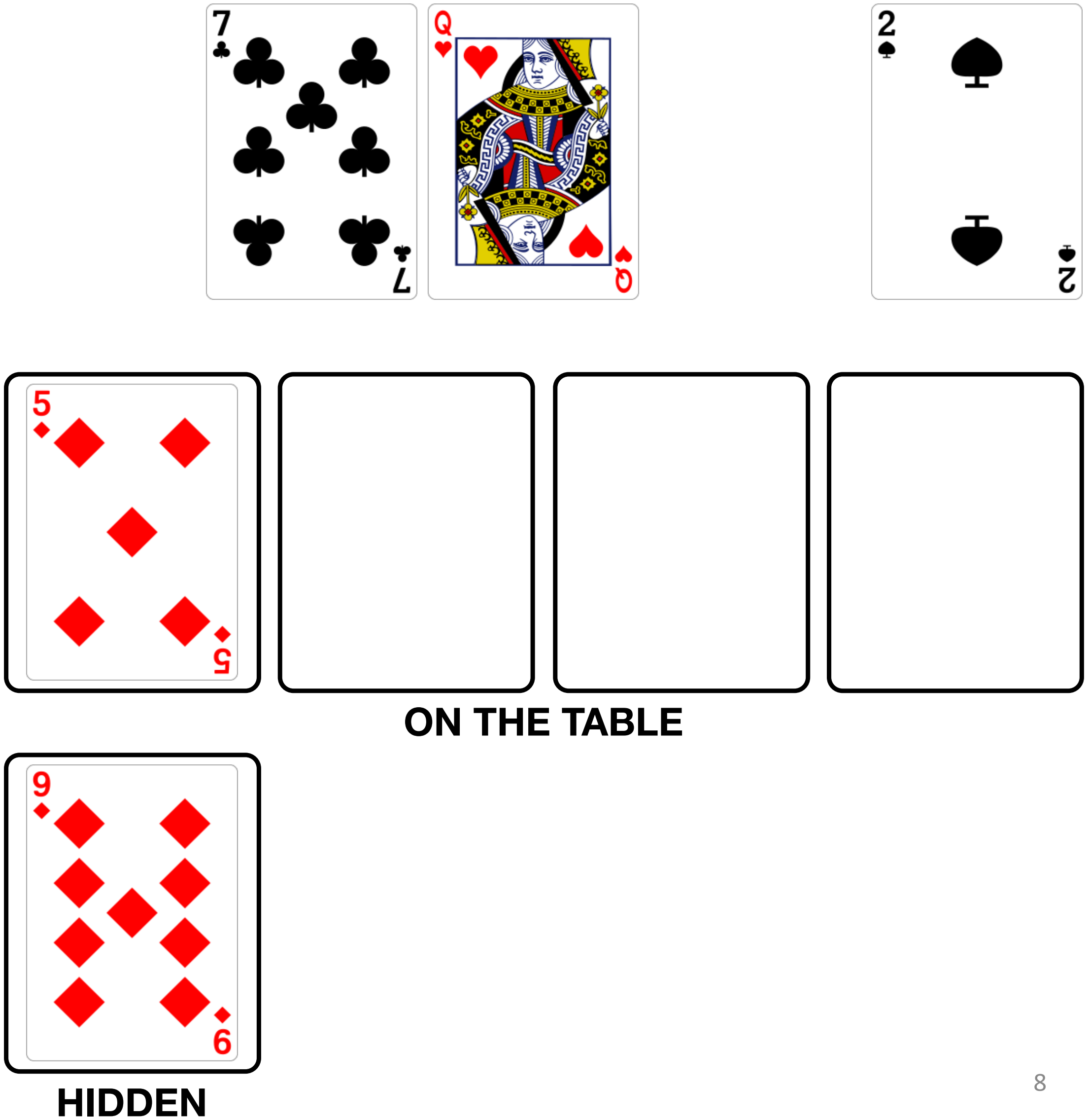
- **Person B** looks at the **4 visible cards** and guesses (correctly!) **Person A**'s **5th card**.



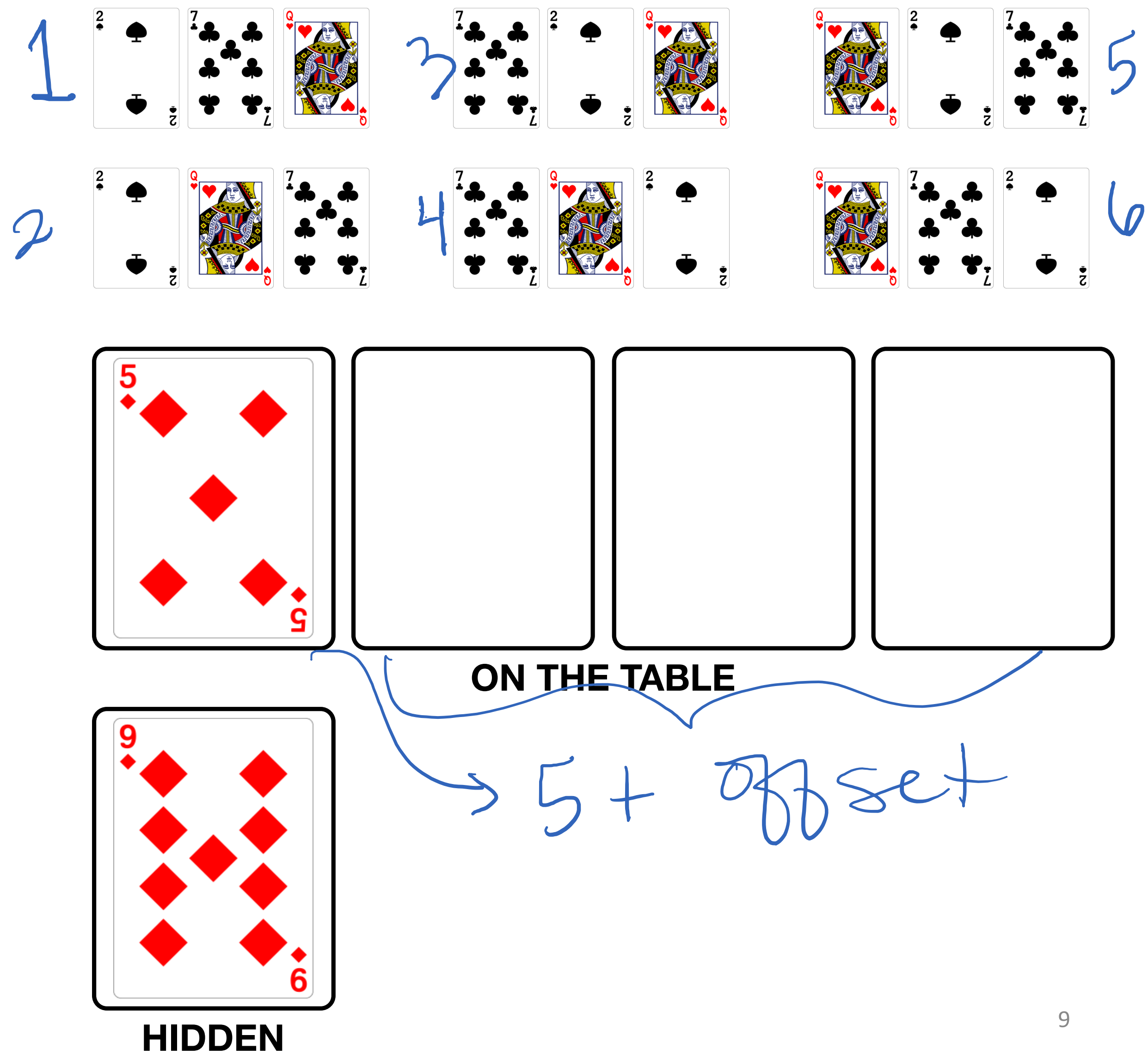
- If you’re **Person A**, you must be prepared to see ***any*** 5-card hand.
- Which card do you keep hidden?
- How can you encode its suit?
- … and encode its number?



- The hidden card is a ♦!

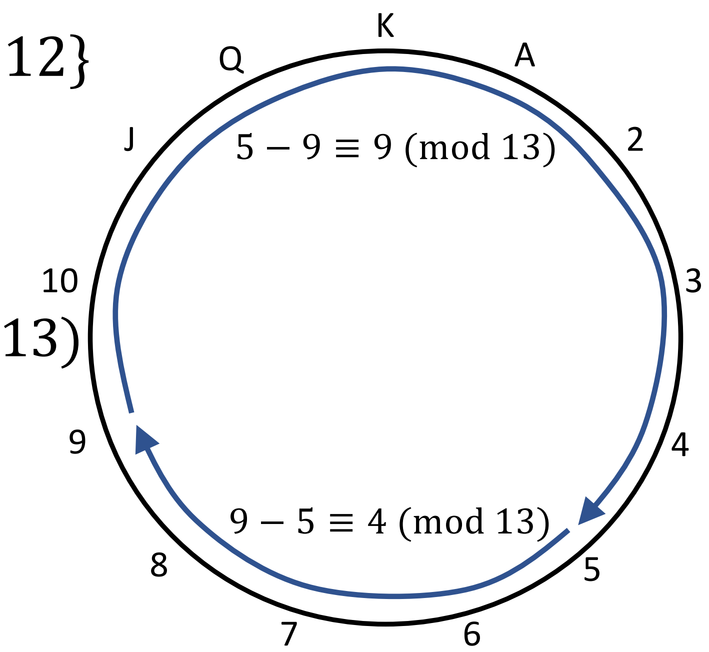


- There are 6 = 3! ways to arrange the remaining 3 cards.

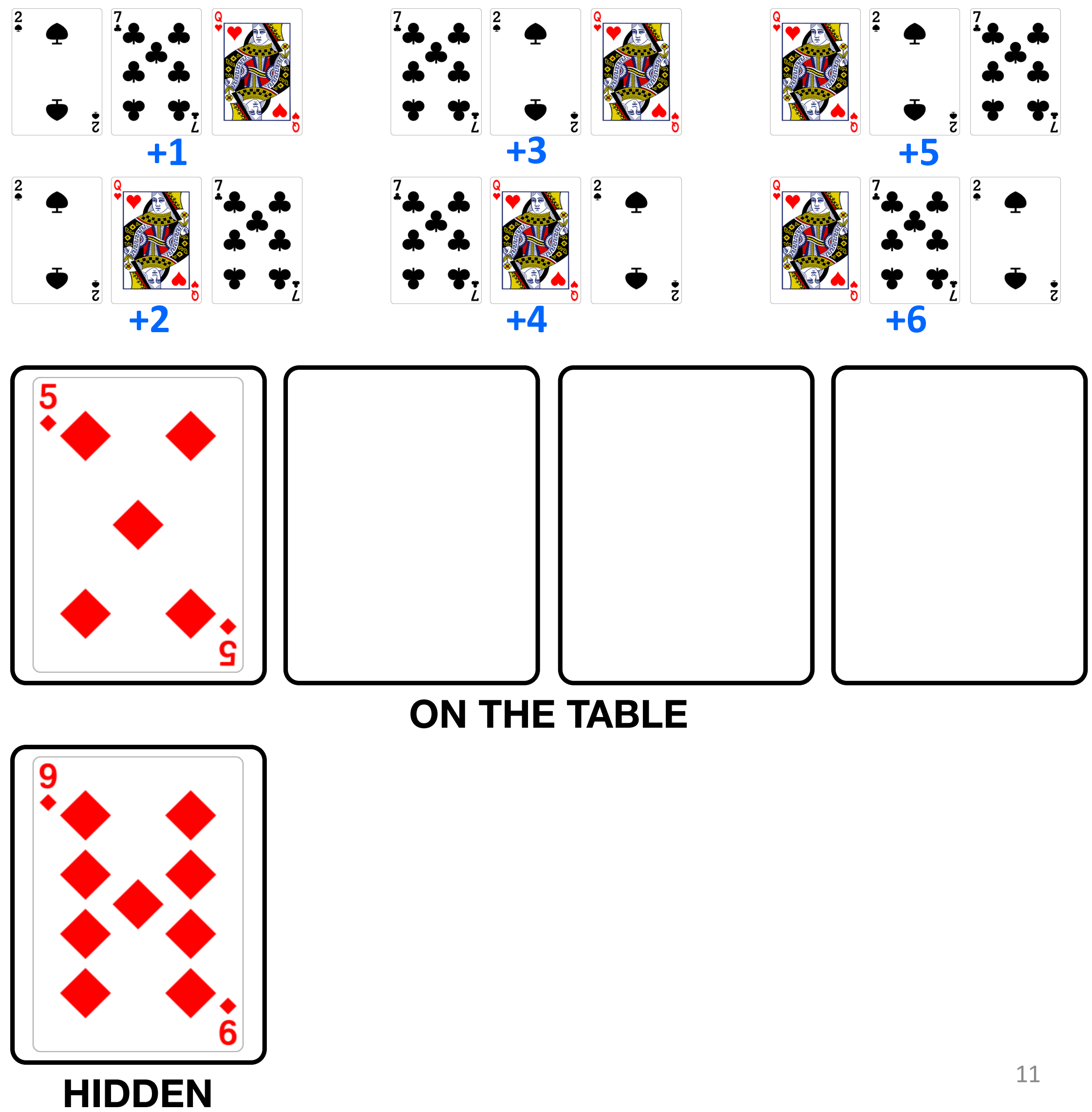


Encoding the number.

- A=1, 2, 3, 4, 5, 6, 7, 8, 9, 10, J=11, Q=12, K=0.
- Note: for any distinct integers $x, y \in \{0,1,\ldots,12\}$
  - Either $x - y \in \{1,2,\ldots,6\} \pmod{13}$
  - Or $x - y \in \{7,8,\ldots,12\} \pmod{13}$
  - But $\{7,\ldots,12\} = \{-6,-5,\ldots,-1\} \pmod{13}$
  - So either $x - y$ or $y - x$ is in $\{1,\ldots,6\} \pmod{13}$



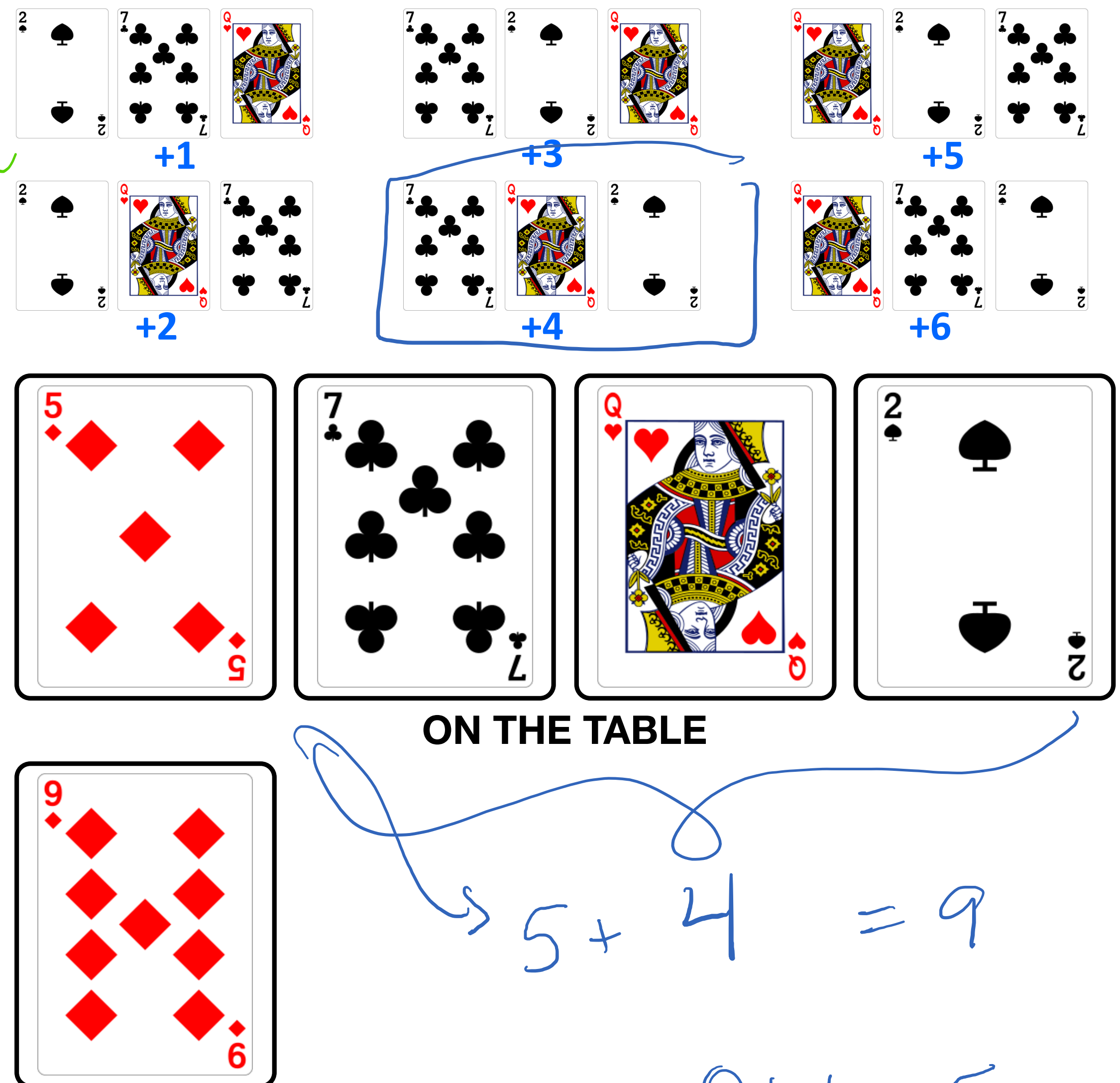
- Encode $9 - 5 \equiv 4 \pmod{13}$ with the remaining 3 cards.



If 2 cards of same rank, use suit "order":

low → clubs, diamonds, hearts, spades ← high

- Encode $9 - 5 \equiv 4 \pmod{13}$ with the remaining 3 cards.



What if 5♢ Q♢ instead of 9

{ Q goes face up
{ 5 goes face down

encode: 6

5 + 4 = 9

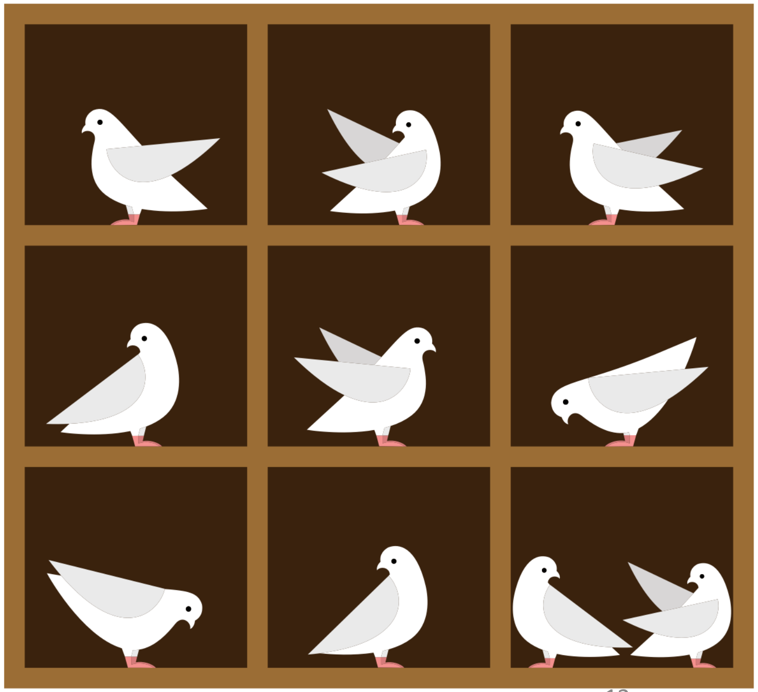
because Q + 6 = 5

The Pigeonhole Principle

- There are N pigeonholes and $N+1$ pigeons. If you place each pigeon in a pigeonhole, at least 2 pigeons must be in some hole.

- *Generalized* Pigeonhole Principle:
  - If at least $kN+1$ pigeons are assigned to N pigeonholes, there must be at least $k+1$ pigeons in some hole.

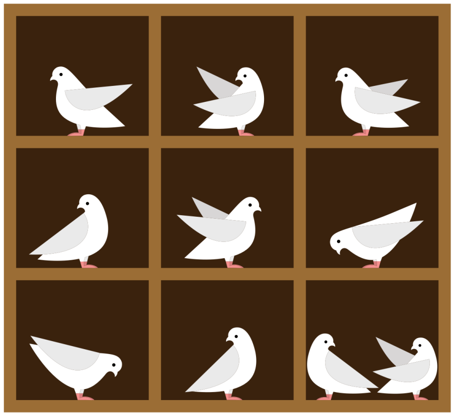
- It looks simple! But it's a powerful proof technique. *The hard part is identifying the pigeon(hole)s.*



Lecture 13 Handout: Pigeonhole Principle

***Generalized* Pigeonhole Principle**

- If there are __KN+1__ pigeons in __N__ pigeonholes, then there must be at least __K+1__ pigeons in some hole.



Handout

Some examples

- How many cards do we need to ***guarantee*** 2 of the same suit?

pigeons = cards → 5
holes = suits → 4

- How many people do you need to ***guarantee*** that 2 have the same birth month?

pigeons = people → 13
holes = months → 12

  - … to ***guarantee*** that **6** people have the same birth month?

pigeons = ppl
holes = months

$5 \cdot 12 + 1 = 61$

- Do 2 students in EECS 203 have the same two-letter initials?

YES, there are >1000 students in 203, we just need 677.

pigeons: students → need $676+1 = 677$

holes: 2-letter initials $= 26^2 = 676$

# of 2-letter initials $= 26^2$

26 choices 26 choices

Some examples

Handout

$\lceil \; \rceil$ = ceiling

- How many cards do we need to ***guarantee*** 2 of the same suit?
  - Pigeons: 5 cards, pigeonholes: 4 suits. $\lceil 5/4 \rceil = 2$.
- How many people do you need to ***guarantee*** that 2 have the same birth month?
    - Pigeons: 13 people, pigeonholes: 12 months. $\lceil 13/12 \rceil = 2$.
  - … to ***guarantee*** that **6** people have the same birth month?
    - Pigeons: 61 people, pigeonholes: 12 months. $\lceil 61/12 \rceil = 6$.
- Do 2 students in EECS 203 have the same two-letter initials?
  - This is not a "universally" true statement, and in fact wasn't necessarily true 10 years ago. *It depends on the number of students enrolled in EECS 203.*
  - Pigeons: > 1000 students, Pigeonholes: $(26)^2 = 676$ possible two-letter initials.

Floor and Ceiling Functions

- The floor function of x, denoted $\lfloor x \rfloor$
  - The largest integer smaller than or equal to x
  - Examples:

    $\lfloor 1.2 \rfloor = 1$, $\lfloor 3 \rfloor =$ **?**, $\lfloor -2.5 \rfloor =$ **?**
- The ceiling function of x, denoted $\lceil x \rceil$
  - The smallest integer greater than or equal to x

    $\lceil 1.2 \rceil = 2$, $\lceil 3 \rceil =$ **?**, $\lceil -2.5 \rceil =$ **?**
- Some useful properties in Rosen Section 2.3

Outline

- The card trick
- **Lunch orders**
- **Sets of numbers**
- Handshakes
- The game of 30 questions
- The card trick revisited

Lunch A E D

Handout

2 choices 2 choices 3 choices.

- A group of G people go to lunch, where the options for each course are [soup or salad] + [chicken or tofu] + [cookie or pie or chocolate mousse]. How big must G be to ensure that two people order the same lunch?

holes : possible meals

$\# \text{holes} = 2 \cdot 2 \cdot 3 = 12$

$M = \{ (\text{salad}, \text{tofu}, \text{cookie}), (\text{soup}, \text{tofu}, \text{pie}), \ldots \}$

$M = A \times E \times D$

$|M| = |A|\,|E|\,|D| = 2 \cdot 2 \cdot 3$

pigeons = people

need 13 people for 2 to have the same meal.

Handout

Lunch

- A group of G people go to lunch, where the options for each course are [soup or salad] + [chicken or tofu] + [cookie or pie or chocolate mousse]. How big must G be to ensure that two people order the same lunch?

- (A) 7
- (B) 8
- (C) 9
- (D) 12
- (E) 13

Sets of numbers

Handout

take or leave } 2 choices

2 choices

- Pick a set S of ten distinct positive 2-digit integers as you like.

  Example: $S = \{11, 24, 39, 46, 47, 50, 71, 88, 94, 97\}$.

claim

- **Lemma.** There are two subsets of S that have the same sum.

  Example: $\{39, 46, 47, 50\}$ and $\{88, 94\}$ both sum up to $\mathbf{182}$.

  - Where are the pigeons and pigeonholes?

pigeons = subsets → # subsets = $|P(S)| = 2^{10} = 1024$

holes = sums

↳ largest sum: $\underbrace{\{90, 91, 92, \dots 99\}}_{\text{sum} < 10 \cdot 99 = 990}$

sums < 990

By PHP, since there are more subsets than sums, 2 subsets have to have the same sum.

Sets of numbers

Handout

- Pick a set S of ten distinct positive 2-digit integers as you like.

  Example: $S = \{11, 24, 39, 46, 47, 50, 71, 88, 94, 97\}$.
- **Lemma.** There are two subsets of S that have the same sum.

  Example: $\{\mathbf{39}, \mathbf{46}, \mathbf{47}, \mathbf{50}\}$ and $\{\mathbf{88}, \mathbf{94}\}$ both sum up to $\mathbf{182}$**.**
  - Where are the pigeons and pigeonholes?
  - Q1. What is the largest sum that you can have?
  - Q2. How many subsets of S are there?

Sets of numbers

- Pick a set S of ten distinct positive 2-digit integers as you like.

  Example: $S = \{11, 24, 39, 46, 47, 50, 71, 88, 94, 97\}$.
- **Lemma.** There are two subsets of S that have the same sum.

  Example: $\mathbf{\{39, 46, 47, 50\}}$ and $\mathbf{\{88, 94\}}$ both sum up to $\mathbf{182}$**.**
- **Proof.**
  - 1. There are ten numbers, each less than 100, so all sums are in the range $\mathbf{[0, 1000]}$.
  - 2. The number of subsets is $|P(S)| = 2^{|S|} = \mathbf{1024}$.
  - 3. By the P.H.P., two subsets must have the same sum.

Outline

- The card trick
- Lunch orders
- Sets of numbers
- **Handshakes**
- **The game of 30 questions**
- The card trick revisited

Handshakes

Handout

- **The set-up:** 50 people are at a party, and every pair of people either ***does*** or ***doesn't*** shake hands. At the end of the evening everyone counts how many people they shook hands with.
- **The problem:** must there be two people who shook hands with the same number of people?

YES!

pigeons = people → 50

holes = # handshakes → $\underbrace{\{0, 1, 2, \ldots, 49\}}_{50}$

2 cases: $\{0, 1, 2, \ldots, 48\}$

49 $\{1, 2, \ldots, 49\}$

Handshakes

Handout

- **The set-up:** 50 people are at a party, and every pair of people either ***does*** or ***doesn't*** shake hands. At the end of the evening everyone counts how many people they shook hands with.
- **The problem:** must there be two people who shook hands with the same number of people?
  - Pigeons: people. 50 of them!
  - Pigeonholes: numbers in the set $\{0,1,\ldots,49\}$. 50 of them!

Handshakes

- **Lemma.** There must be two people who shook hands with the same number of people.

- **Proof by cases.**
  - Case 1: There is ***someone*** who shook hands with ***no one***. This implies that the maximum hand-shake-number is 48 (not 49).
    - Pigeons: 50 people. Pigeonholes: hand-shake numbers in $\{0,1,\ldots,48\}$.
    - By PHP, two people shook hands the same number of times.
  - Case 2: There is ***no one*** who shook hands with ***no one***.
    - Pigeons: 50 people. Pigeonholes: by assumption, all numbers are in $\{1,2,\ldots,48,49\}$.
    - By PHP, two people shook hands the same number of times.
  - Case 1 and 2 are exhaustive, and both cases imply the lemma.

The Game of 30 Questions.

Handout

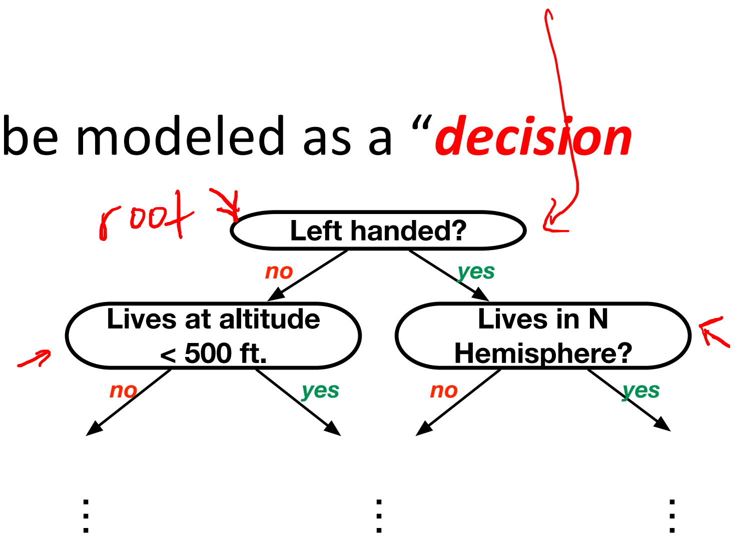
- The game:
  - the ***chooser*** picks a living person on earth.
  - the ***guesser*** asks the chooser up to 30 yes/no questions and then guesses the person.
- Is it possible for the ***guesser*** to always win this game?
- As always: the tricky part is identifying the pigeons and pigeonholes!

The Game of 30 Questions.

- Suppose that ***guesser*** has a winning strategy. It can be modeled as a "***decision tree***".
  - Begin at the root.
  - Repeat 30 times:
    - Ask the question at the current node.
    - If the answer is ***no***, move to its left child.
    - If the answer is ***yes***, move to its right child.
  - Make a guess depending on your final leaf node.

Q1

Q2

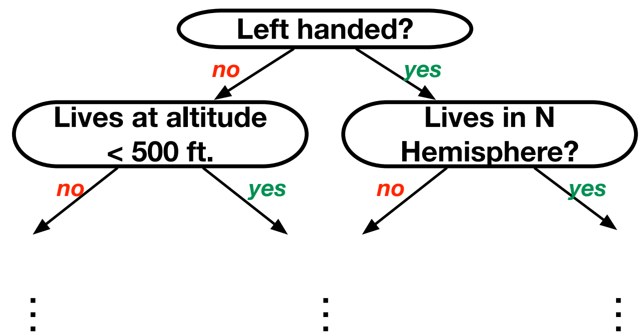


The Game of 30 Questions.

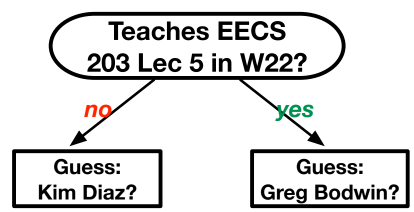
- Suppose that ***guesser*** has a winning strategy. It can be modeled as a "***decision tree***".
  - Begin at the root.
  - Repeat 30 times:
    - Ask the question at the current node.
    - If the answer is ***no***, move to its left child.
    - If the answer is ***yes***, move to its right child.
  - Make a guess depending on your final leaf node.

Q1

Q2



Q30



  - How many final leaf-nodes are there?

$2^{30} \approx 1.07$ billion final nodes

But over 7 billion people on earth.

The Game of 30 Questions.

- **Lemma.** ***Guesser*** does not have a winning strategy.
- **Proof.**
  - By contradiction. Suppose that ***guesser*** has a winning strategy. It can be written down as a decision tree T.
  - This decision tree T has $N = 2^{30} \approx 1.073$ billion final leaf nodes.
  - There are $> 7.5 \text{ billion} \geq N + 1$ people on earth.
  - Since the decision tree always wins the game, everyone on earth is mapped to *some* leaf node.
  - By the P.H.P., 2 people get mapped to the same leaf node.
  - But each leaf node can only make a ***single guess***, a contradiction. Hence ***guesser*** does ***not*** have a winning strategy.

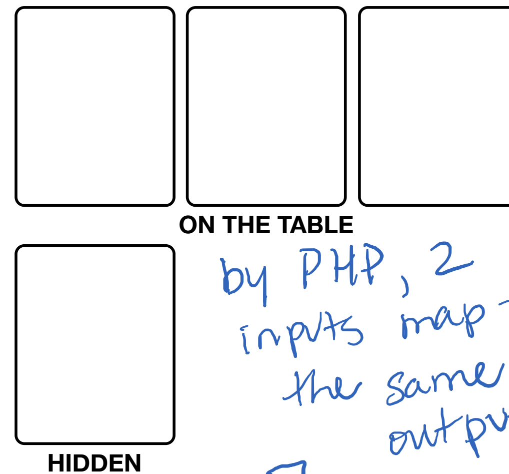
Outline

- The card trick
- Lunch orders
- Sets of numbers
- Handshakes
- The game of 30 questions
- **The card trick revisited**

Back to the card trick…

Handout

- Could you do it with 4 cards instead of 5 cards?
- **<u>Claim.</u>** The trick can't be done with 4 cards.
- The "input" is a ***set*** of 4 cards.
  - Let X = the number of 4-card hands.
- The "output" is a ***sequence*** of 3 cards.
  - Let Y = the number of 3-card sequences.
- What would $X > Y$ imply?



by PHP, 2 inputs map to the same output

⇒ more unique inputs than unique outputs

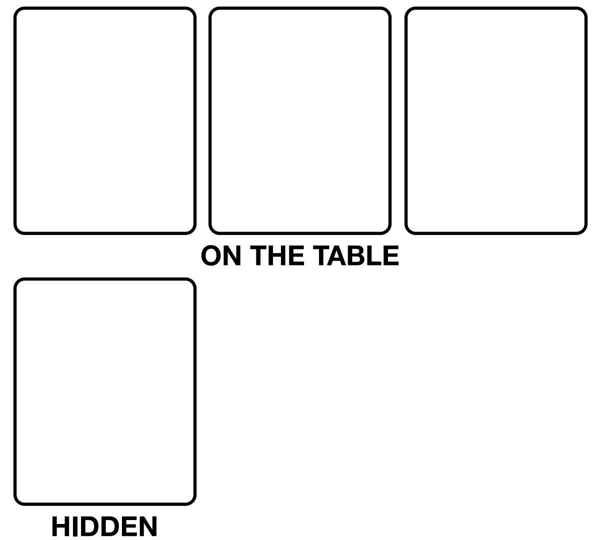
Person B can't look at the output and determine the input.

- $Y = 52 \cdot 51 \cdot 50$
- $X = 52 \cdot 51 \cdot 50 \cdot 49 / 4! = $ 52 choose 4

4! arrangements

Back to the card trick…

- Could you do it with 4 cards instead of 5 cards?
- **Claim.** The trick can't be done with 4 cards.
- The "input" is a ***set*** of 4 cards.
  - Let X = the number of 4-card hands.
- The "output" is a ***sequence*** of 3 cards.
  - Let Y = the number of 3-card sequences.
- (1) We'll calculate X and Y (in a moment)
- (2) What would $X > Y$ imply?
  - By the P.H.P.: there must be *two 4-card hands* that result in showing the *same 3-card sequence*. You can't guess the hidden card in both situations!



Back to the card trick…

- **Claim.** The trick can't be done with 4 cards.
- **Proof.**
- Let Y = be the number of ***3-card sequences***.
  - 52 options for 1st card; 51 options for 2nd, etc.: $Y = 52 \cdot 51 \cdot 50.$
- Let X = be the number of ***4-card hands***.
  - 52 options for 1st card, 51 options, etc. *Is* $X = 52 \cdot 51 \cdot 50 \cdot 49$ *??*
  - *No!* Order doesn't matter. How many times did we over-count each hand? The 4 cards could have been picked in $4! = 4 \cdot 3 \cdot 2 \cdot 1$ ways.
  - $X = \frac{52 \cdot 51 \cdot 50 \cdot 49}{4 \cdot 3 \cdot 2 \cdot 1}$. $49 > 4!$, so $X > Y$.
- By the P.H.P., there must be 2 four-card hands that result in the same 3-card sequence. Thus, we cannot always guess the hidden card!

Additional Exercises:

Handout

1. A sock drawer contains 100 socks, each one of 7 colors. There are at least 2 of each color. How many socks do you need to draw until you can be sure of having
   a) a pair of the same color?
   b) a pair of different colors?

2. A "flush" in poker is 5 cards of the same suit. How many cards do you need to draw until a subset of 5 of these cards forms a flush?

3. Suppose you pick 5 numbers in the set {1,2,3,4,5,6,7,8}. Show that two of them sum to 9.
   Hint: you need 4 pigeonholes.

4. Suppose you place 5 points in the unit square $[0,1]^2$. Prove that at least 2 of them are at distance at most $\sqrt{1/2}$.
   Hint: Divide the square into 4 pieces.

Bonus Material

- **Erdos-Szekeres Theorem (Increasing and decreasing subsequences)**

Increasing/Decreasing Subsequences

[Erdos-Szekeres theorem]

- Claim. Every sequence of $n^2 + 1$ distinct integers must contain ***<u>either</u>*** an ***increasing*** or ***decreasing*** subsequence of length $n + 1$.
- Example. Pick $n = 3$. Here is a sequence of length $3^2 + 1 = 10$:

$$5 \quad 0 \quad 4 \quad \boxed{9} \quad \boxed{8} \quad 1 \quad \boxed{3} \quad 7 \quad \boxed{2} \quad 6$$

<u>Note:</u> subsequences don't have to be contiguous!

Increasing/Decreasing Subsequences

[Erdos-Szekeres theorem]

- Claim. Every sequence of $n^2 + 1$ distinct numbers must contain ***either*** an ***increasing*** or ***decreasing*** subsequence of length $n + 1$.
- **Proof.**
  - By contradiction. Suppose there is a sequence $a_1, a_2, \ldots, a_{n^2+1}$ such that every incr./decr. subsequence of it has length at most n.
  - Label each number a_j with a pair of numbers (i_j, d_j) where:
    - i_j : the length of the longest increasing subsequence *that ends at a_j.*
    - d_j : the length of the longest decreasing subsequence *that ends at a_j.*
  - What are the ranges of i_j, d_j? How many possible labels are there?
    - By assumption: $1 \le i_j \le n$ and $1 \le d_j \le n$, so there are at most n^2 labels.
  - By the P.H.P.: there must be two numbers, say j, k, with the same labels!
    - Suppose $j < k$ and $(i_j, d_j) = (i_k, d_k)$.

Increasing/Decreasing Subsequences

[Erdos-Szekeres theorem]

- Claim. Every sequence of $n^2 + 1$ distinct numbers must contain ***<u>either</u>*** an ***increasing*** or ***decreasing*** subsequence of length $n + 1$.
- **Proof continued.**
  - There are indices $j < k$ with the same label $(i_j, d_j) = (i_k, d_k)$.

    $a_1 \quad \ldots \quad a_j \quad \ldots \quad a_k \quad \ldots \quad a_{n^2+1}.$
  - Case 1: $a_j < a_k$.
    - By definition of i_j, there is an increasing subsequence of length i_j that ends at a_j. Append a_k to this sequence and you have an increasing subsequence of length $i_j + 1 > i_k$ ending at a_k, contradicting the definition of i_k.
  - Case 2: $a_j > a_k$.
    - By definition of d_j, there is a decreasing subsequence of length d_j that ends at a_j. Append a_k to this sequence and you have a decreasing subsequence of length $d_j + 1 > d_k$ ending at a_k, contradicting the definition of d_k.
- Both case 1 and case 2 lead to a contradiction!