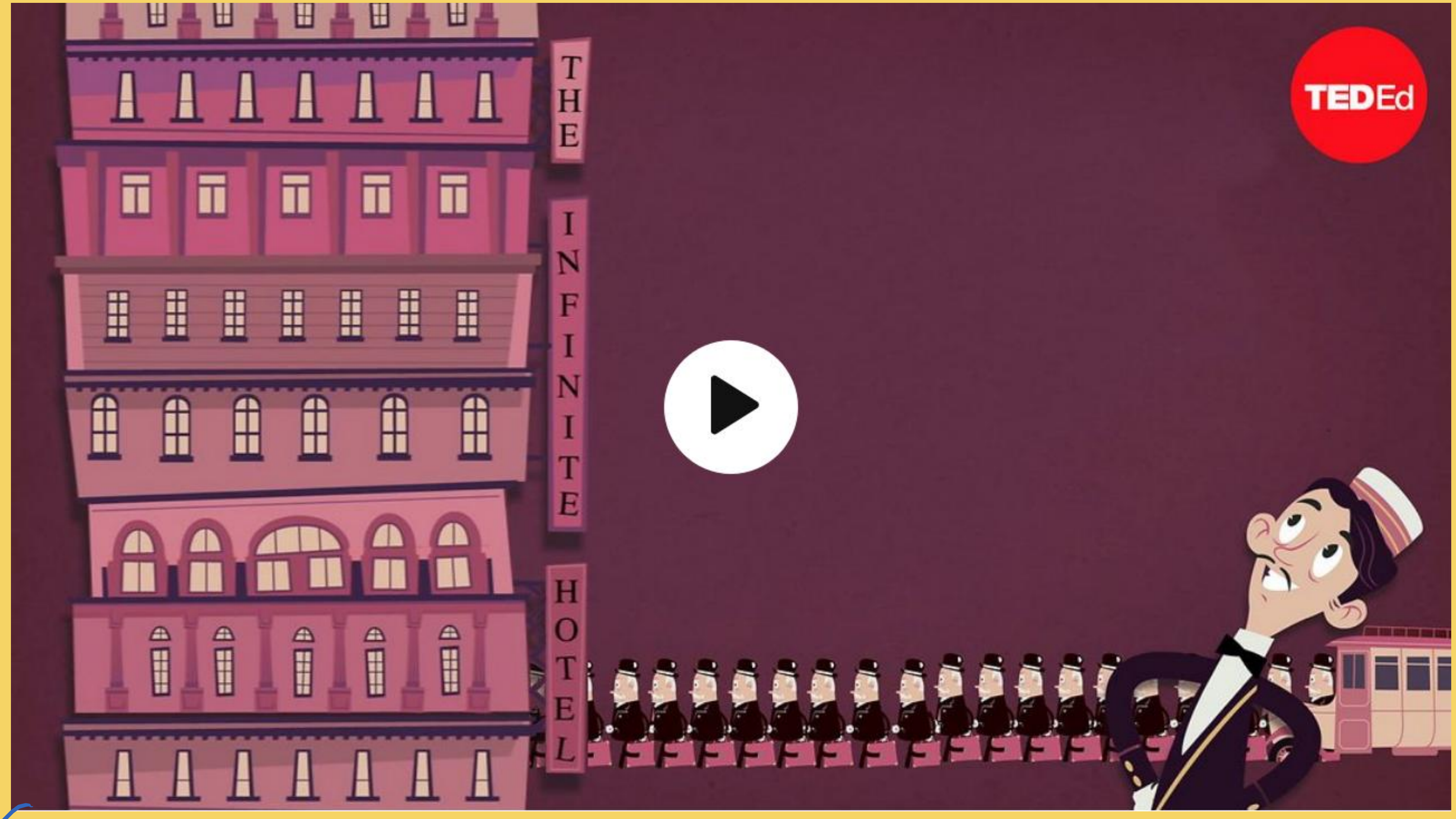
HW 0 is posted with online autograder, due next Friday, Jan 23 (10:00pm)

Quiz 1 will be posted in 24 hours, due next Mon, Jan 26 (10:00pm)



https://www.ted.com/talks/jeff_dekofsky_the_infinite_hotel_paradox

watch it

Class 2: *What Can Be Represented by Bits?*

University of Virginia
CS3120: DMT2
https://weikailin.github.io/cs3120-toc
Wei-Kai Lin

Quiz 0: Some questions

free drink
food

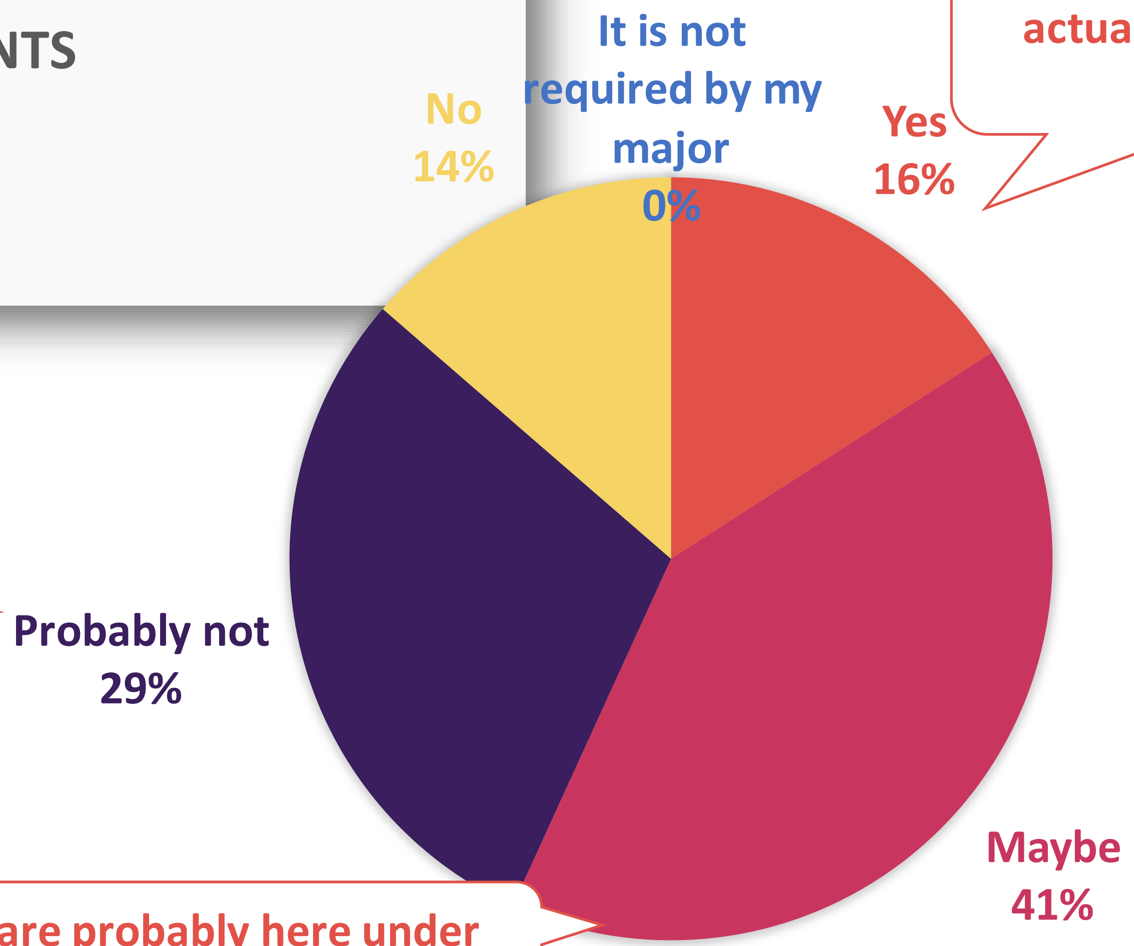
- Talk about yourself a bit more
  - Undergrad major: chemistry. Master: a CS group in EE. (Taiwan)
  - PhD: CS at Cornell, focused on theoretical cryptography and algo. rithms
- Can we work ahead?
  - Yes
- What’s the AI policy, not to get solutions, but to get better understanding?
  - There is no restriction.
    Eg, “What are the meanings or purposes of Babylonian clay tablets?”
    Eg, “Find a practice question for constructive definitions.”

practice for this course

Q2.2 Would you be taking this class if it was not required for your major (or desired major)?

0 Points

- It is not required by my major
- Yes
- Maybe
- Probably not
- No



Maybe 16% people actually want to be in this class!

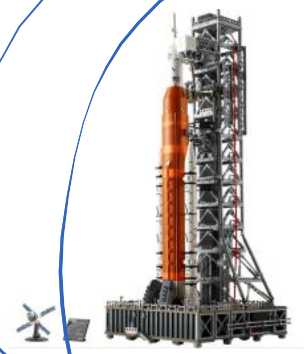
43% are here under duress!

84% are probably here under duress!

Why study theory?

My Answer

It is fun! Like LEGO

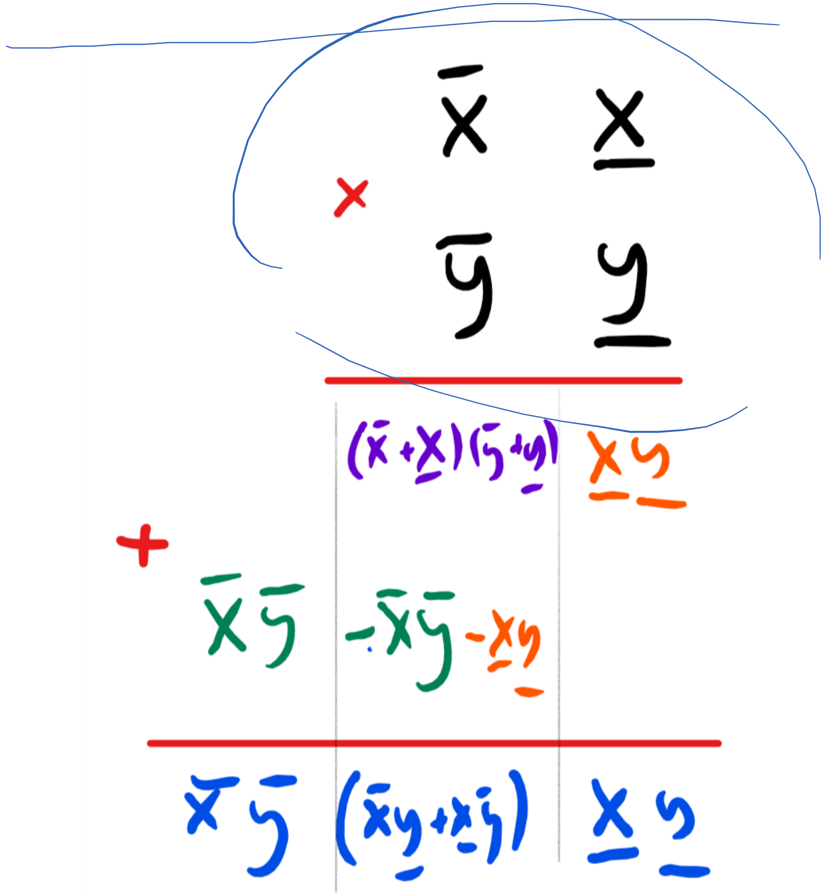


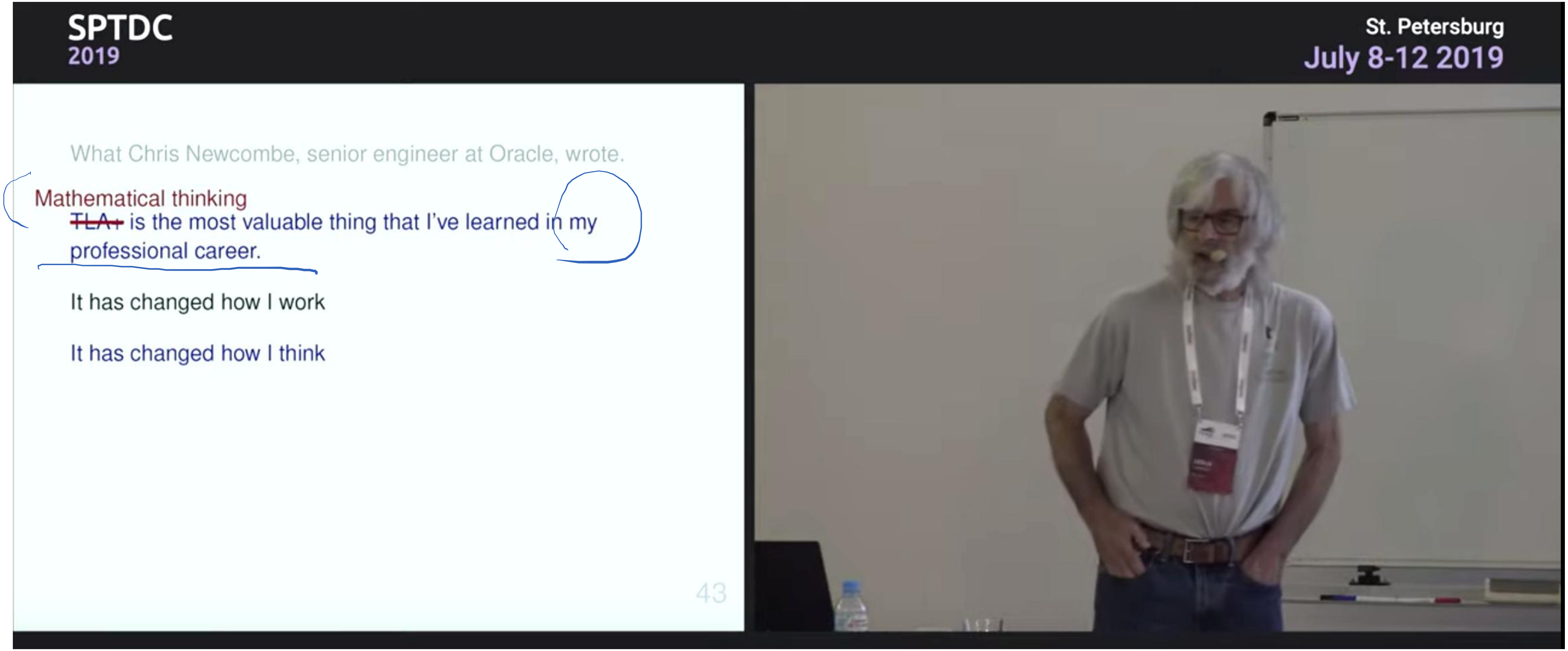
E.g., Karatsuba built the same multiply but taking fewer bricks

(Maybe) Your Answer

1. Think **Better** and **Bigger**

Karatsuba's multiplication





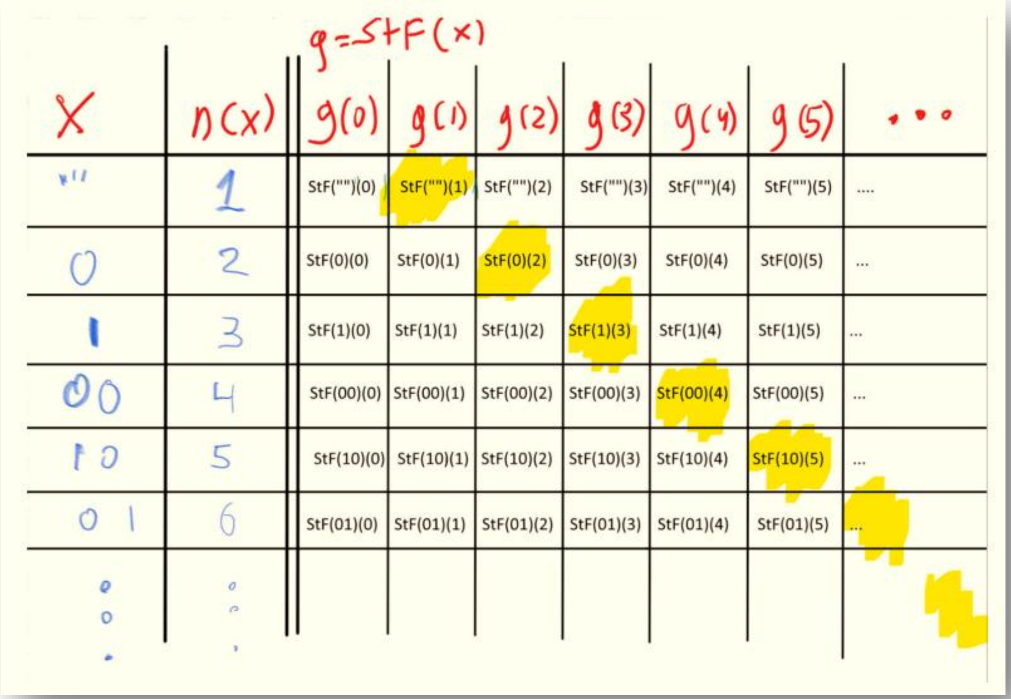
https://lamport.azurewebsites.net/tla/paxos-algorithm.html
https://www.youtube.com/watch?v=tw3gsBms-f8

Leslie Lamport
(2013 Turing Awardee)

(Maybe) Your Answer

1. Think **Better** and **Bigger**

2. Intrinsic **Truth** and **Beauty**



$g = StF(x)$

x	n(x)	g(0)	g(1)	g(2)	g(3)	g(4)	g(5)	...
""	1	StF("")(0)	StF("")(1)	StF("")(2)	StF("")(3)	StF("")(4)	StF("")(5)	....
0	2	StF(0)(0)	StF(0)(1)	StF(0)(2)	StF(0)(3)	StF(0)(4)	StF(0)(5)	...
1	3	StF(1)(0)	StF(1)(1)	StF(1)(2)	StF(1)(3)	StF(1)(4)	StF(1)(5)	...
00	4	StF(00)(0)	StF(00)(1)	StF(00)(2)	StF(00)(3)	StF(00)(4)	StF(00)(5)	...
10	5	StF(10)(0)	StF(10)(1)	StF(10)(2)	StF(10)(3)	StF(10)(4)	StF(10)(5)	...
01	6	StF(01)(0)	StF(01)(1)	StF(01)(2)	StF(01)(3)	StF(01)(4)	StF(01)(5)	...
⋮	⋮							

Cantor's Diagonalization Proof
(Chapter 2 and upcoming class)

(Maybe) Your Answer

1. Think **Better** and **Bigger**

2. Intrinsic **Truth** and **Beauty**

3. Track record for **Impact** and **Success**

Definition 1 *Let $E(u)$ be some vector over the Web pages that corresponds to a source of rank. Then, the PageRank of a set of Web pages is an assignment, R', to the Web pages which satisfies*

$$R'(u) = c \sum_{v \in B_u} \frac{R'(v)}{N_v} + cE(u) \tag{1}$$

such that c is maximized and $||R'||_1 = 1$ ($||R'||_1$ denotes the L_1 norm of R').

Definition 1 *Let $E(u)$ be some vector over the Web pages that corresponds to a source of rank. Then, the PageRank of a set of Web pages is an assignment, R', to the Web pages which satisfies*

$$R'(u) = c \sum_{v \in B_u} \frac{R'(v)}{N_v} + cE(u) \tag{1}$$

such that c is maximized and $||R'||_1 = 1$ ($||R'||_1$ denotes the L_1 norm of R').

Google

The PageRank Citation Ranking: Bringing Order to the Web

Lawrence Page, Sergey Brin,
Rajeev Motwani, Terry Winograd

January 29, 1998

Definition 1 *Let $E(u)$ be some vector over the Web pages that corresponds to a source of rank. Then, the PageRank of a set of Web pages is an assignment, R', to the Web pages which satisfies*

$$R'(u) = c \sum_{v \in B_u} \frac{R'(v)}{N_v} + cE(u) \tag{1}$$

such that c is maximized and $||R'||_1 = 1$ ($||R'||_1$ denotes the L_1 norm of R').

No theorems or proofs in this paper! (but makes claims about convergence)

6.1 Manipulation by Commercial Interests

These types of personalized PageRanks are virtually immune to manipulation by commercial interests. For a page to get a high PageRank, it must convince an important page, or a lot of non-important pages to link to it. At worst, you can have manipulation in the form of buying advertisements(links) on important sites. But, this seems well under control since it costs money. This immunity to manipulation is an extremely important property. This kind of commercial manipulation is causing search engines a great deal of trouble, and making features that would be great to have very difficult to implement. For example fast updating of documents is a very desirable feature, but it is abused by people who want to manipulate the results of the search engine.

Dynamic Itemset Counting and Implication Rules for Market Basket Data

Sergey Brin * Rajeev Motwani Jeffrey D. Ullman
Department of Computer Science
Stanford University
{sergey,rajeev,ullman}@cs.stanford.edu

Shalom Tsur
R&D Division, Hitachi America Ltd.
tsur@hitachi.com

Copy Detection Mechanisms for Digital Documents *

Sergey Brin, James Davis, Hector Garcia-Molina
Department of Computer Science
Stanford University
Stanford, CA 94305-2140
e-mail: sergey@cs.stanford.edu

Near Neighbor Search in Large Metric Spaces

Sergey Brin*
Department of Computer Science
Stanford University
sergey@cs.stanford.edu

Discrete Mathematics 27 (1979) 47–57.

BOUNDS FOR SORTING BY PREFIX REVERSAL

Bill

William H. GATES

Microsoft, Albuquerque, New Mexico

Christos H. PAPADIMITRIOU*†

Department of Electrical Engineering, University of California, Berkeley, CA 94720, U.S.A.

Received 18 January 1978
Revised 28 August 1978

For a permutation σ of the integers from 1 to n, let $f(\sigma)$ be the smallest number of prefix reversals that will transform σ to the identity permutation, and let $f(n)$ be the largest such $f(\sigma)$ for all σ in (the symmetric group) S_n. We show that $f(n) \leqslant (5n+5)/3$, and that $f(n) \geqslant 17n/16$ for n a multiple of 16. If, furthermore, each integer is required to participate in an even number of reversed prefixes, the corresponding function $g(n)$ is shown to obey $3n/2-1 \leqslant g(n) \leqslant 2n+3$.

Theorem 1. *Algorithm $\mathscr{A}$ creates a permutation with $n-1$ adjacencies by at most $(5n-7)/3$ moves.*

Proof. First, it is clear that if we have a permutation σ with less than $n-1$ adjacencies, one of the cases 1 through 7 is applicable. Hence, the algorithm does not halt unless $n-1$ adjacencies have been created. Obviously the algorithm will eventually halt, since at each execution of the main loop at least one new

Theorem 2. $19n/16 \geq f(\chi) \geq 17n/16$.

Proof. To show the upper bound, we first do the following sequences of moves

$$\chi \to \tau_2\tau_1^R\tau_3 \cdots \to \tau_2^R\tau_1^R\tau_3 \cdots \to \tau_1\tau_2\tau_3 \cdots$$

and so on, bringing the even-indexed τ's in front and then back with the reversal cancelled in three moves. Thus, in $3n/16$ moves we obtain $\chi' = \tau_1\tau_2\tau_3 \cdots \tau_m$. Then, for each copy of τ in χ' we repeat the following sequence of eight moves (among a number of possibilities)

Lemma 1. *Let A_j be an (a, b) cut.*
(1) It $a, b > 1$, then both A_j, A_{j+1} are wastes,
(2) If the only one of $a, b > 1$, then either A_j or A_{j+1} is a waste.

Proof. An easy case-by-case analysis.

Theorem 3. $g(n) \leq 2n+3$.

Proof. First observe that $g(\sigma)$ is not greater than $f(\sigma')$ where $\sigma' \in S_{2n}$ is defined as follows, for each $\sigma \in S_n : \sigma'(2i-1) = 2\sigma(i) - 1$ and $\sigma'(2i) = 2\sigma(i)$ for all $i = 1, \ldots, n$. The complexity of sorting σ' without the restriction can now be bounded from above by the algorithm $\mathscr{A}$ of Section 2. The equations governing the complexity of $\mathscr{A}$ when applied to σ' are (1), (2), and (3) of Section 2 with n replaced by $2n$, $b = a = n$, and also noting that only x_5 can be nonzero, since all other actions are possible only in the presence of free elements. The maximum is therefore $2n-2$. Allowing five more moves to sort the resulting permutation, we get the claimed bound.

Claim 1. *There are exactly m events.*

To prove Claim 1, we notice that χ_0 is k-stable for $k = 1, \ldots, m$, and $\chi_{f(\chi)}$ is not k-stable for any k. Furthermore, no permutation can cease being k_1-stable *and* k_2-stable, $k_1 \neq k_2$, in only one move.

Theorem 4. $g(e_n^R) \geq \frac{3}{2}n - 1$.

Claim 3. *For all j, $1 \leq j \leq m-1$, there exists a waste χ_l with $i_j \leq l \leq i_{j+1}$.*

To prove Claim 3, suppose that it fails. In other words, suppose that there is an event i_j other than the last one, such that all moves χ_l, $i_j \leq l \leq i_{j+1}$ construct a new adjacency without destroying an existing adjacency. Suppose that k is the

Discrete Mathematics 27 (1979) 47–57.

BOUNDS FOR SORTING

William H. GATES
Microsoft, Albuquerque, New Mexico

Christos H. PAPADIMITRIOU*
Department of Electrical Engineering, U

Received 18 January 1978
Revised 28 August 1978

For a permutation σ of the integers fr
reversals that will transform σ to the ident
for all σ in (the symmetric group) S_n. We show that $f(n) \leq (5n+5)/3$, and that $f(n) \geq 17n/16$ for n a multiple of 16. If, furthermore, each integer is required to participate in an even number of reversed prefixes, the corresponding function $g(n)$ is shown to obey $3n/2-1 \leq g(n) \leq 2n+3$.

"Two years later, I called to tell him our paper had been accepted to a fine math journal. He sounded eminently disinterested. He had moved to Albuquerque, New Mexico to run a small company writing code for microprocessors, of all things. I remember thinking: *`Such a brilliant kid. What a waste.'"*

BOUNDS FOR SORTING BY PREFIX REVERSAL

William H. GATES
Microsoft, Albuquerque, New Mexico

1. Introduction

We introduce our problem by the following quotation from [1]

> The chef in our place is sloppy, and when he prepares a stack of pancakes they come out all different sizes. Therefore, when I deliver them to a customer, on the way to the table I rearrange them (so that the smallest winds up on top, and so on, down to the largest at the bottom) by grabbing several from the top and flipping them over, repeating this (varying the number I flip) as many times as necessary. If there are n pancakes, what is the maximum number of flips (as a function $f(n)$ of n) that I will ever have to use to rearrange them?

largest at the bottom) by
this (varying the number
is the maximum number
rearrange them?

References

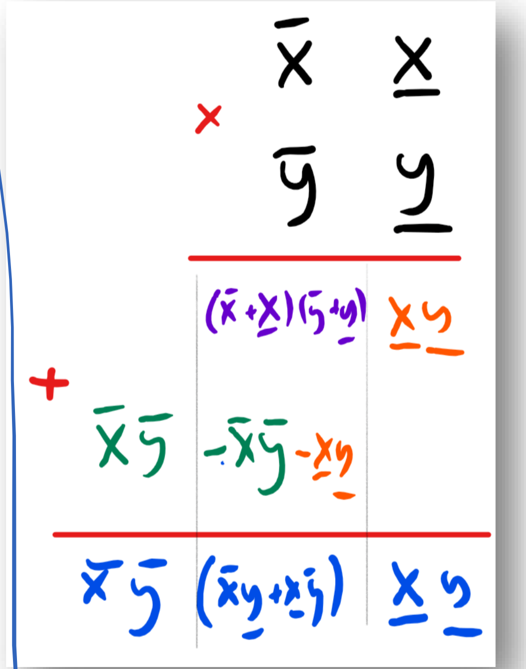
[1] Amer. Math. Monthly 82 (1) (1975) 1010.

(Maybe) Your Answer

1. Think **Better** and **Bigger**
2. Intrinsic **Truth** and **Beauty**
3. Track record for **Impact** and **Success**
4. Being **wrong** often and **admitting** it

Kolmogorov

Karatsuba



Plan

Represent Binary Strings

And more

Cardinality of Sets

Countable

Cantor’s Theorem

Today: Chapter 2 in the TCS book

https://introtcs.org/public/lec_02_representation.html#cantorsec

Recap: Natural Number, Binary String

Constructive defn

Rule 1: $\mathbf{0}$ is a Natural Number
Rule 2: If n is a Natural Number, $\boldsymbol{S}(n)$ is a Natural Number.

Rule 1: "" is a Binary String
Rule 2: If ***s*** is a Binary String, both **0*s*** and **1*s*** are Binary Strings.

Notations

Natural Number, $\mathbb{N}$.

Binary String, $\{\mathbf{0},\mathbf{1}\}^*$

Recap: Set

Set Builder Notation

A set can be defined based on any predicate using **set-builder notation**: $\{f(x) \mid P(x)\}$ is the set of all $f(x)$ where $P(x)$ is true.

Example —

$\{x+2 \mid 1 < x \leq 2\}$ is the set of all numbers greater than 3 and no greater than 4.

$x \in \{x \mid P(x)\}$ is a long way of writing $P(x)$.

$\{x \mid x \in S\}$ is a long way of writing S.

Eg, even natural numbers, $\{x \in \mathbb{N} \mid x/2 = 0\}$

Eg, primes, $\{x \in \mathbb{N} \mid x \neq 0 \wedge x \neq 1 \wedge \forall y < x, y \in \mathbb{N}, y > 1\}$

$x \% y \neq 0$

2, 3, 5, 7, ...

Can all Binary Strings be represented by Natural Numbers?

Rule 1: “” is a Binary String
Rule 2: If ***s*** is a Binary String, both **0*s*** and **1*s*** are Binary Strings.

Rule 1: $\mathbf{0}$ is a Natural Number
Rule 2: If n is a Natural Number, $\boldsymbol{S}(n)$ is a Natural Number.

Want: each ***s*** is *mapped* to a unique n

“”	0
0	1
1	2
00	3
01	4
10	5
11	6
000	7
001	8
010	9
…	…

Product Set (notation)

Definition: Any sets S, T, the product set $S \times T$ is defined to be

$$S \times T := \{(a, b) \mid a \in S, b \in \mathrm{T}\}.$$

Syntax: Any set S, let $S^1 := S$, and
let $S^{n+1} := S \times S^n$ for all $n \in \mathbb{N}, n > 0$.

$S^1, S^2, S^3, \ldots$ $\cancel{S \times S}$

$S \times S, S \times S \times S, \ldots$

$\{\mathbf{0}, \mathbf{1}\}^*$

Rule 1: “” is a Binary String
Rule 2: If ***s*** is a Binary String, both **0*s*** and **1*s*** are Binary Strings.

Definition: Any *finite* set S,
let $S^* := \{a \in S^n \mid n \in \mathbb{N}, n > 0\}$.

Another definition: Any *finite* set S,
let $S^* := S^0 \cup S^1 \cup S^2 \cup \cdots = \bigcup_{n \in \mathbb{N}} S^n$

Called *Kleene Star.*

$\{\mathbf{0}, \mathbf{1}\}^*$

Rule 1: “” is a Binary String
Rule 2: If ***s*** is a Binary String, both **0*s*** and **1*s*** are Binary Strings.

Is each of the following in $\{0,1\}^*$?

- “” Yes
- “010111” Yes
- 010001100010010011111 YES
- 111111111… NO
- Is “111111111…” a Binary String?

Constructive definitions are **finite.** We can produce a binary string of *any* length from these rules but cannot produce an infinite binary string.

$\{0,1\}^*$ vs. $\{0,1\}^\infty$

Definition 2.7

We denote by $\{0,1\}^\infty$ the set $\{f \mid f : \mathbb{N} \to \{0,1\}\}$.

$\{0,1\}^*$ finite binary strings of any length

$\{0,1\}^\infty$ "infinite binary strings"

$f = 1\ 1\ 1\ 1\ 1\ 1 \ldots$

$\mathbb{N} = 0, 1, 2, 3, \ldots\ 100, \ldots\ 1000$

$f[1000]$

A set can be defined based on any predicate using **set-builder notation**: $\{f(x) \mid P(x)\}$ is the set of all $f(x)$ where $P(x)$ is true.

DMT1-OTB

this course

rare

$\{\mathbf{0},\mathbf{1}\}^*$ vs. $\{\mathbf{0},\mathbf{1}\}^\infty$

Definition 2.7

We denote by $\{0,1\}^\infty$ the set $\{f \mid \underbrace{f:\mathbb{N}\to\{0,1\}}_{\text{often}}\}$.

$\{\mathbf{0},\mathbf{1}\}^*$ finite binary strings of any length

$\{\mathbf{0},\mathbf{1}\}^\infty$ "infinite binary strings"

Why we need infinite binary strings, $\{\mathbf{0},\mathbf{1}\}^\infty$?

Cardinality of Sets

Cantor

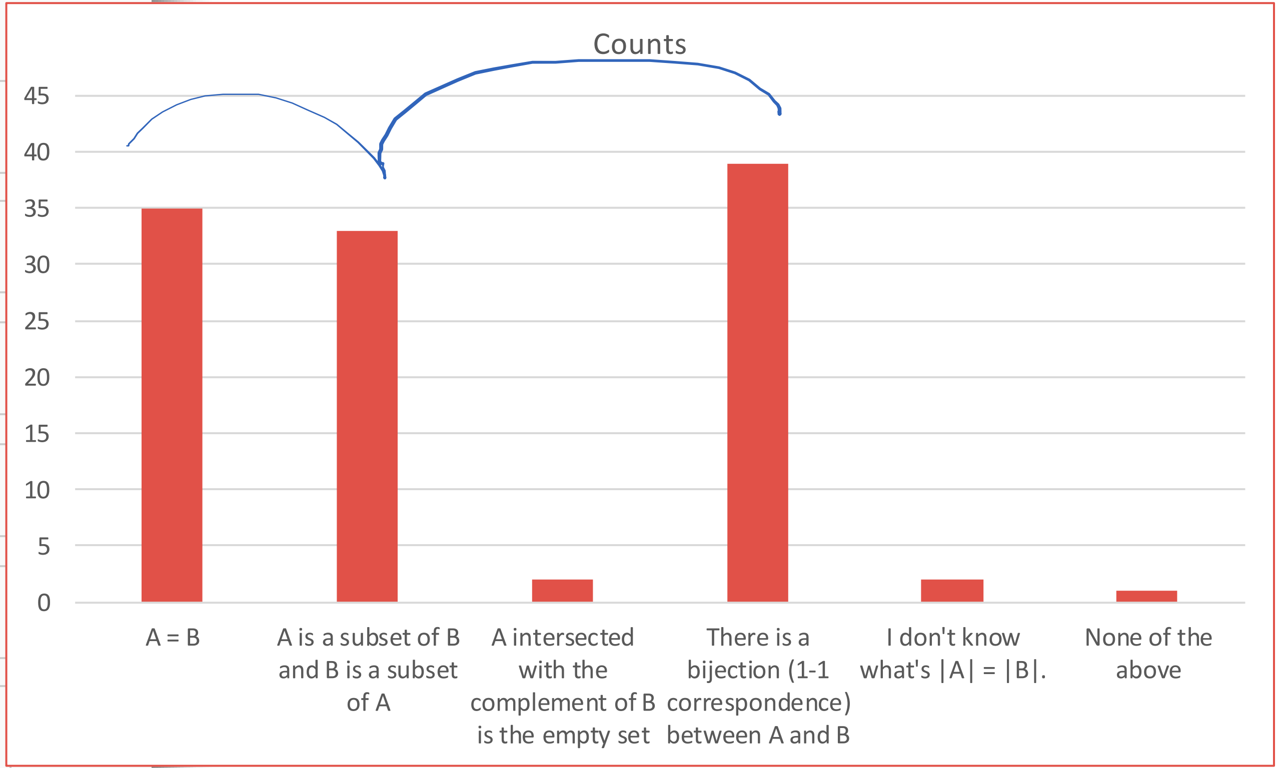
$$|S|$$

Q3.4 Sets

1 Point

For two sets A and B, which of the following implies $\boxed{|A| = |B|}$? (select all that apply)

- [ ] A = B
- [ ] A is a subset of B and B is a subset of A
- [ ] A intersected with the complement of B is the empty set
- [ ] There is a bijection (1-1 correspondence) between A and B
- [ ] I don't know what's |A| = |B|.
- [ ] None of the above.



Cardinality of Finite Sets

Definition. Two sets have the *same cardinality* if there is a bijection between the two sets.

Another

The ***cardinality*** of the set

$$[k] := \{ n \mid n \in \mathbb{N} \wedge n < k \}$$

is k. for all $k \in \mathbb{N}$

Example: $S = \{C, S, 3120\}$, $|S| = ?$ 3

$S \xleftrightarrow{\text{biject}} [3]$

$|S| = |[3]| = 3$

Cardinality of Infinite Sets

Definition. Two sets have the *same cardinality* if there is a bijection between the two sets.

The ***cardinality*** of the set

$$[k] = \{\, n \mid n \in \mathbb{N} \wedge n < k \,\}$$

is k.

$$|\{0,1\}^*| = ?$$

$\{0,1\}^* \leftrightarrow [k]$?

$\exists s \in$ bijct$(s) \notin [k]$

Cardinality of Infinite Sets

Definition. Two sets have t… …*rdinality* if there is a bijection b… … two sets.

The ***cardin…*** … set

$$\ldots \mid n \in \mathbb{N} \wedge n < k \}$$

is k.

A set S is infinite, if there is no bijection between S and any $[k]$.

Cardinality of (Infinite) Sets

Definition. Two sets have the *same cardinality* if there is a bijection between the two sets.

If $A \subsetneq B$, then $|A|$ $\overset{?}{<}$ $|B|$

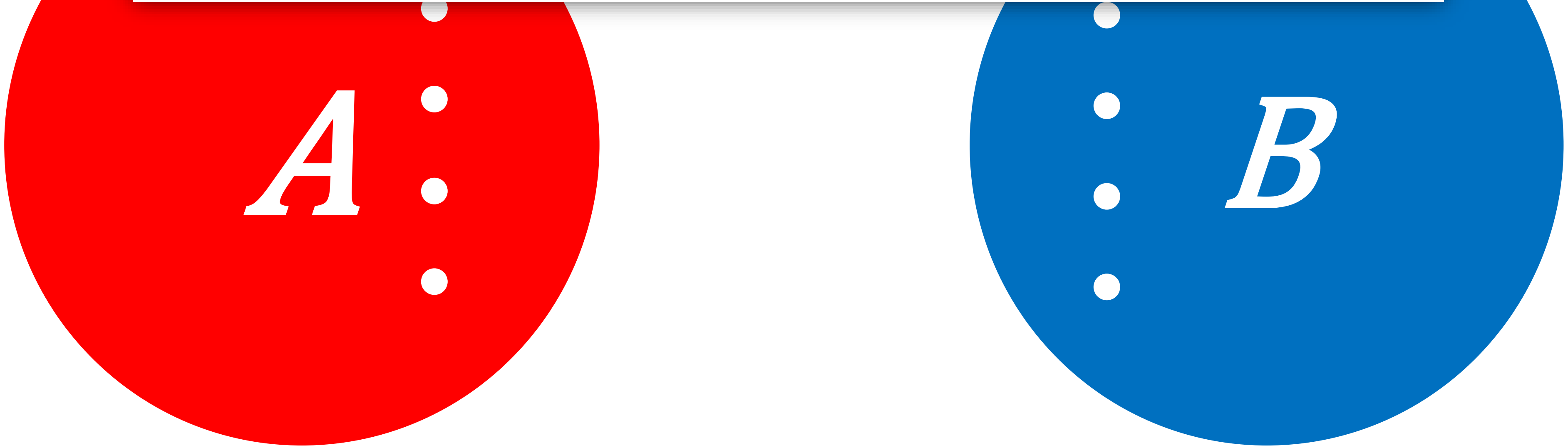
$\exists x \in B,\ x \notin A$ $\neq$?

Function (DMT1 Review)

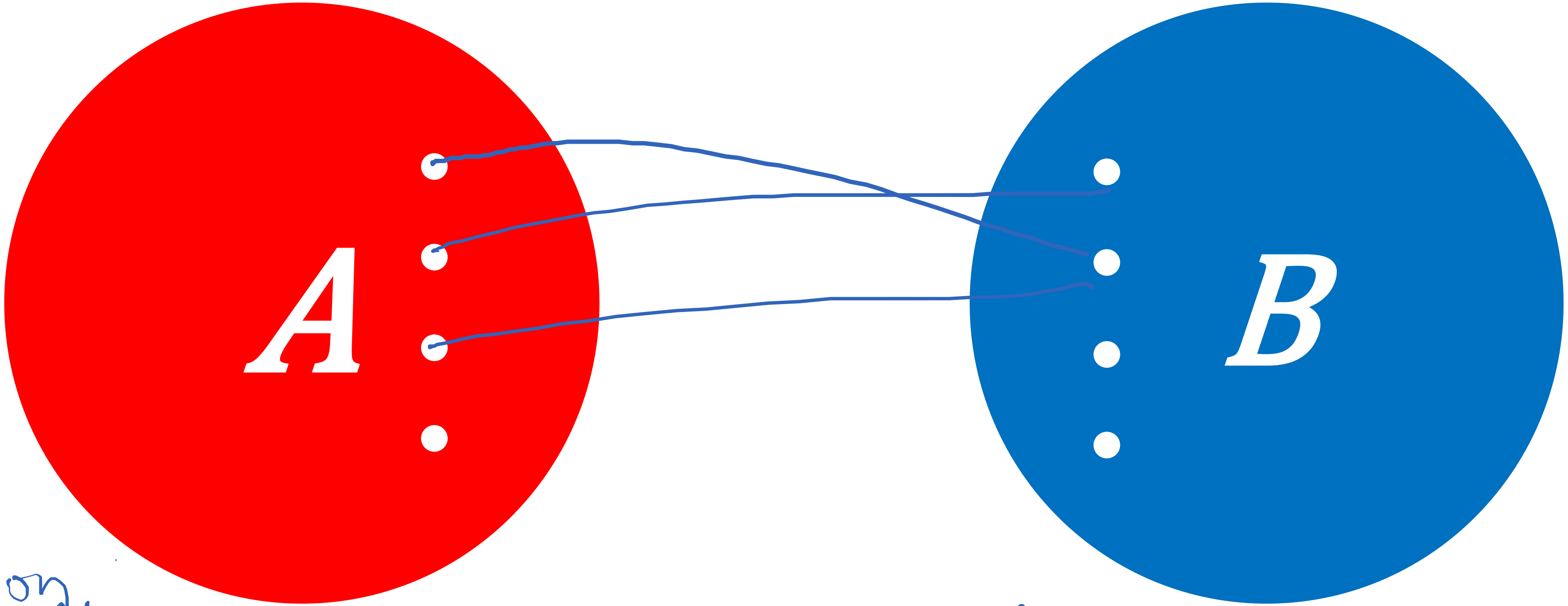
DMT1-OTB

A *function* $f : D \to C$ is

- **total** if $f(x)$ is defined for all $x \in D$.
- **partial** means “either total or not total” and is used in contexts where most functions are assumed to be total to identify subcontexts where that assumption does not apply.
- **injective** or **1-to-1** if $f(x) = f(y)$ implies $x = y$.
- **surjective** or **onto** if the *co-domain* is equal to the *range*.
- **bijective**, **invertible**, or a **1-to-1 correspondence** if it is *total* and *injective* and *surjective*.



Function Properties



function is total

function: $\leq$ 1 out

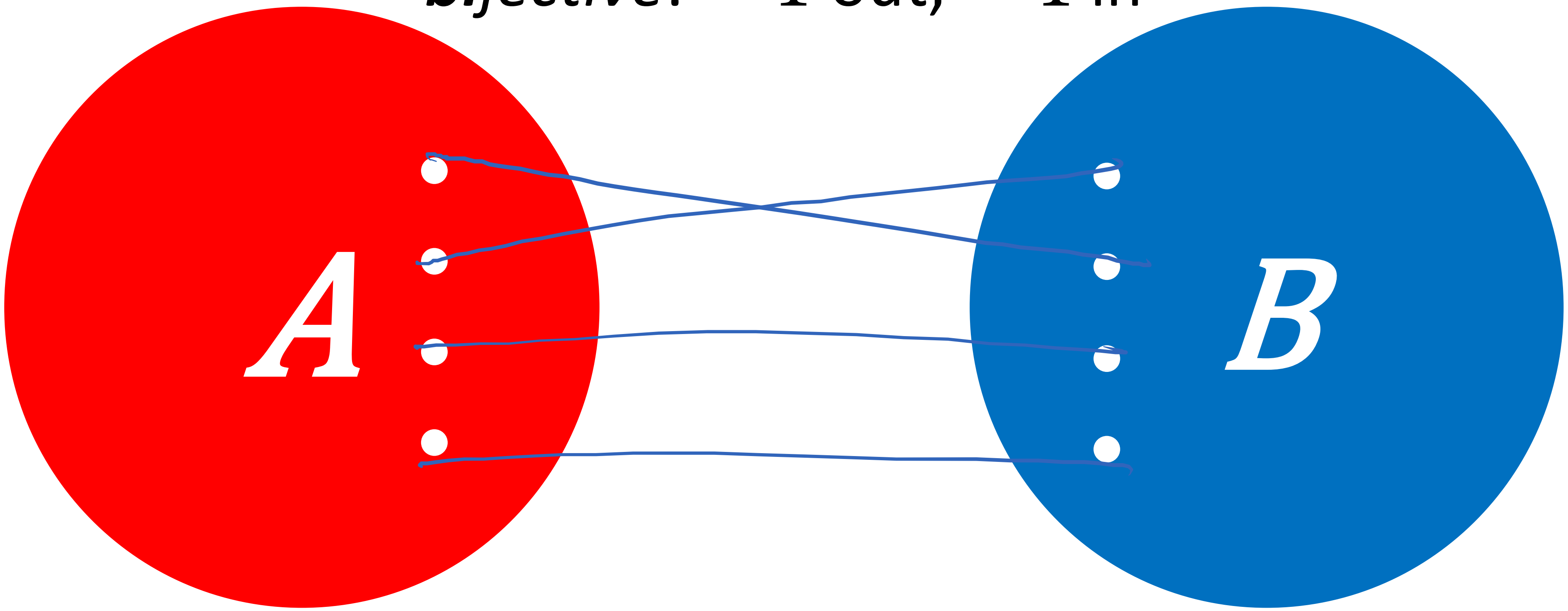
total: $\geq$ 1 out

injective: $\leq$ 1 in

surjective: $\geq$ 1 in

Function Properties

***bi**jective*: $=1$ out, $=1$ in



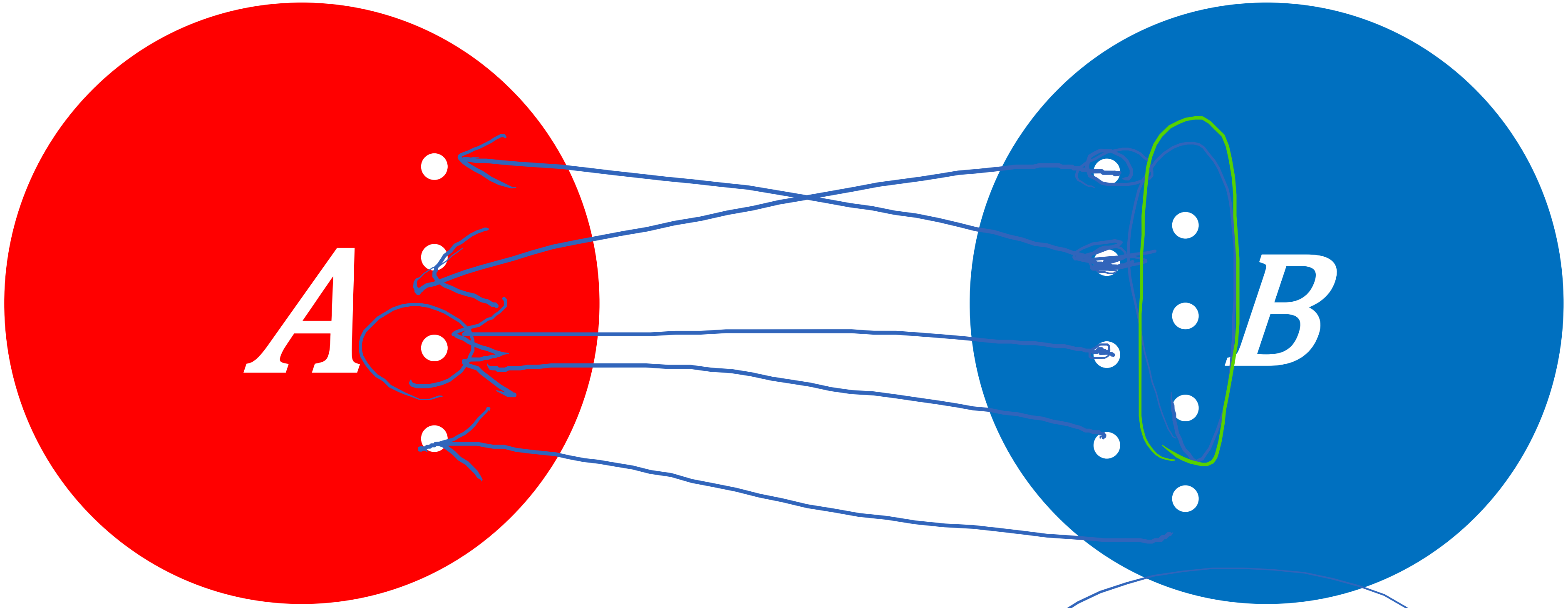
function: ≤ 1 out

total: ≥ 1 out

***in**jective*: ≤ 1 in

***sur**jective*: ≥ 1 in

Surjective function from B to A



surjective: ≥ 1 in

function: ≤ 1 out

$|Even| \leq^{g} |\mathbb{N}|$ $g(x) = 2x$ g surject $\Rightarrow$

$|Nat| \leq$ $g^{-1}(x) = h(x) = x/2$

Cardinality of (Infinite) Sets

Definition. Two sets have the *same cardinality* if there is a bijection between the two sets.

$|Even| \geq |\mathbb{N}|$

&

Definition. If there exists a **surjective function** from sets B to A, then we say the cardinality of B is ***greater than or equal to*** the cardinality of A.

We denote this as $|A| \leq |B|$.

Two Useful (Intuitively obvious?) Facts

Thms.

1. If $\underbrace{|A| = |B|}_{\text{bij}}$ then $\underbrace{|A| \leq |B|}_{\text{surj}}$ and $\underbrace{|B| \leq |A|}_{\text{surj}}$.

easy

2. If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.

Are these “obvious” facts, or do we need a proof?

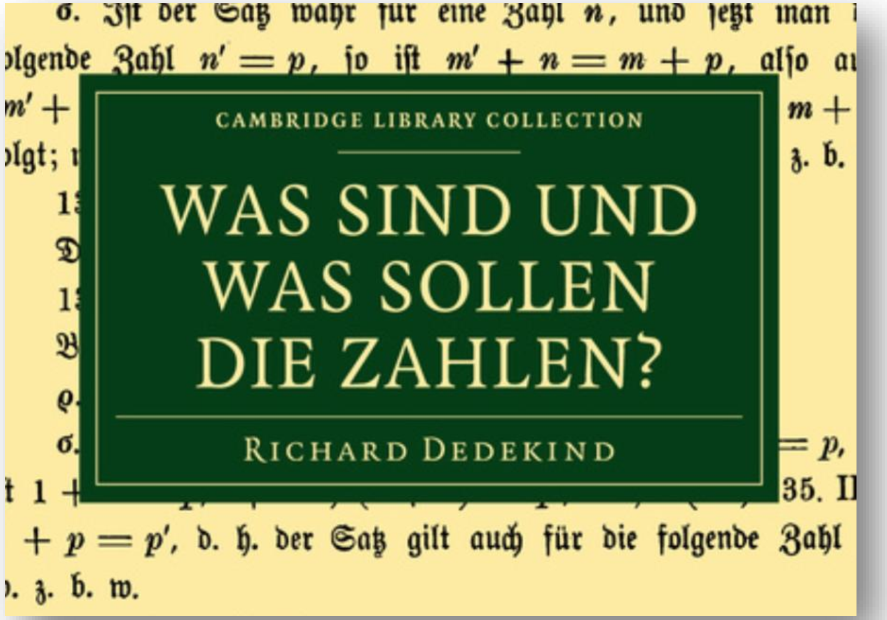
Two Useful Theorems

1. If $|A| = |B|$ then $|A| \leq |B|$ and $|B| \leq |A|$.

2. If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.

(Cantor-) Schröder–Bernstein theorem

First stated (but not proven) by Cantor (1887)
First proven (but not published) by Dedekind (1887)
First proof announced by Schroder in 1896 (but its incorrect)
First correct proof by Bernstein (1897) (student in Cantor's class)



Do infinite sets even *exist*?


Was sind und was sollen
die
Zahlen?

Von

Richard Dedekind
Professor an der technischen Hochschule zu Braunschweig

Dritte unveränderte Auflage

Braunschweig
Druck und Verlag von Friedr. Vieweg & Sohn
1911

Pronounce “Dedekind”: https://www.youtube.com/watch?v=s5pbbwGrH14

Do infinite sets even *exist*?

¶64. *Definition.* A set S is said to be *infinite* when it is similar to a proper subset of itself, otherwise it is said to be *finite.* Dedekind's footnote to this definition contains some important historical notes.

> In this form I submitted the definition of the infinite which forms the core of my whole investigation in September, 1882, to G. Cantor and several years earlier to Schwarz and Weber. All other attempts that have come to my knowledge to distinguish the infinite from the finite seem to me to have met with so little success that I think I may be permitted to forego any criticism of them.

David Joyce's *Notes on Richard Dedekind's Was sind und was sollen die Zahlen?*
https://mathcs.clarku.edu/~djoyce/numbers/dedekind.pdf

Do infinite sets even *exist*?

¶64. *Definition.* A set S is said to be *infinite* when it is similar to a proper subset of itself, otherwise it is said to be *finite.* Dedekind's footnote to this definition contains some important historical notes.

Definition. A set is ***Dedekind***-*infinite* if and only if it has the same cardinality as some **strict subset** of itself.

bij

$\subsetneq$

David Joyce's *Notes on Richard Dedekind's Was sind und was sollen die Zahlen?*
https://mathcs.clarku.edu/~djoyce/numbers/dedekind.pdf

Logistics

- Quiz 1 will be released soon (in 24 hours)
  Due next Monday, Jan 26, 10pm.
- HW 0 due this Friday, 10 pm.
  Autograder is online.

Course Outline

1. Sets and Cardinality $\neq$ infinity
2. Regular Expression and Finite Automata
3. Circuits (major diff from Prof Pettit)
4. Turing Machines and Computability
5. Complexity

Meetings

- Attend the class of Prof Pettit as you like
- HW & Exams are still here
- We have TAs (phd students) in person

will

Recap

Today: Chapter 2 in the TCS book

Represent Binary Strings

And more

Cardinality of Sets

Countable and Infinite

Power Set

Cantor's Theorem