$\binom{n}{2}$ Themes

Why discrete math and counting ?

- Makes you think in a special way
- Essential for algorithms

$$
\begin{aligned}
&= \binom{k}{2}\frac{1}{n} \\
&= \frac{k(k-1)}{2n} .
\end{aligned}
$$

When $k(k-1) \geq 2n$, therefore, the expected number of pairs of people with the same birthday is at least 1. Thus, if we have at least $\sqrt{2n}+1$ individuals in a room, we can expect at least two to have the same birthday. For $n = 365$, if $k = 28$, the expected number of pairs with the same birthday is $(28 \cdot 27)/(2 \cdot 365) \approx 1.0356$. Thus, with at least 28 people, we expect to find at least one matching pair of birthdays. On Mars, with 669 days per year, we need at least 38 Martians.

The first analysis, which used only probabilities, determined the number of people required for the probability to exceed 1/2 that a matching pair of birthdays exists, and the second analysis, which used indicator random variables, determined the number such that the expected number of matching birthdays is 1. Although the exact numbers of people differ for the two situations, they are the same asymptotically: $\Theta(\sqrt{n})$.

$1 \leq i \leq n-2$. The probability that the edge randomly chosen in the first step is in C is at most $k/(nk/2) = 2/n$, so that $\mathbf{Pr}[\mathcal{E}_1] \geq 1-2/n$. Assuming that $\mathcal{E}_1$ occurs, during the second step there are at least $k(n-1)/2$ edges, so the probability of picking an edge in C is at most $2/(n-1)$, so that $\mathbf{Pr}[\mathcal{E}_2 \mid \mathcal{E}_1] \geq 1-2/(n-1)$. At the ith step, the number of remaining vertices is $n-i+1$. The size of the min-cut is still at least k, so the graph has at least $k(n-i+1)/2$ edges remaining at this step. Thus, $\mathbf{Pr}[\mathcal{E}_i \mid \cap_{j=1}^{i-1}\mathcal{E}_j] \geq 1-2/(n-i+1)$. What is the probability that no edge of C is ever picked in the process? We invoke (1.6) to obtain

$$\mathbf{Pr}[\cap_{i=1}^{n-2}\mathcal{E}_i] \geq \prod_{i=1}^{n-2}\left(1-\frac{2}{n-i+1}\right) = \frac{2}{n(n-1)}.$$

The probability of discovering a particular min-cut (which may in fact be the unique min-cut in G) is larger than $2/n^2$. Thus our algorithm may err in declaring the cut it outputs to be a min-cut. Suppose we were to repeat the above algorithm $n^2/2$ times, making independent random choices each time. By (1.4), the probability that a min-cut is not found in any of the $n^2/2$

$$
\begin{aligned}
\mathrm{E}[X] &= \binom{n}{2} \cdot \frac{1}{n^2} \\
&= \frac{n^2 - n}{2} \cdot \frac{1}{n^2} \\
&< 1/2 .
\end{aligned}
$$

(Note that this analysis is similar to the analysis of the birthday paradox
tion 5.4.1.) Applying Markov's inequality (C.29), $\Pr\{X \geq t\} \leq \mathrm{E}[X]/$
$t = 1$ completes the proof.

In the situation described in Theorem 11.9, where $m = n^2$, it follows tha
function h chosen at random from $\mathcal{H}$ is more likely than not to have *no* co
Given the set K of n keys to be hashed (remember that K is static), it is t
to find a collision-free hash function h with a few random trials.

When n is large, however, a hash table of size $m = n^2$ is excessive. T
we adopt the two-level hashing approach, and we use the approach of Theo

hash functions h_k have a rather simple description since they are completely determined by the value of k. Since $k \in \{1, \ldots, m\}$, this description can be encoded into a key value in $M = \{0, \ldots, m-1\}$ and stored in a single cell in the table. (The function h_0 is identically 0, and this is why we choose k from the set $\{1, \ldots, m\}$ instead of from M.) The following lemma summarizes the critical property of these hash functions that motivates their use in this application. For $b_i < 2$, we define $\binom{b_i}{2}$ to be 0.

Lemma 8.17: *For all $V \subseteq M$ of size v, and all $r \geq v$,*

$$\sum_{k=1}^{p-1} \sum_{i=0}^{r-1} \binom{b_i(k, r, V)}{2} < \frac{(p-1)v^2}{r} = \frac{mv^2}{r}. \tag{8.2}$$

PROOF: The left-hand side of (8.2) counts the number of tuples $(k, \{x, y\})$ such that h_k causes x and y to collide. Equivalently, it is the number of tuples that satisfy the following two conditions:

1. $x, y \in V$ with $x \neq y$, and
2. $((kx \bmod p) \bmod r) = ((ky \bmod p) \bmod r)$.

Fix any (unordered) pair $\{x, y\} \subseteq V$ with $x \neq y$. The total contribution of this pair to the summation is the number of choices of k satisfying the second condition. In other words, this pair's contribution is the number of choices of k such that

$$
\begin{aligned}
&= 1 + \sum_{q=0}^{m-1} \sum_{j=1}^{n} \sum_{i=1}^{j-1} \frac{1}{nm^2} \\
&= 1 + m \cdot \frac{n(n-1)}{2} \cdot \frac{1}{nm^2} \qquad \text{(by equation (A.2) on page 1141)} \\
&= 1 + \frac{n-1}{2m} \\
&= 1 + \frac{n}{2m} - \frac{1}{2m} \\
&= 1 + \frac{\alpha}{2} - \frac{\alpha}{2n} .
\end{aligned}
$$

Thus, the total time required for a successful search (including the time for computing the hash function) is $\Theta(2 + \alpha/2 - \alpha/2n) = \Theta(1 + \alpha)$. ■

What does this analysis mean? If the number of elements in the table is at most proportional to the number of hash-table slots, we have $n = O(m)$ and, consequently, $\alpha = n/m = O(m)/m = O(1)$. Thus, searching takes constant time and deletion takes $O(1)$

We have

$$\begin{aligned}
\mathrm{E}\left[\sum_{j=0}^{m-1} n_j^2\right] &= \mathrm{E}\left[\sum_{j=0}^{m-1}\left(n_j + 2\binom{n_j}{2}\right)\right] && \text{(by equation (11.6))} \\
&= \mathrm{E}\left[\sum_{j=0}^{m-1} n_j\right] + 2\,\mathrm{E}\left[\sum_{j=0}^{m-1}\binom{n_j}{2}\right] && \text{(by linearity of expectation)} \\
&= \mathrm{E}[n] + 2\,\mathrm{E}\left[\sum_{j=0}^{m-1}\binom{n_j}{2}\right] && \text{(by equation (11.1))} \\
&= n + 2\,\mathrm{E}\left[\sum_{j=0}^{m-1}\binom{n_j}{2}\right] && \text{(since } n \text{ is not a random variable)}\ .
\end{aligned}$$

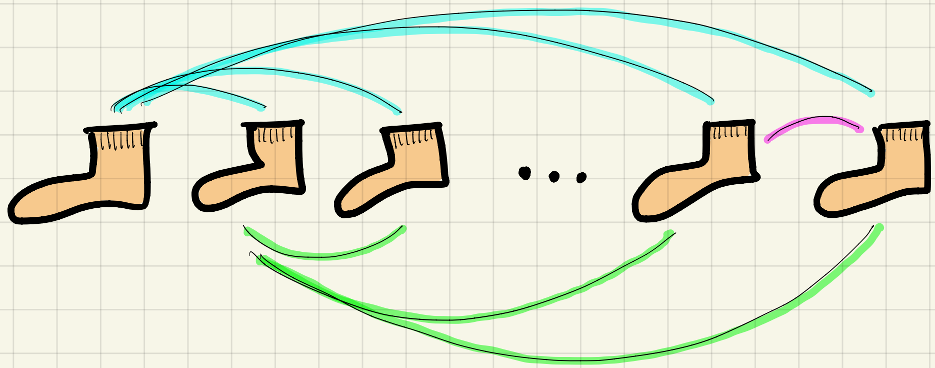
To evaluate the summation $\sum_{j=0}^{m-1}\binom{n_j}{2}$, we observe that it is just the total number of collisions. By the properties of universal hashing, the expected value of this summation is at most

$$\binom{n}{2}\frac{1}{m} = \frac{n(n-1)}{2m} = \frac{n-1}{2}\ ,$$

since $m = n$. Thus,

When did we see $\binom{n}{2}$ for the first time?

- We counted pairs of socks



$$1 + 2 + \cdots + (n-2) + (n-1) = \frac{n(n-1)}{2} = \binom{n}{2}$$

- we also counted snakes on a board with n squares

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n} 1 = \sum_{i=1}^{n} (n-i) = (n-1) + \cdots + 1 + 0 = \binom{n}{2}$$

Handshakes

n people met at a party. They all shook hands.

How many handshakes was there ?

$$\binom{n}{2}$$

Handshake $\equiv$ pair of people

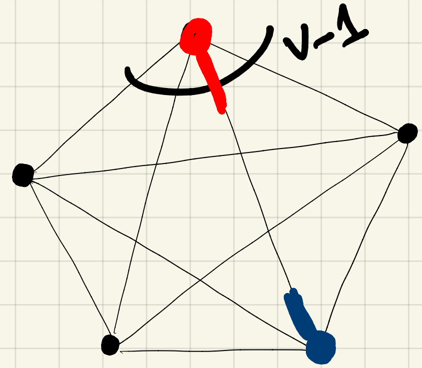
We used abtraction / generalization !

Apply a concept to different setting

Same example in graphs

Given a graph with v vertices and <u>all</u> possible edges. How many edges are there?

Example: $v = 5$

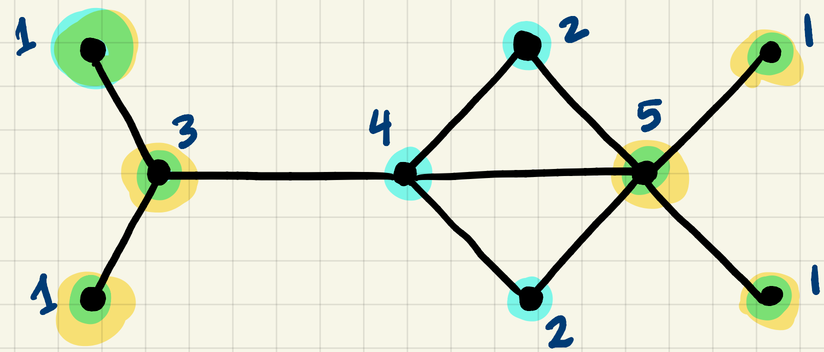


"degree of the vertex" is # edges connected to it

$e = 10$

In general, $e = \binom{v}{2} = \frac{v(v-1)}{2} = \frac{5 \times 4}{2} = 10$

$$\underbrace{2e}_{\text{twice \# edges}} = \underbrace{v(v-1)}_{\text{sum of degrees}}$$



- Add up all degrees : $1+1+3+4+2+2+5+1+1 = 20$
- Every edge is counted exactly twice in above sum
- Let d_i be degree of vertex i :

$$\sum_{i=1}^{v} d_i = 2e$$ Handshake Lemma

(say it in English)

- $\sum_{i=1}^{v} d_i$ is even
- The # vertices with odd degree is even

Abstraction

$$\binom{n}{2} = \frac{n(n-1)}{2}$$ is the number of

unordered pairs

we can make given n "things"

But what does that really mean?

To many, it can still be a little confusing!

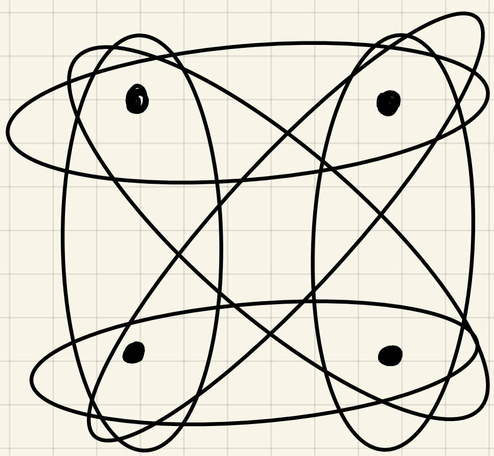
Be careful how we look at pairs

$$\boxed{\binom{n}{2} \neq \frac{n}{2}}$$

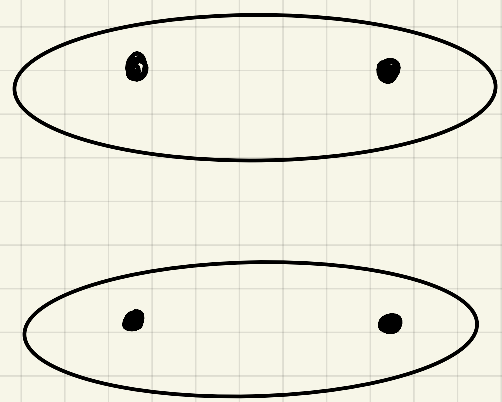
$\binom{n}{2}$: # ways we can select a pair

$\frac{n}{2}$: (when n is even) # pairs that exist simultaneously

Ex. $n = 4$

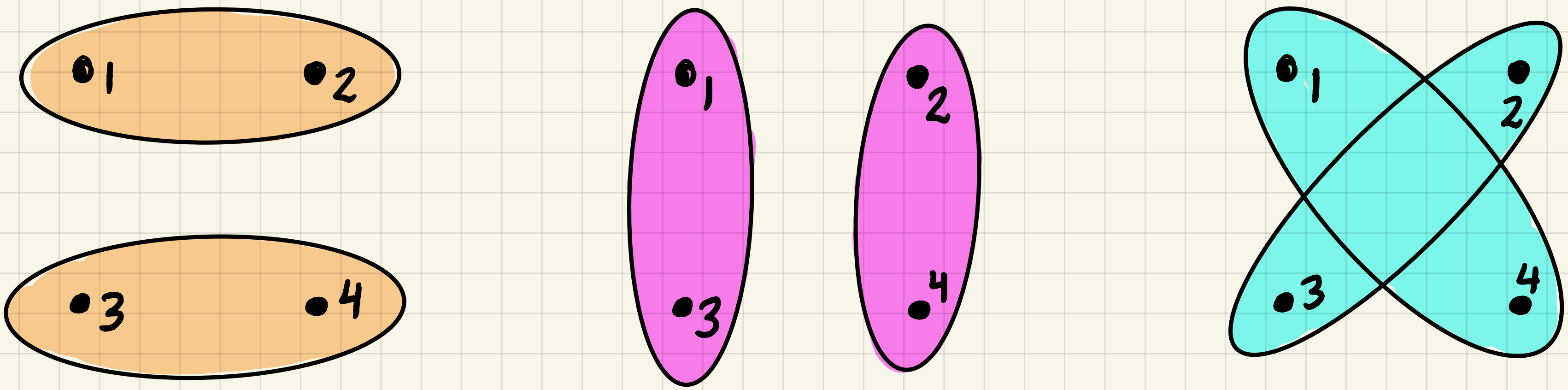


$\binom{4}{2} = 6$



$\frac{4}{2} = 2$

... But, in how many ways can we make simultaneous pairs?



3 ways $\neq \frac{n}{2} \neq \binom{n}{2}$

In general, if we have $2n$ (even) people, in how many ways can we make teams of 2?

We don't have an existing abstraction or framework.

go to scratch ...

\# ways

1. choose two people ..... $\binom{2n}{2}$

2. choose another two ..... $\binom{2n-2}{2}$

$n!$

⋮

n. choose another two ...... $\binom{2}{2}$

overcounting

$\binom{2n}{2}\binom{2n-2}{2}\cdots\binom{2}{2}$

Answer: $\dfrac{\binom{2n}{2}\binom{2n-2}{2}\cdots\binom{2}{2}}{n!} = \dfrac{(2n)!}{2^n n!}$

(see below)

$$\binom{2n}{2}\binom{2n-2}{2}\cdots\binom{2}{2} = \frac{(2n)!}{2!\,\cancel{(2n-2)!}}\,\frac{\cancel{(2n-2)!}}{2!\,\cancel{(2n-4)!}}\cdots\frac{\cancel{2!}}{2!\,0!}$$

$$= \frac{(2n)!}{\underbrace{2\cdots 2}_{n}} = \frac{(2n)!}{2^n}$$

Select k from n	ordered	unordered
no repetition	$\frac{n!}{(n-k)!}$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$
repetition	n^k	?

How many words of length 2 can we make using $\{a, b, c\}$?

No repetition : tuples (a,b) (b,a) (a,c) (c,a) (b,c) (c,b) $\frac{3!}{(3-2)!}$

No repetition, alphabetical : sets $\{a,b\}$ $\{a,c\}$ $\{b,c\}$ $\binom{3}{2}$

Repetition : tuples with repetition, add (a,a) (b,b) (c,c) 3^2

Repetition, alphabetical : Multisets $\{a,a\}$ $\{a,b\}$ $\{a,c\}$ $\{b,b\}$ $\{b,c\}$ $\{c,c\}$

we did not see a formula yet!