15-251

Some ~~Great~~ Theoretical Ideas ~~in~~ Computer Science for

"My friends keep asking me what 251 is like. I link them to this video: http://youtube.com/watch?v=M275PjvwRlI"

Probability Refresher

What's a Random Variable?

A Random Variable is a real-valued function on a sample space S

E[X+Y] = E[X] + E[Y]

Probability Refresher

What does this mean: E[X | A]?

Is this true:

Pr[A] = Pr[A | B] Pr[B] + Pr[A | $\overline{B}$] Pr[$\overline{B}$]

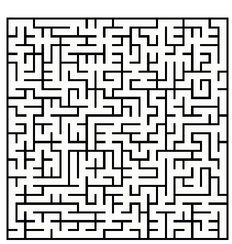
Yes!

Similarly:

E[X] = E[X | A] Pr[A] + E[X | $\overline{A}$] Pr[$\overline{A}$]

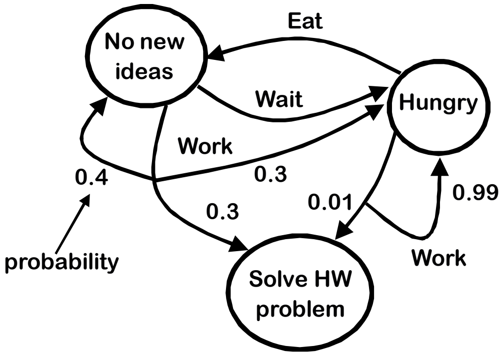
Random Walks

Lecture 13 (February 26, 2008)

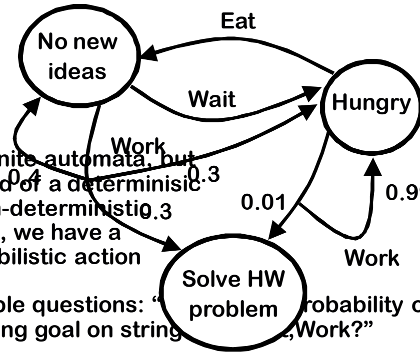


How to walk home drunk

Abstraction of Student Life



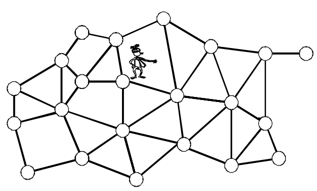
Abstraction of Student Life



Like finite automata, but instead of a determinisic or non-deterministic action, we have a probabilistic action

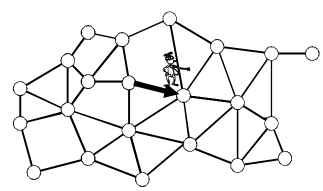
Example questions: “ robability of reaching goal on string ,Work?”

Simpler: Random Walks on Graphs



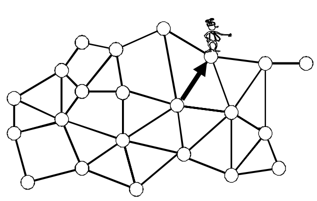
At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs



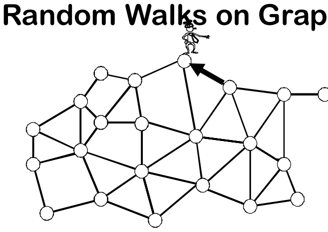
At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs



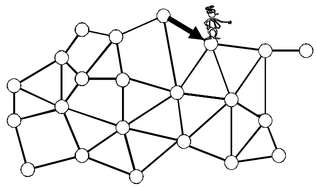
At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs



At any node, go to one of the neighbors of the node with equal probability

Simpler: Random Walks on Graphs

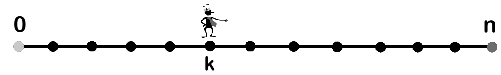


At any node, go to one of the neighbors of the node with equal probability

Random Walk on a Line

You go into a casino with \$k, and at each time step, you bet \$1 on a fair game

You leave when you are broke or have \$n



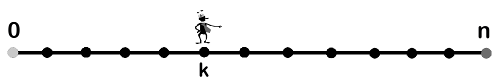
Question 1: what is your expected amount of money at time t?

Let X_t be a R.V. for the amount of \$\$\$ at time t

Random Walk on a Line

You go into a casino with \$k, and at each time step, you bet \$1 on a fair game

You leave when you are broke or have \$n



$X_t = k + \delta_1 + \delta_2 + \ldots + \delta_t,$

(δ_i is RV for change in your money at time i)

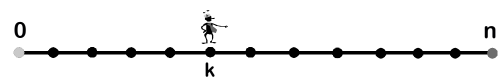
$E[\delta_i] = 0$

So, $E[X_t] = k$

Random Walk on a Line

You go into a casino with \$k, and at each time step, you bet \$1 on a fair game

You leave when you are broke or have \$n



Question 2: what is the probability that you leave with \$n?

Random Walk on a Line

Question 2: what is the probability that you leave with \$n?

$E[X_t] = k$

$E[X_t] = E[X_t | X_t = 0] \times Pr(X_t = 0)$
$+ E[X_t | X_t = n] \times Pr(X_t = n)$
$+ E[X_t | \text{neither}] \times Pr(\text{neither})$

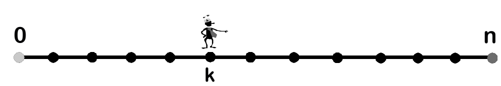
$k = n \times Pr(X_t = n)$
$+ (\text{something}_t) \times Pr(\text{neither})$

As $t \to \infty$, Pr(neither) $\to 0$, also $\text{something}_t < n$
Hence $Pr(X_t = n) \to k/n$

Another Way To Look At It

You go into a casino with \$k, and at each time step, you bet \$1 on a fair game

You leave when you are broke or have \$n

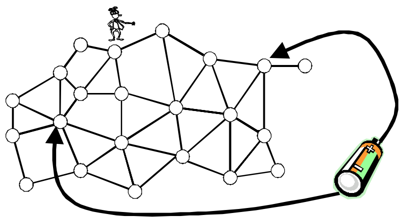


Question 2: what is the probability that you leave with \$n?

= probability that I hit green before I hit red

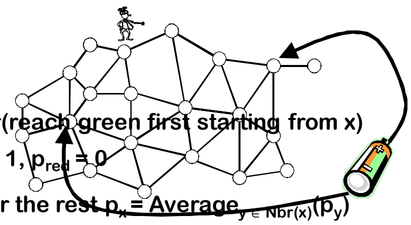
Random Walks and Electrical Networks

What is chance I reach green before red?



Same as voltage if edges are resistors and we put 1-volt battery between green and red

Random Walks and Electrical Networks



p_x = Pr(reach green first starting from x)

p_{green}= 1, p_{red} = 0

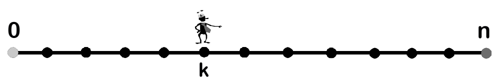
And for the rest $p_x = \text{Average}_{y \in Nbr(x)}(p_y)$

Same as equations for voltage if edges all have same resistance!

Another Way To Look At It

You go into a casino with \$k, and at each time step, you bet \$1 on a fair game

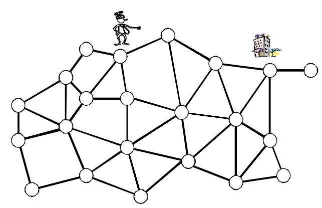
You leave when you are broke or have \$n



Question 2: what is the probability that you leave with \$n?

voltage(k) = k/n
= Pr[hitting n before 0 starting at k] !!!

Getting Back Home

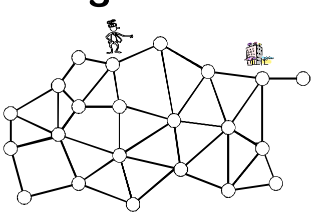


Lost in a city, you want to get back to your hotel
How should you do this?

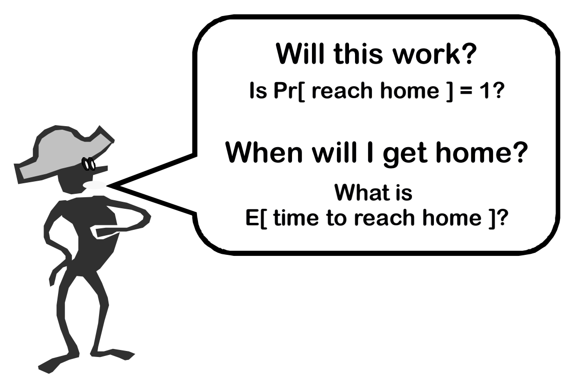
Depth First Search!

Requires a good memory and a piece of chalk

Getting Back Home



How about walking randomly?



Pr[will reach home] = 1

We Will Eventually Get Home

Look at the first n steps

There is a non-zero chance p_1 that we get home

Also, $p_1 \geq (1/n)^n$

Suppose we fail

Then, wherever we are, there is a chance $p_2 \geq (1/n)^n$ that we hit home in the next n steps from there

Probability of failing to reach home by time kn

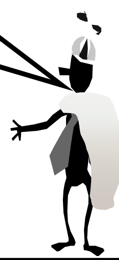
$= (1 - p_1)(1 - p_2) \ldots (1 - p_k) \rightarrow 0$ as $k \rightarrow \infty$

Furthermore:

If the graph has
n nodes and m edges, then

E[time to visit all nodes]
$\leq 2m \times (n-1)$

E[time to reach home] is at most this



Cover Times

Cover time (from u)
C_u = E [time to visit all vertices | start at u]

Cover time of the graph
$C(G) = \max_u \{ C_u \}$

(worst case expected time to see all vertices)

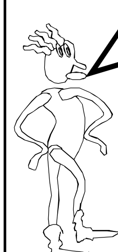
Cover Time Theorem

If the graph G has
n nodes and m edges, then
the cover time of G is

$C(G) \leq 2m (n - 1)$

Any graph on n vertices has $< n^2/2$ edges

Hence $C(G) < n^3$ for all graphs G



Actually, we get home
pretty fast…

Chance that we don't hit home by
$(2k)2m(n-1)$ steps is $(½)^k$

A Simple Calculation

True of False:

If the average income of people is \$100 then more than 50% of the people can be earning more than \$200 each

False! else the average would be higher!!!

Markov's Inequality

If X is a non-negative r.v. with mean E[X], then

$\Pr[\, X > 2\, E[X]\,] \leq ½$

$\Pr[\, X > k\, E[X]\,] \leq 1/k$

Andrei A. Markov

Markov's Inequality

Non-neg random variable X has expectation A = E[X]

$A = E[X] = E[X \mid X > 2A]\Pr[X > 2A] + E[X \mid X \leq 2A]\Pr[X \leq 2A]$

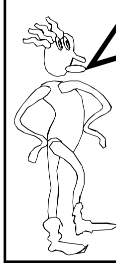
$\geq E[X \mid X > 2A]\Pr[X > 2A]$ (since X is non-neg)

Also, $E[X \mid X > 2A] > 2A$

$\Rightarrow A \geq 2A \times \Pr[X > 2A]$

$\Rightarrow ½ \geq \Pr[X > 2A]$

$\Pr[\, X > k \times \text{expectation}\,] \leq 1/k$



Actually, we get home pretty fast…

Chance that we don't hit home by $(2k)2m(n-1)$ steps is $(½)^k$

An Averaging Argument

Suppose I start at u

E[time to hit all vertices | start at u] $\leq$ C(G)

Hence, by Markov's Inequality:
Pr[time to hit all vertices > 2C(G) | start at u] $\leq ½$

So Let's Walk Some Mo!

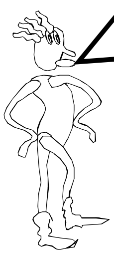
Pr [time to hit all vertices > 2C(G) | start at u] $\leq ½$

Suppose at time 2C(G), I'm at some node with more nodes still to visit

Pr [haven't hit all vertices in 2C(G) more time | start at v] $\leq ½$

Chance that you failed both times $\leq ¼ = (½)^2$

Hence,
Pr[havent hit everyone in time k × 2C(G)] $\leq (½)^k$



Hence, if we know that

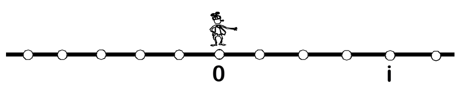
Expected Cover Time
$C(G) < 2m(n-1)$

then

Pr[home by time $4k\,m(n-1)$]
$\geq 1 - (½)^k$

Random walks
on infinite graphs

Random Walk On a Line

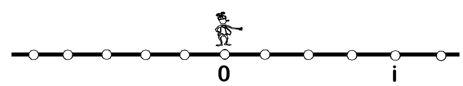


Flip an unbiased coin and go left/right

Let X_t be the position at time t

$$\Pr[\,X_t = i\,] = \Pr[\,\#\text{heads} - \#\text{tails} = i\,]$$
$$= \Pr[\,\#\text{heads} - (t - \#\text{heads}) = i\,]$$
$$= \binom{t}{(t+i)/2} / 2^t$$

Random Walk On a Line

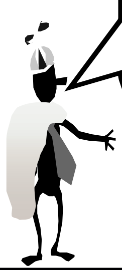


$$\Pr[\,X_{2t} = 0\,] = \binom{2t}{t} / 2^{2t} \leq \Theta(1/\sqrt{t})$$ Sterling's approx

Y_{2t} = indicator for $(X_{2t} = 0)$ $\Rightarrow E[\,Y_{2t}\,] = \Theta(1/\sqrt{t})$

Z_{2n} = number of visits to origin in 2n steps

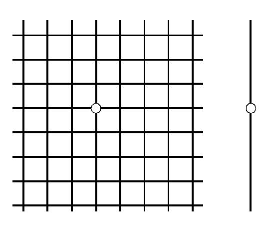
$$E[\,Z_{2n}\,] = E[\,\textstyle\sum_{t = 1\ldots n} Y_{2t}\,]$$
$$\leq \Theta(1/\sqrt{1} + 1/\sqrt{2} + \ldots + 1/\sqrt{n}) = \Theta(\sqrt{n})$$



In n steps, you expect to return to the origin $\Theta(\sqrt{n})$ times!

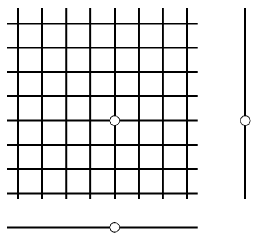
How About a 2-d Grid?

Let us simplify our 2-d random walk:
move in both the x-direction and y-direction…



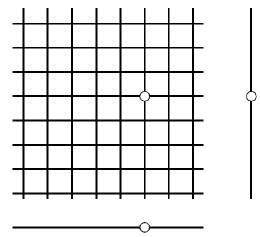
How About a 2-d Grid?

Let us simplify our 2-d random walk:
move in both the x-direction and y-direction…



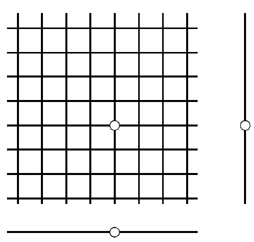
How About a 2-d Grid?

Let us simplify our 2-d random walk:
move in both the x-direction and y-direction…



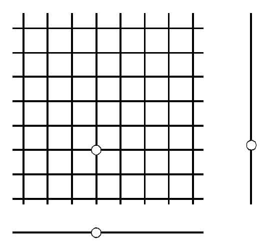
How About a 2-d Grid?

Let us simplify our 2-d random walk:
move in both the x-direction and y-direction…



How About a 2-d Grid?

Let us simplify our 2-d random walk:
move in both the x-direction and y-direction…



In The 2-d Walk

Returning to the origin in the grid
$\Leftrightarrow$ both "line" random walks return to their origins

Pr[visit origin at time t] $= \Theta(1/\sqrt{t}) \times \Theta(1/\sqrt{t})$
$= \Theta(1/t)$

E[# of visits to origin by time n]
$= \Theta(1/1 + 1/2 + 1/3 + \ldots + 1/n) = \Theta(\log n)$

But In 3D

Pr[visit origin at time t] $= \Theta(1/\sqrt{t})^3 = \Theta(1/t^{3/2})$

$\lim_{n \to \infty}$ E[# of visits by time n] < K (constant)

Hence Pr[never return to origin] > 1/K

Drunk man will find way home, but drunk bird may get lost forever

- Shizuo Kakutani

Dot Proofs

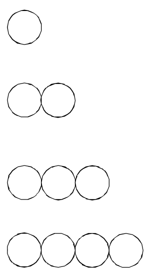
Prehistoric Unary

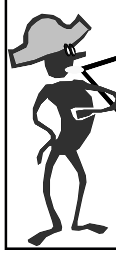
1

2

3

4





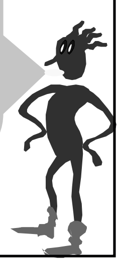
Hang on a minute!

Isn't unary too literal as a representation?

Does it deserve to be an "abstract" representation?

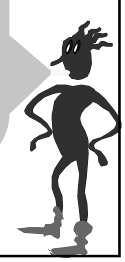
It's important to respect each representation, no matter how primitive

Unary is a perfect example



Consider the problem of finding a formula for the sum of the first n numbers

You already used induction to verify that the answer is ½n(n+1)

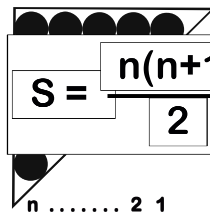


$$1 + 2 + 3 + \ldots + n\text{-}1 + n = S$$
$$n + n\text{-}1 + n\text{-}2 + \ldots + 2 + 1 = S$$

$$n\text{+}1 + n\text{+}1 + n\text{+}1 + \ldots + n\text{+}1 + n\text{+}1 = 2S$$
$$n(n+1) = 2S$$

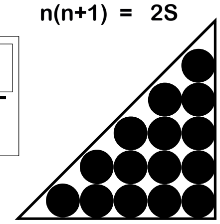
$$S = \frac{n(n+1)}{2}$$

$$\boxed{1 + 2 + 3 + \ldots + n\text{-}1 + n = S}$$
$$\boxed{n + n\text{-}1 + n\text{-}2 + \ldots + 2 + 1 = S}$$



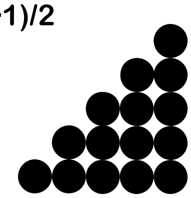
$$S = \frac{n(n+1)}{2}$$

$$n(n+1) = 2S$$



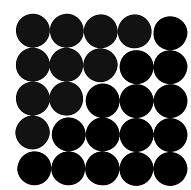
n^{th} Triangular Number

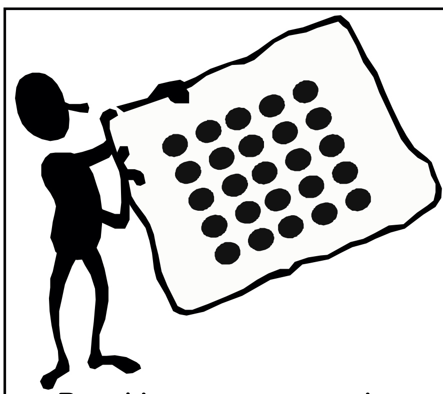
$$\Delta_n = 1 + 2 + 3 + \ldots + n\text{-}1 + n$$
$$= n(n+1)/2$$



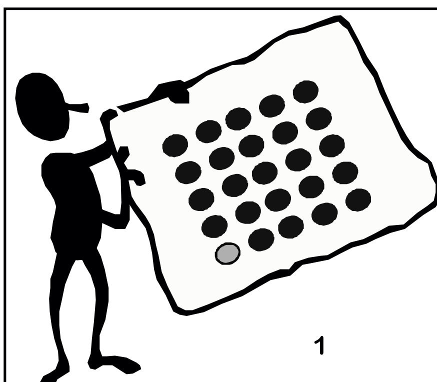
n^{th} Square Number

$$_n = n^2$$
$$= \Delta_n + \Delta_{n-1}$$



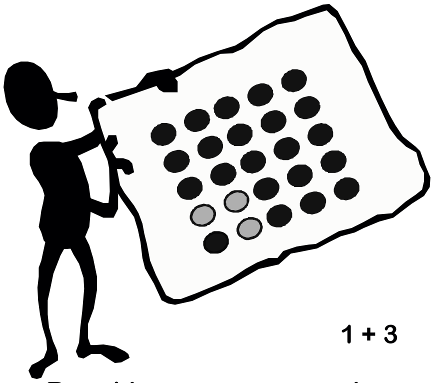


Breaking a square up in a new way



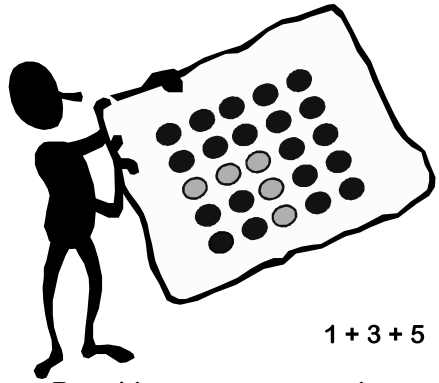
1

Breaking a square up in a new way



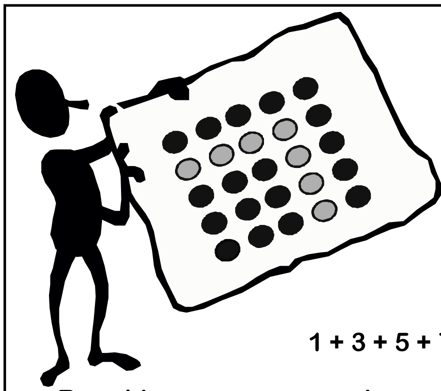
$1 + 3$

Breaking a square up in a new way



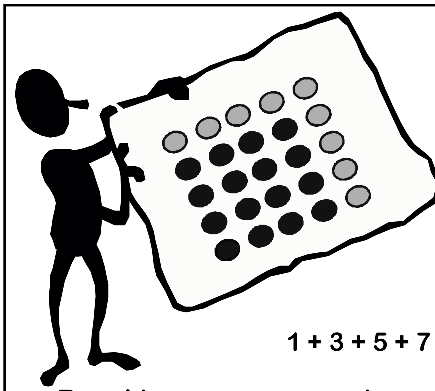
$1 + 3 + 5$

Breaking a square up in a new way



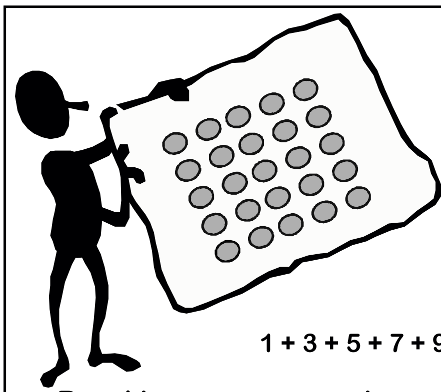
$1 + 3 + 5 + 7$

Breaking a square up in a new way



$1 + 3 + 5 + 7 + 9$

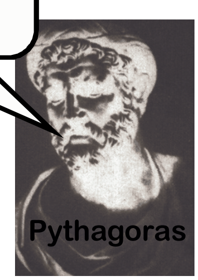
Breaking a square up in a new way



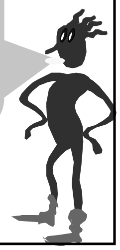
$1 + 3 + 5 + 7 + 9 = 5^2$

Breaking a square up in a new way

The sum of the first n odd numbers is n^2



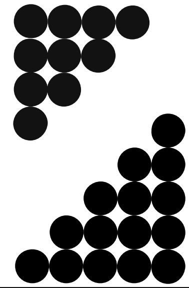
Here is an alternative dot proof of the same sum....



n^{th} Square Number

$n = \Delta_n + \Delta_{n-1}$

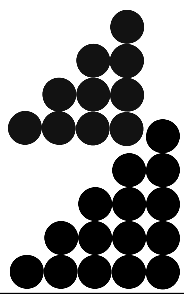
$= n^2$



n^{th} Square Number

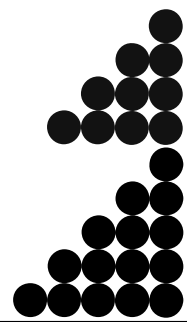
$n = \Delta_n + \Delta_{n-1}$

$= n^2$



n^{th} Square Number

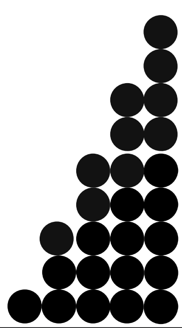
$n = \Delta_n + \Delta_{n-1}$



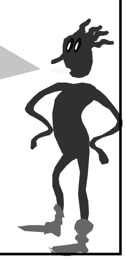
n^{th} Square Number

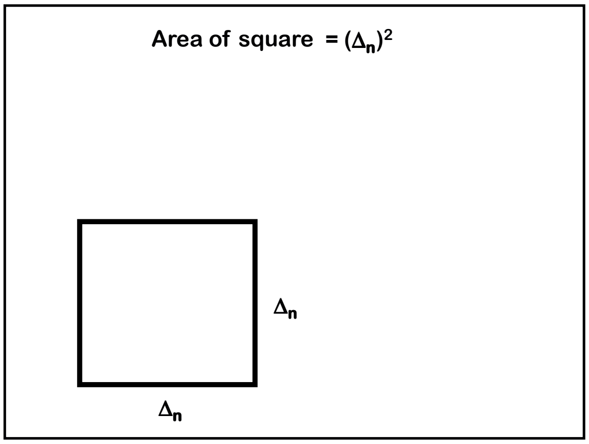
$n = \Delta_n + \Delta_{n-1}$

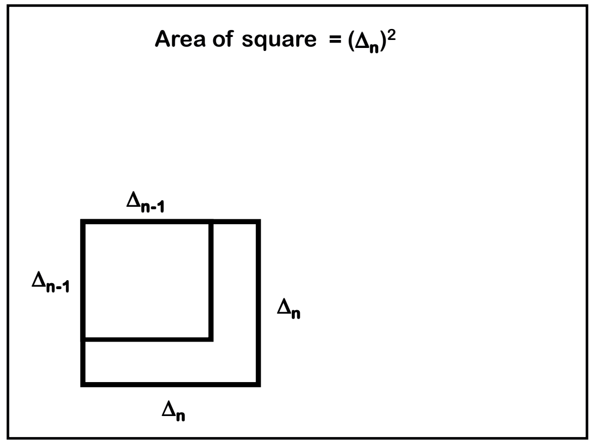
= Sum of first n odd numbers

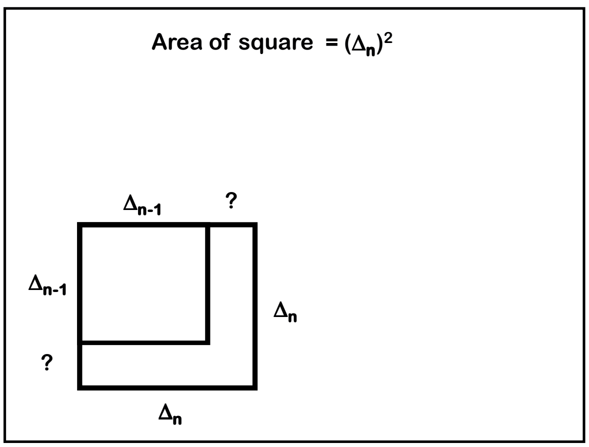


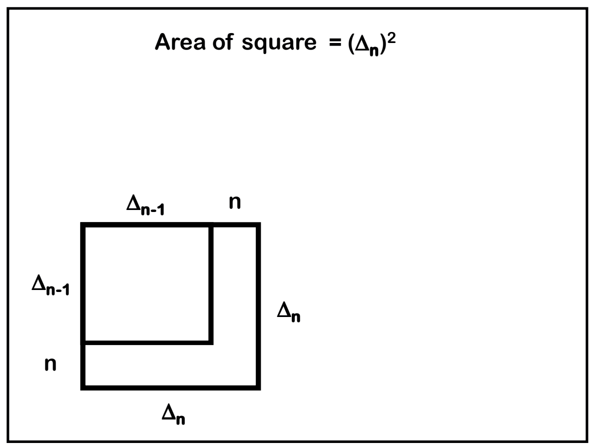
Check the next one out...

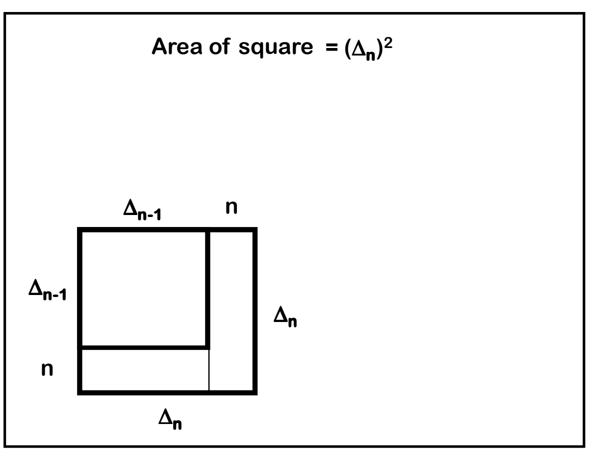


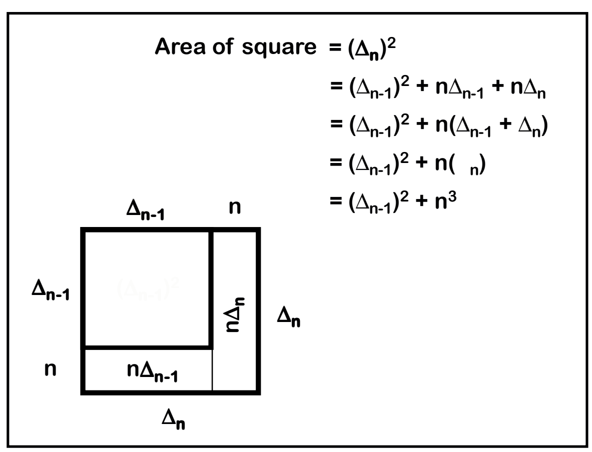

Area of square $= (\Delta_n)^2$
Δ_n
Δ_n


Area of square $= (\Delta_n)^2$
Δ_{n-1}
Δ_{n-1}
Δ_n
Δ_n


Area of square $= (\Delta_n)^2$
Δ_{n-1}
?
Δ_{n-1}
Δ_n
?
Δ_n

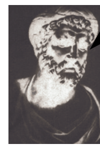

Area of square $= (\Delta_n)^2$
Δ_{n-1}
n
Δ_{n-1}
Δ_n
n
Δ_n


Area of square $= (\Delta_n)^2$
Δ_{n-1}
n
Δ_{n-1}
Δ_n
n
Δ_n


Area of square $= (\Delta_n)^2$
$= (\Delta_{n-1})^2 + n\Delta_{n-1} + n\Delta_n$
$= (\Delta_{n-1})^2 + n(\Delta_{n-1} + \Delta_n)$
$= (\Delta_{n-1})^2 + n(\ _n)$
$= (\Delta_{n-1})^2 + n^3$
Δ_{n-1}
n
Δ_{n-1}
$n\Delta_n$
Δ_n
n
$n\Delta_{n-1}$
Δ_n

$$
\begin{aligned}
(\Delta_n)^2 &= n^3 + (\Delta_{n-1})^2 \\
&= n^3 + (n-1)^3 + (\Delta_{n-2})^2 \\
&= n^3 + (n-1)^3 + (n-2)^3 + (\Delta_{n-3})^2 \\
&= n^3 + (n-1)^3 + (n-2)^3 + \ldots + 1^3
\end{aligned}
$$

$$
\begin{aligned}
(\Delta_n)^2 &= 1^3 + 2^3 + 3^3 + \ldots + n^3 \\
&= [\, n(n+1)/2 \,]^2
\end{aligned}
$$



Here's What You Need to Know…

Random Walk in a Line

Cover Time of a Graph

Markov's Inequality

Dot Proofs