Great Theoretical Ideas In Computer Science

Steven Rudich, Anupam Gupta CS 15-251 Spring 2004

Lecture 22 April 6, 2004 Carnegie Mellon University

The probabilistic method & infinite probability spaces

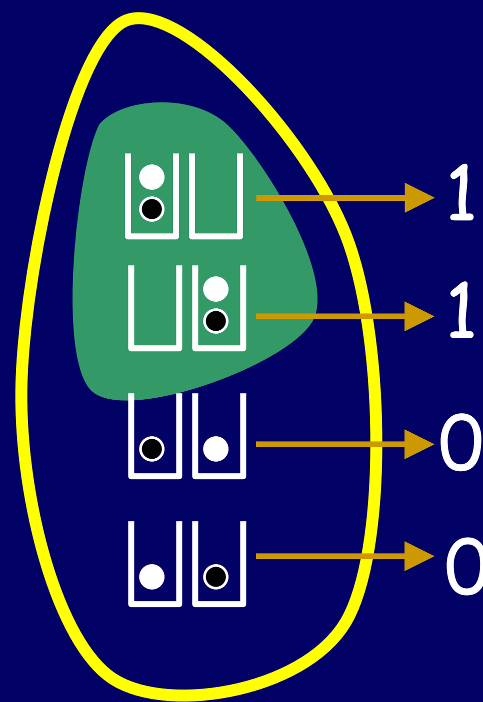
Events and Random Variables

- An event is a subset of sample space S.
- A random variable is a (real-valued) function on S.

Eg: we throw a black and a white ball into 2 equally likely bins.

event E = {both balls fall in same bin}

R.V. X = number of empty bins.



Events and Random Variables

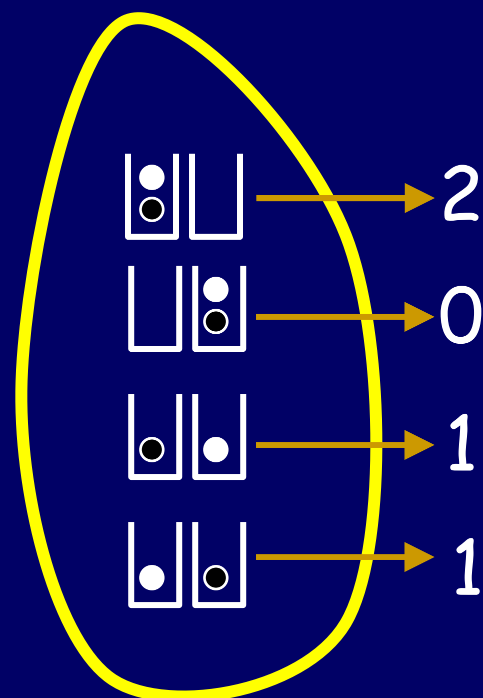
- An event is a subset of S.
- A random variable is a (real-valued) function on S.

Eg: we throw a black and a white ball into 2 equally likely bins.

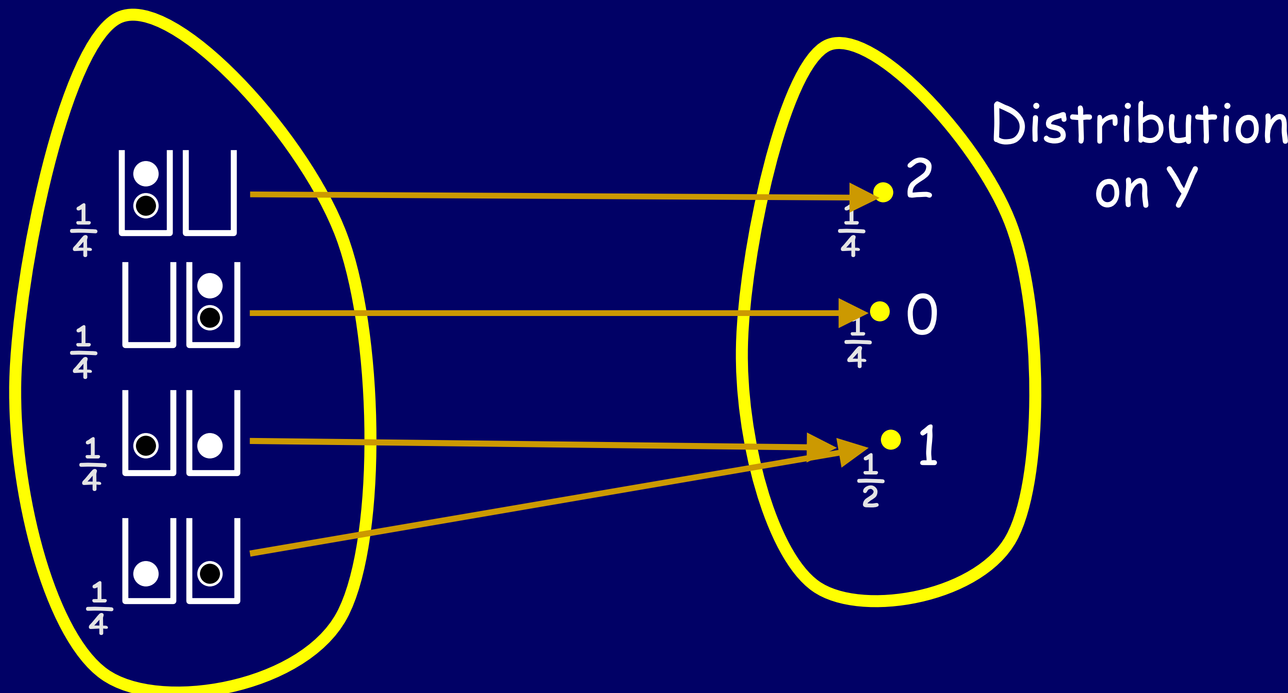
event E = {both balls fall in same bin}

R.V. X = number of empty bins.

Y = number of balls in bin #1



Thinking about R.V.'s

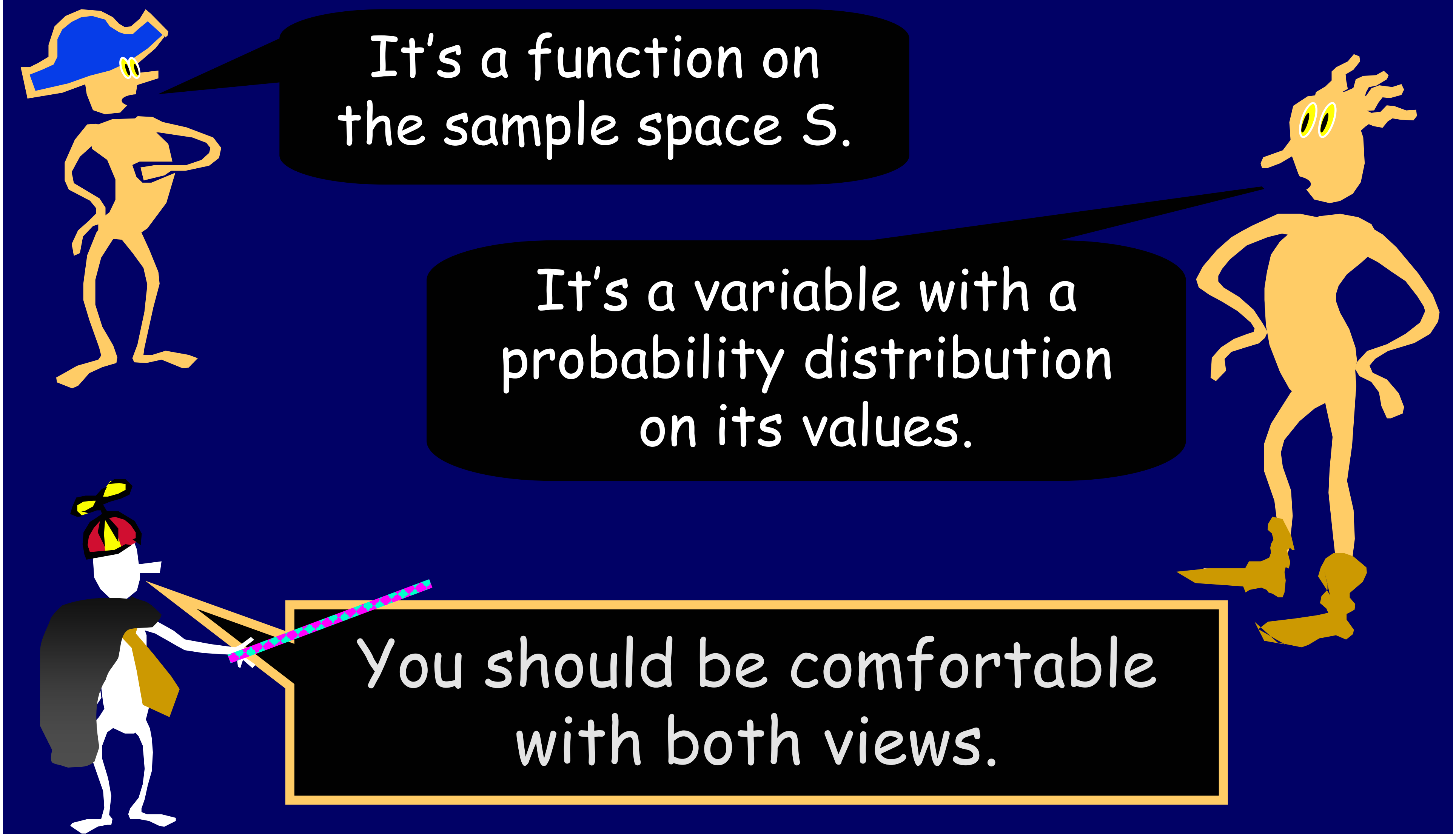


Distribution
+ real-valued function

Distribution
on the reals

Y = number of balls in bin #1

The guises of a random variable



Definition: expectation

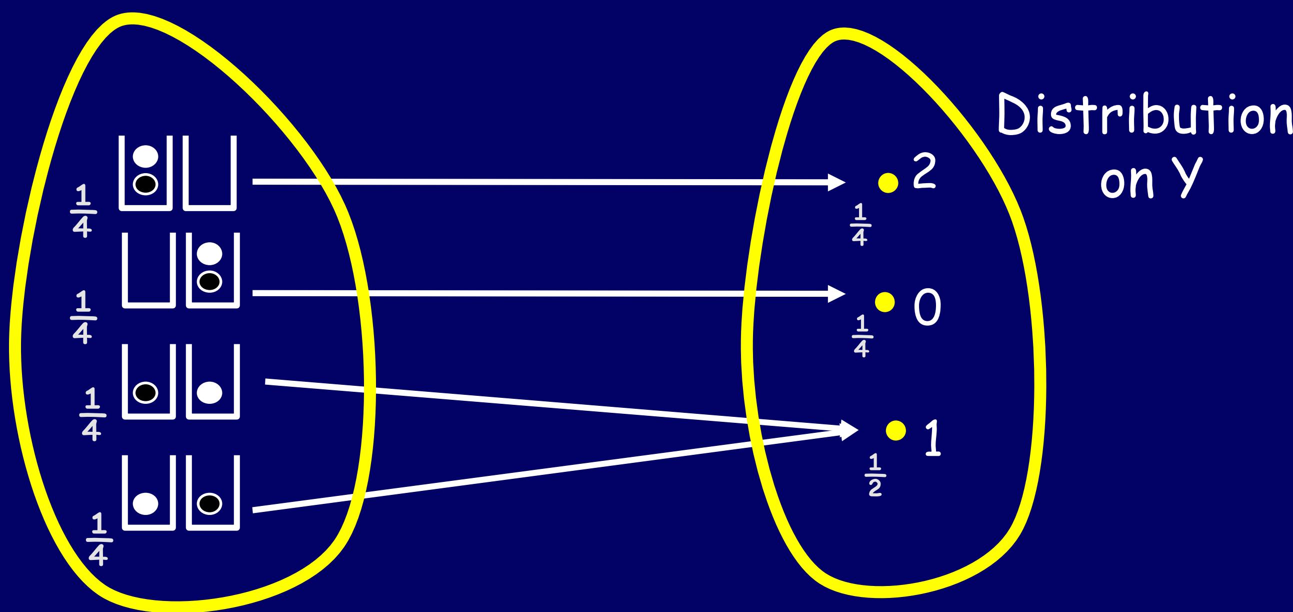
The <u>expectation</u>, or <u>expected value</u> of a random variable Y is

$$E[Y] = \sum_{x \in S} \Pr(x) \times Y(x)$$

$$= \sum_{k} \Pr(Y = k) \times k$$

$$\sum_{(x \in S \mid Y(x) = k)} \Pr(x)$$

Thinking about expectation

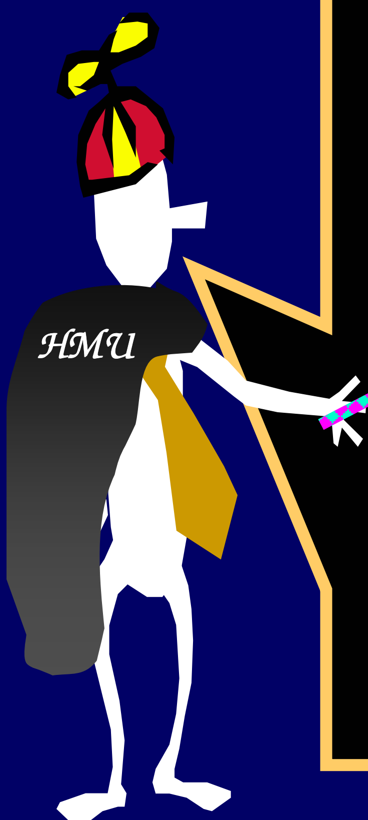


$$\boxed{\sum_{x \in S} Y(x)\,\Pr(x)} \qquad \boxed{\sum_k k \Pr(Y = k)}$$

$E[Y] = \frac{1}{4}*2 + \frac{1}{4}*0 + \frac{1}{4}*1 + \frac{1}{4}*1 = 1.$

$E[Y] = \frac{1}{4}*0 + \frac{1}{2}*1 + \frac{1}{4}*2 = 1.$

Linearity of Expectation



If $Z = X+Y$, then

$$E[Z] = E[X] + E[Y]$$

Even if X and Y are not independent.

Question

There are 156 students in a class

There are 156 "alphabet cookies"
(six for each letter of the alphabet)

I hand the cookies randomly to students, one to each.

What is the average number of students who have a cookie with letter = first letter of their name?

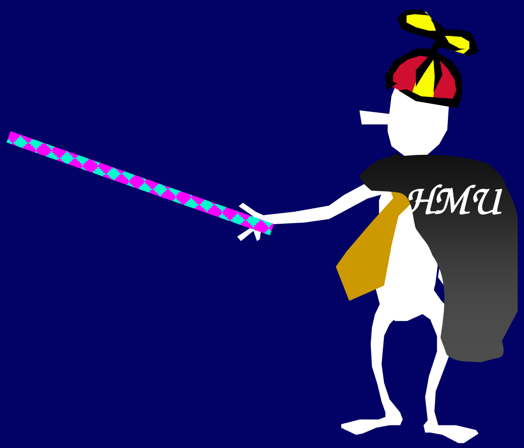
Use Linearity of Expectations

X_j = 1 if student j got a cookie with the right letter
0 otherwise

$X = \Sigma_{j = 1\ldots156} X_j$ = # lucky students

$E[X_j] = 6/156 = 1/26.$

$E[X] = E\,[\Sigma_{j = 1\ldots156} X_j] =$
$\Sigma_{j = 1\ldots156} E[X_j] = 156 * 1/26 = 6.$



Some things to note

Random variables X_i and X_j not independent

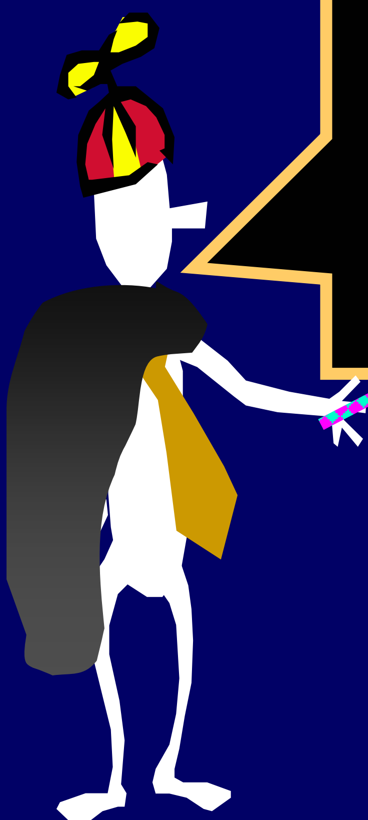
But we don't care!!!

E[Y+Z] = E[Y] + E[Z] even if Y and Z are dependent.

E[X] does not depend on distribution of people's names

We are correct even when all students have names beginning with A!!!

New topic: The probabilistic method

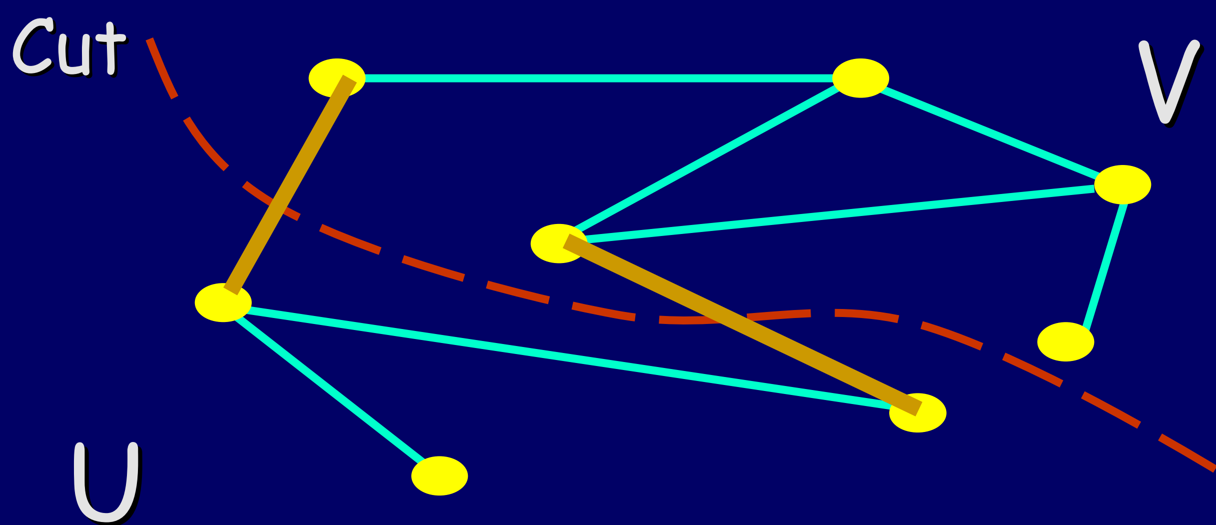


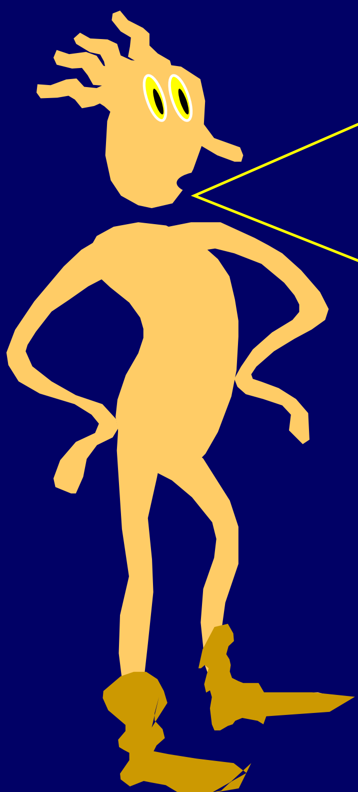
Use a probabilistic argument to prove a non-probabilistic mathematical theorem.

Definition: A cut in a graph.

A *cut* is a partition of the nodes of a graph into two sets: U and V.

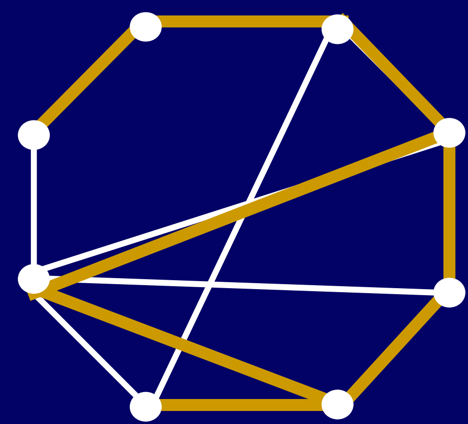
We say that an edge crosses the cut if it goes from a node is U to a node in V.





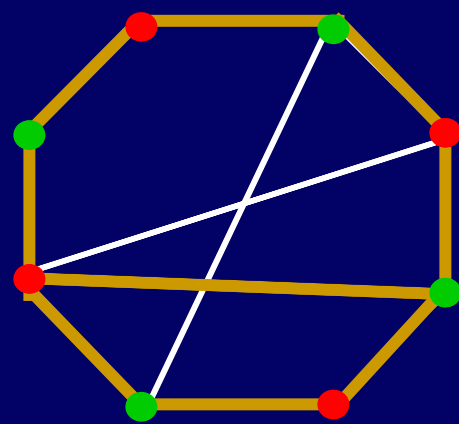
Theorem:

In any graph, there exists a cut such that at least half the edges cross the cut.



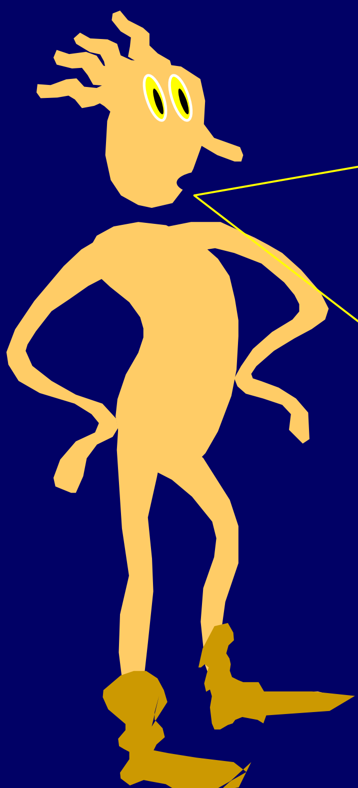
G has 11 edges.

This is a cut with $8 \geq \frac{1}{2}(11)$ edges.



G has 11 edges.

This is a cut with $9 \geq \frac{1}{2}(11)$ edges.

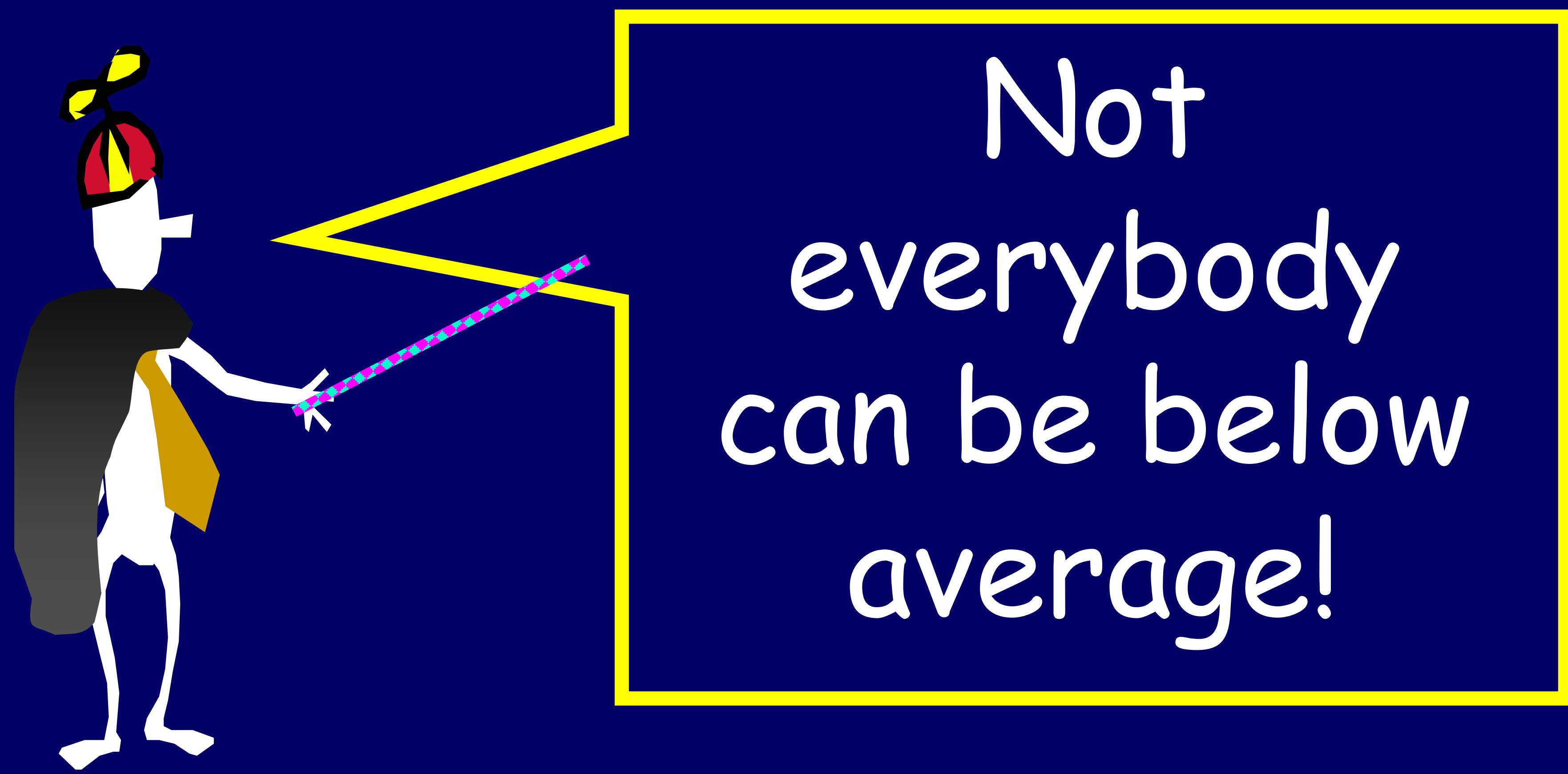


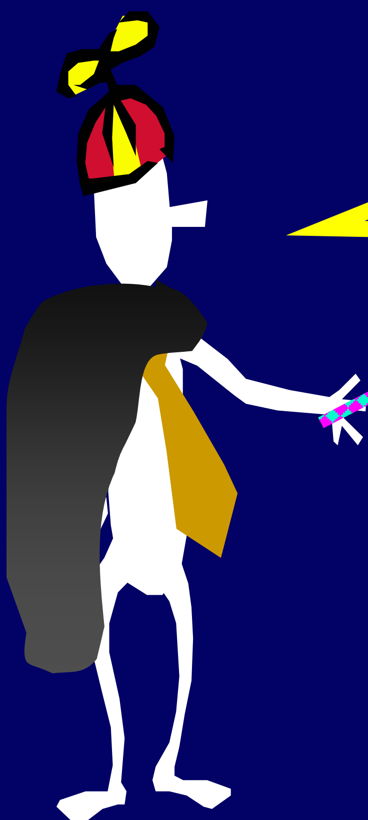
Theorem:

In any graph, there exists a cut such that ≥ ½(# edges) cross the cut.

How are we going to prove this?

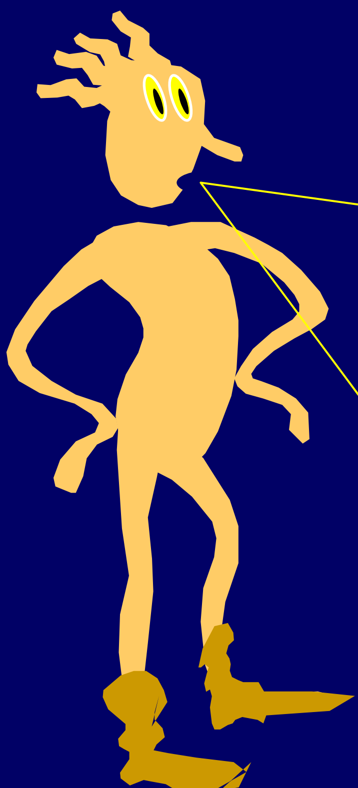
We will show that if we pick a cut <u>at random</u>, the <u>expected</u> number of edges crossing is ½(# edges).


Not everybody can be below average!



What might be is surely possible!

The Probabilistic Method



Theorem:

In any graph, there exists a cut such that ≥ ½(# edges) cross the cut.

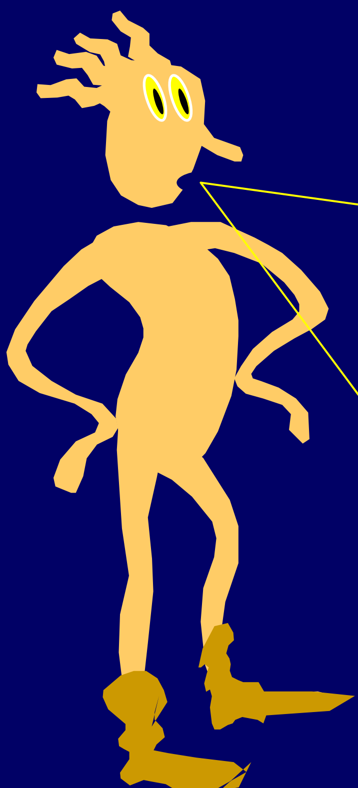
Proof:

Pick a cut of G uniformly at random. I.e., for each node, flip a fair coin to determine if it is in U or V.

Let X_e be the indicator RV for the event that edge e crosses the cut.

What is $E[X_e]$?

Ans: ½.



Theorem:

In any graph, there exists a cut such that ≥ ½(# edges) cross the cut.

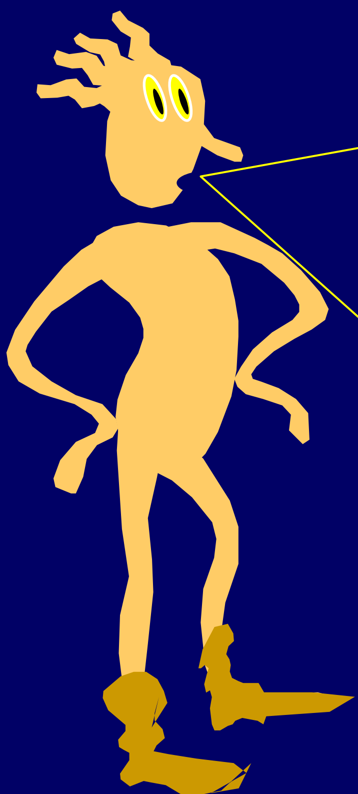
Proof:

- Pick random cut.
- Let X_e=1 if e crosses, else X_e=0.
- Let X = #(edges crossing cut).
- So, $X = \sum_e X_e$.
- Also, $E[X_e] = \frac{1}{2}$.
- By linearity of expectation,

$$E[X] = \tfrac{1}{2}(\text{total \#edges}).$$

E[# of edges crossing cut]
= $\frac{1}{2}$(# of edges)

The sample space of all possible cuts must contain at least one cut that at least half the edges cross: if not, the average number of edges would be less than half!



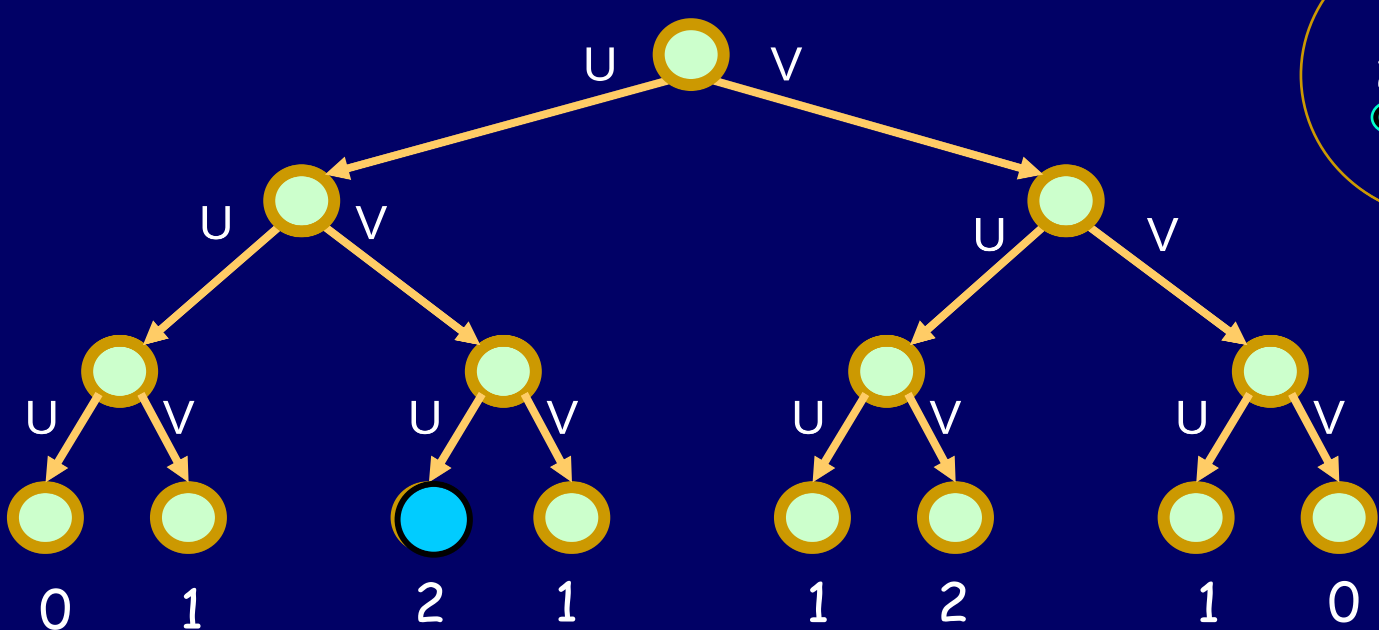
Pictorial view (important!)

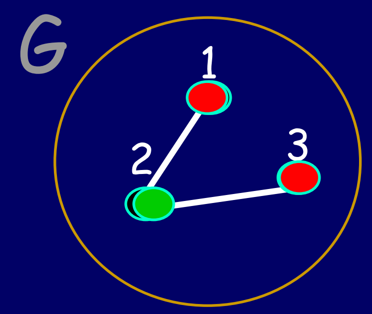
View all the cuts as leaves of a choice tree

where the i^{th} choice is where to put node i.

Label each leaf by value of X

$\Rightarrow$ E[X] = avg leaf value.





The Probabilistic Method

Goal: show that there exists an object of value at least v.

Proof strategy:

- Define distribution D over objects.
- Define a RV X:
  X(object) = value of object.
- Show $E[X] \geq v$. Conclude it must be possible to have $X \geq v$.

Probabilistic Method for MaxCut

Theorem:

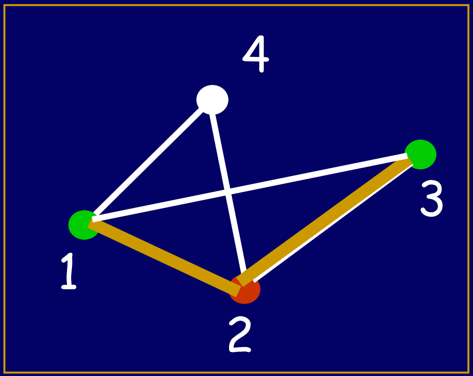
If we take a random cut in a graph G, then
E[cut value] ≥ ½(# of edges of G).

Theorem:

Any graph G has a cut which contains half its edges.

And furthermore…

Suppose I already placed nodes 1,2,…,k into U and V in some way. M of the edges between these nodes are already cut.



If we were to now split the nodes k+1 through n randomly between U and V, we'd get

E[edges cut] = M edges already cut
+ ½ (number of edges not already determined)

Can we use the probabilistic method to find a cut of size $\frac{1}{2}$(# of edges)?

In this case you can, through a neat strategy called the conditional expectation method

Idea: make decisions in greedy manner to maximize expectation-to-go.

$$\Pr(A \cap B) = \Pr(A)\,\Pr(B|A)$$

If B independent of A,
then $\Pr(B|A) = \Pr(B)$

and then
$$\Pr(A \cap B) = \Pr(A)\,\Pr(B)$$

Conditional Probabilities

Def: Conditional Expectation

For RV X and event A,
the "<u>conditional expectation</u> of X given A" is:

$$E[X \mid A] = \sum_k \; k \times \Pr(X = k \mid A)$$

E.g., roll two dice. X = sum of dice, E[X] = 3.5+3.5

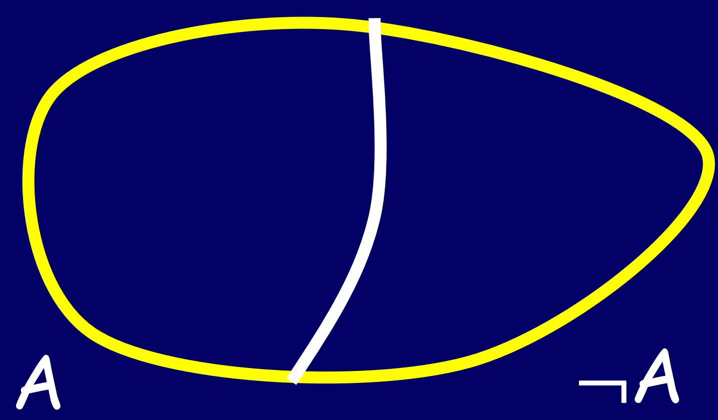
Let A be the event that the first die is 5.

$$E[X|A] = 5+3.5 = 8.5$$

A very useful formula using Conditional Expectation

If S is partitioned into two events A and $\neg A$, then

$$E[X] = E[X \mid A]\,\Pr(A) + E[X \mid \neg A]\,\Pr(\neg A)$$

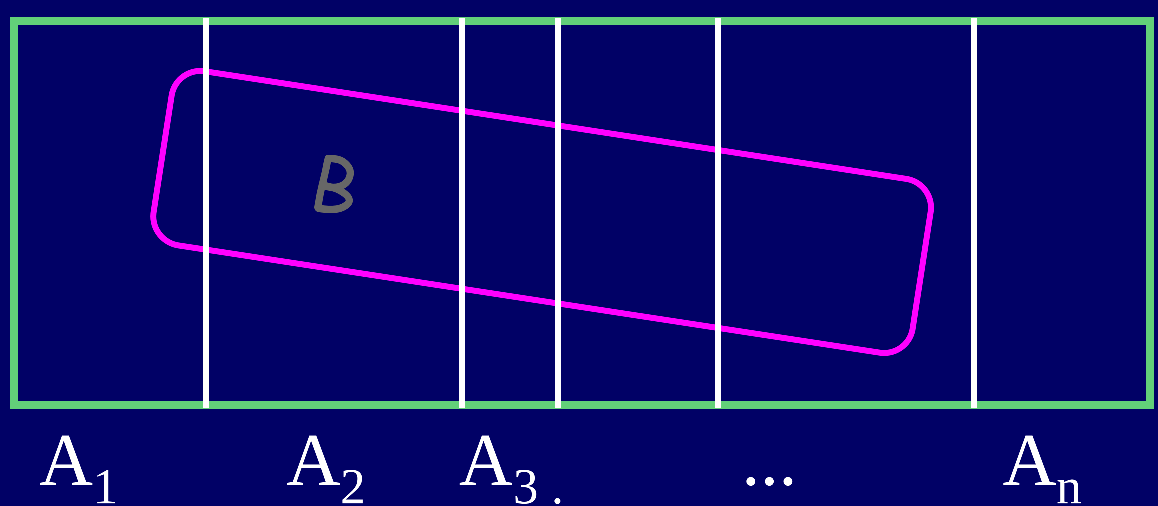


the proof uses a convenient fact

For any partition of the sample space S into disjoint events $A_1, A_2, \ldots, A_n$, and any event B,

$$Pr(B) = \sum_i Pr(B \cap A_i)$$

$$= \sum_i Pr(B|A_i)\, Pr(A_i).$$



Proof

For any partition of S into $A_1, A_2, \ldots$, we have

$$E[X] = \sum_i E[X \mid A_i] \Pr(A_i).$$

$$\begin{aligned}
E[X] &= \sum_k k \times \Pr(X = k) \\
&= \sum_k k \times \left[\sum_i \Pr(X = k \mid A_i) \Pr(A_i) \right] \\
&\qquad\qquad \text{(changing order of summation)} \\
&= \sum_i \left[\sum_k k \times \sum_i \Pr(X = k \mid A_i) \right] \times \Pr(A_i) \\
&= \sum_i E[X \mid A_i] \Pr(A_i).
\end{aligned}$$

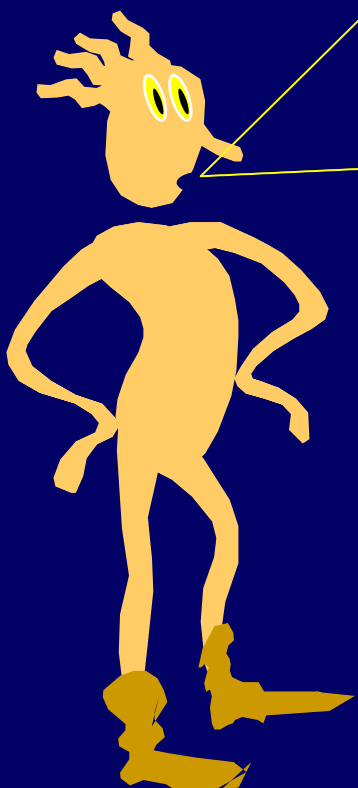
Conditional Expectation

If S is partitioned into two events A and $\neg A$, then

$$E[X] = E[X \mid A]\,\Pr(A) + E[X \mid \neg A]\,\Pr(\neg A)$$

Hence: both $E[X \mid A]$ and $E[X \mid \neg A]$ cannot be less than $E[X]$.

Recap of cut argument



Pick random cut.

- Let X_e=1 if e crosses, else X_e=0.
- Let X = total #edges crossing.
- So, $X = \sum_e X_e$.

- $E[X_e] = \frac{1}{2}$.

- By linearity of expectation,

  $E[X] = \frac{1}{2}$(total #edges).

Conditional expectation method

Let us have already decided fate of nodes 1 to i-1.

Let X = number of edges crossing cut if we place rest of nodes into U or V <u>at random</u>.

So, $E[X] = \frac{1}{2} E[\, X \mid \text{node } i \text{ is put into } U \,] + \frac{1}{2} E[\, X \mid \text{node } i \text{ is } \underline{\text{not}} \text{ put into } U \,]$

One of the terms on the RHS is at least E[X].
Put node i into the side which makes the conditional expectation larger.

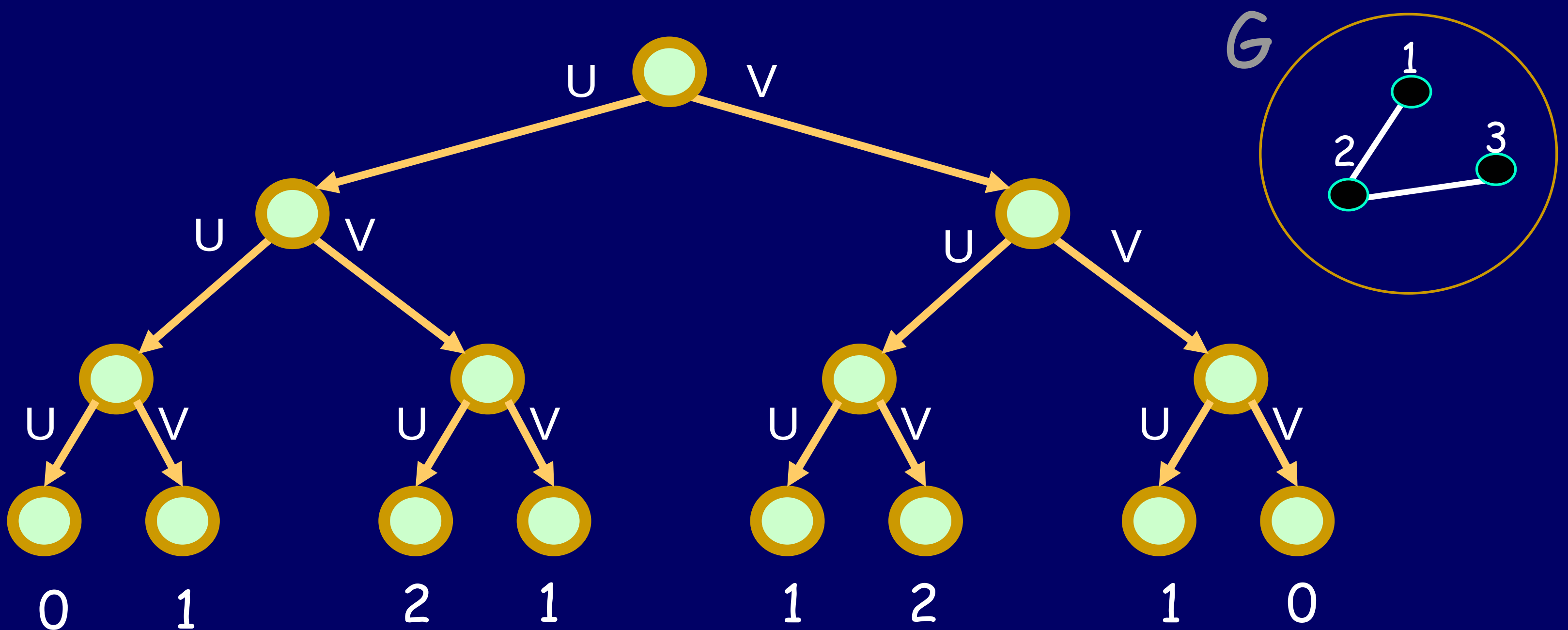
Pictorial view (important!)

View sample space S as leaves of choice tree

the i^{th} choice is where to put node i.

Label leaf by value of X

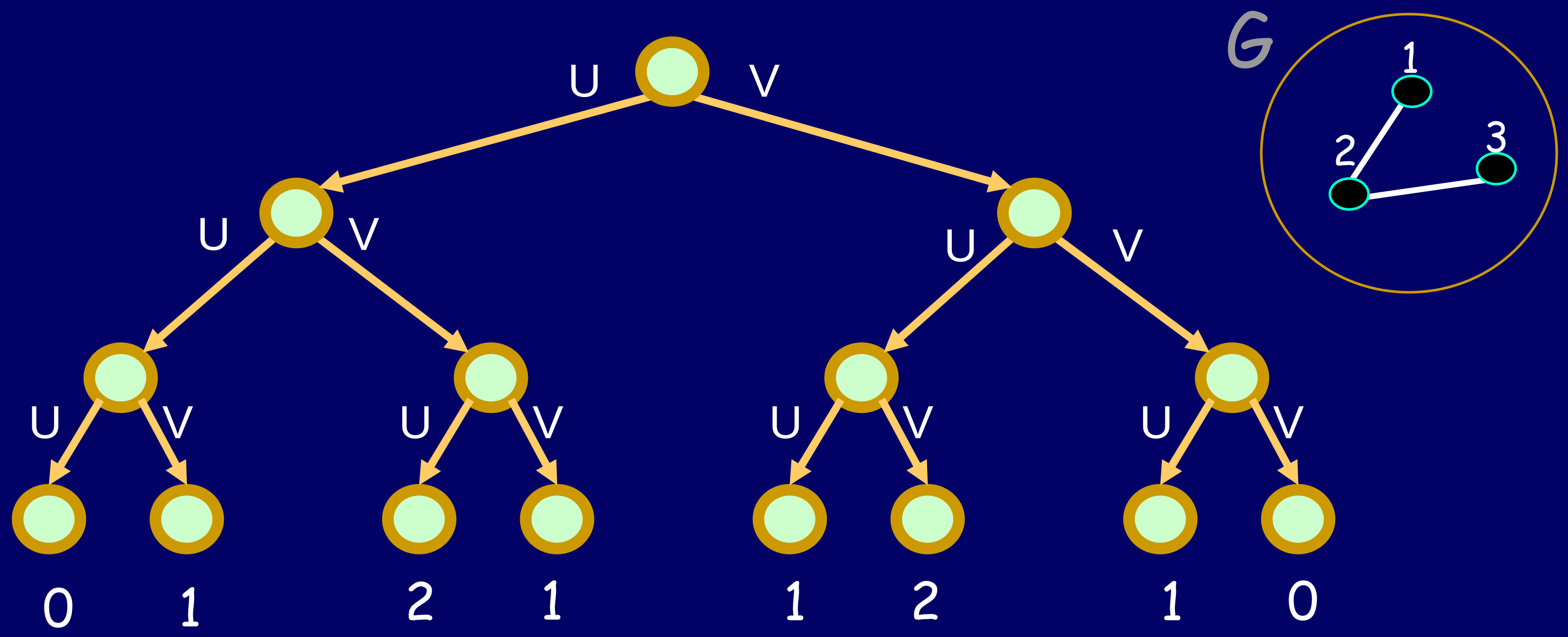
hence E[X] = avg leaf value.



Pictorial view (important!)

If A is some node (event corresponding to choices made already), then $E[X \mid A]$ = average value of leaves under it.

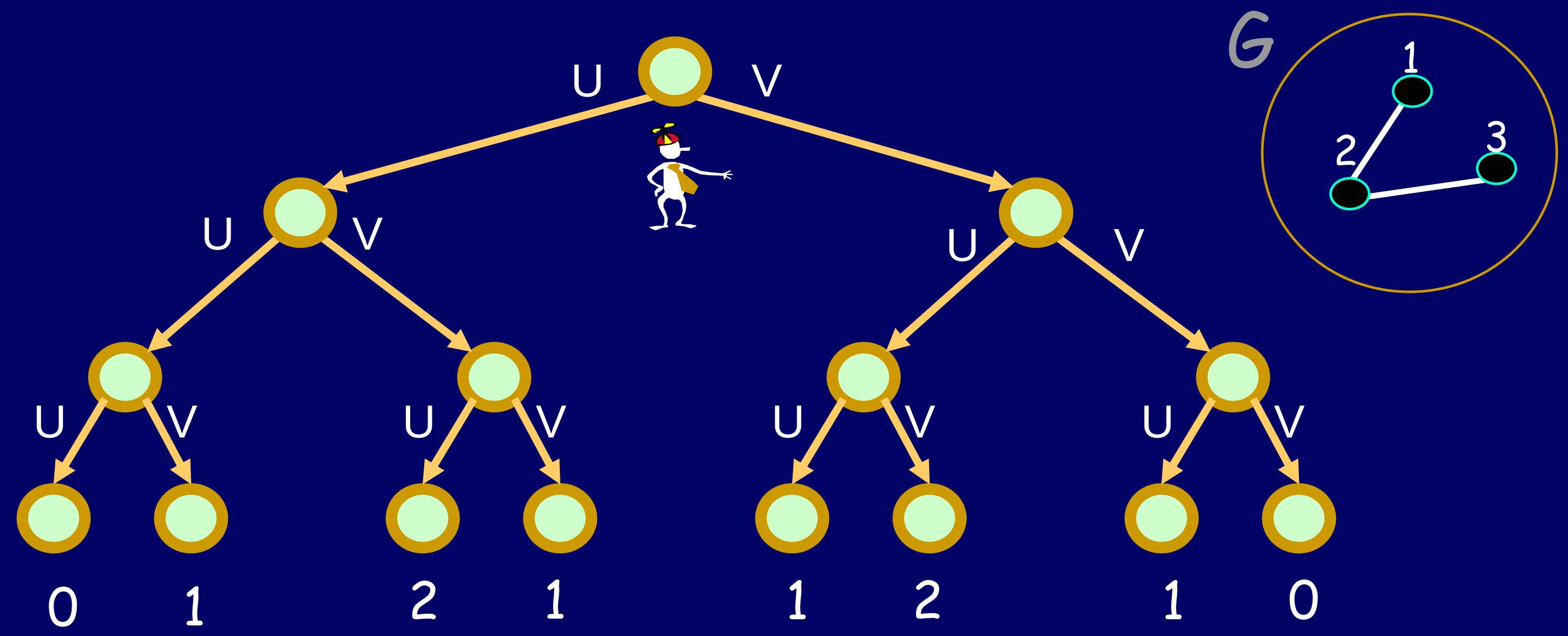
$\Rightarrow$ Algorithm = greedily go to side maximizing $E[X \mid A]$



Pictorial view (important!)

If A is some node (event corresponding to choices made already), then E[X | A] = average value of leaves under it.

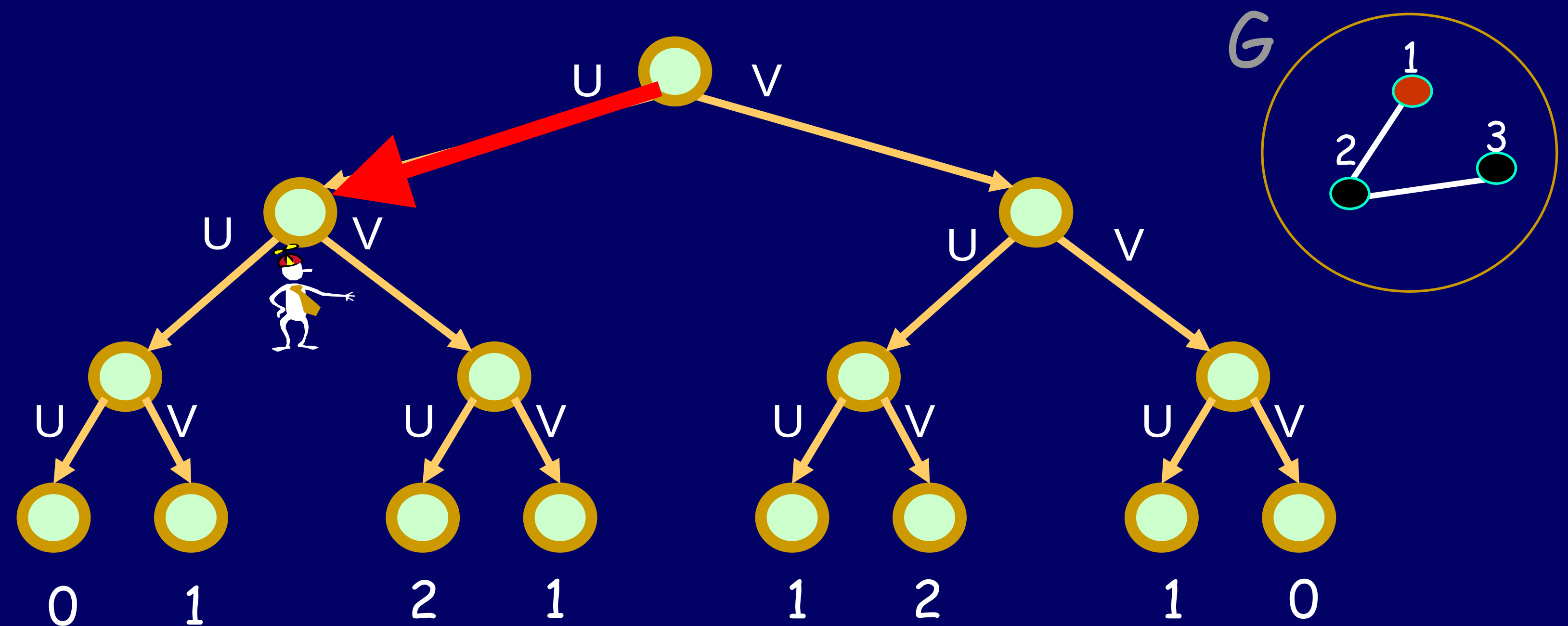
$\Rightarrow$ Algorithm = greedily go to side maximizing E[X | A]



Pictorial view (important!)

If A is some node (event corresponding to choices made already), then E[X | A] = average value of leaves under it.

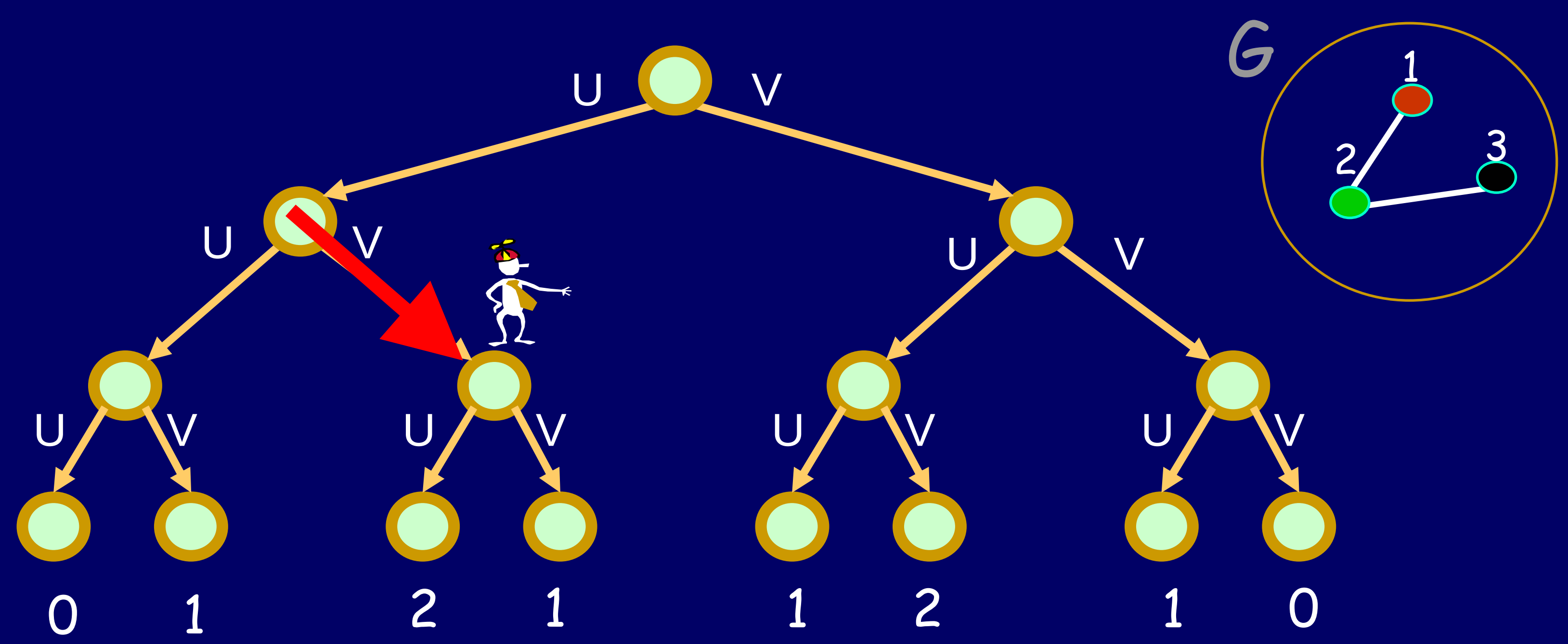
$\Rightarrow$ Algorithm = greedily go to side maximizing E[X | A]



Pictorial view (important!)

If A is some node (event corresponding to choices made already), then E[X | A] = average value of leaves under it.

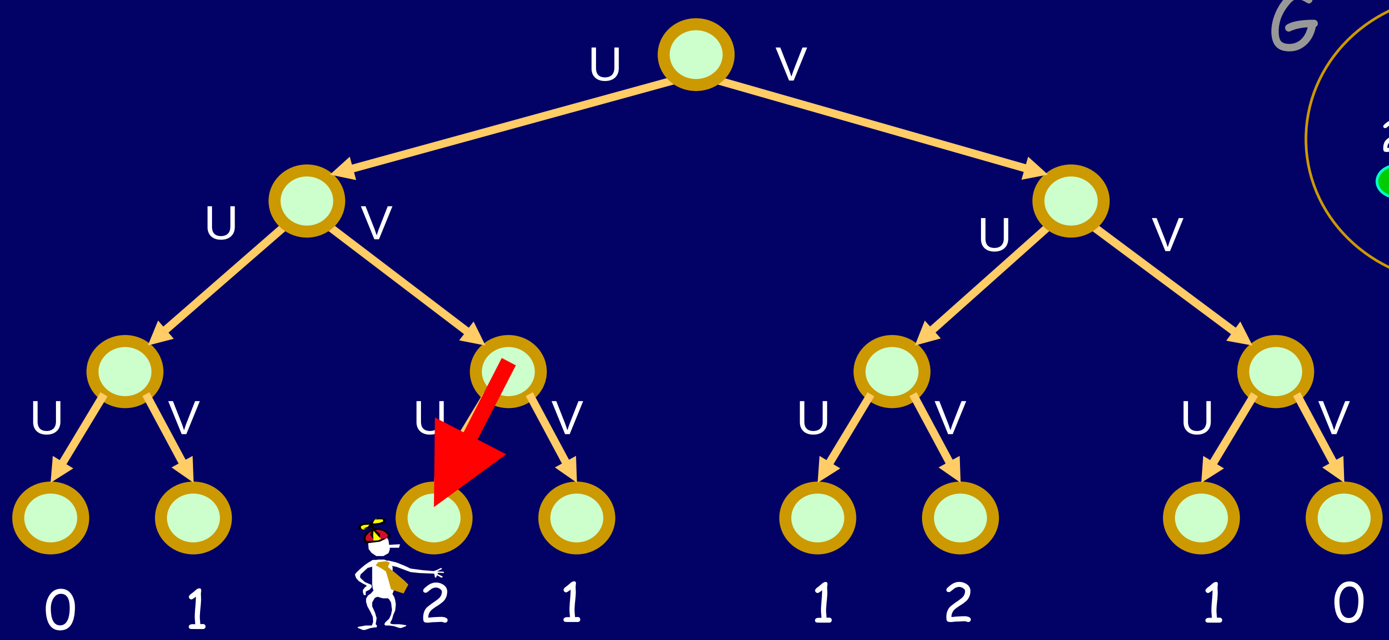
⇒ Algorithm = greedily go to side maximizing E[X | A]

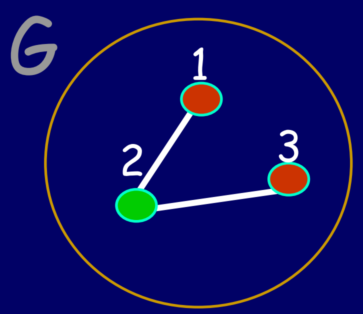


Pictorial view (important!)

If A is some node (event corresponding to choices made already), then E[X | A] = average value of leaves under it.

$\Rightarrow$ Algorithm = greedily go to side maximizing E[X | A]

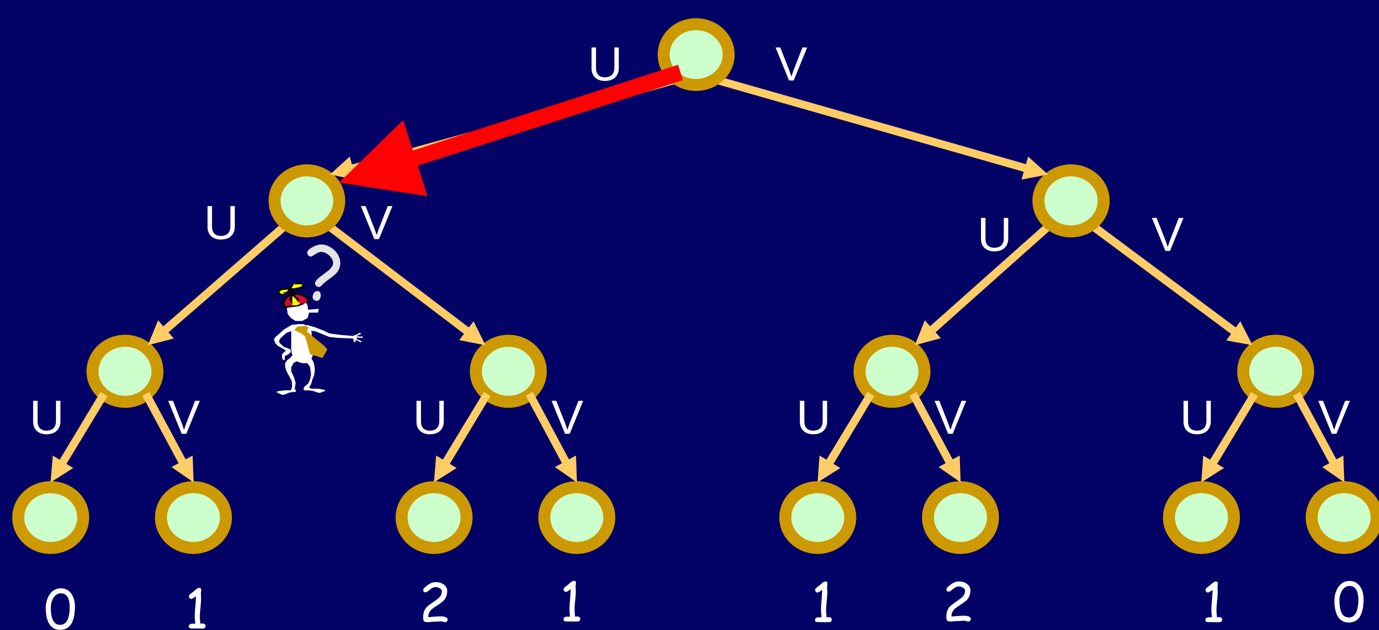




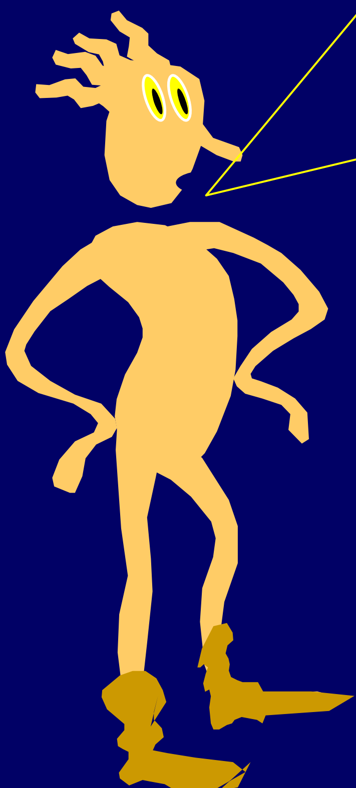
Pictorial view (important!)

How to figure out $E[X \mid A]$?   (Note: tree has 2^n leaves)

$$E[X|A] = \text{edges already cut in } A + \tfrac{1}{2}E[\text{ edges not yet determined}]$$



Conditional expectation method



When the dust settles, our algorithm is just this:

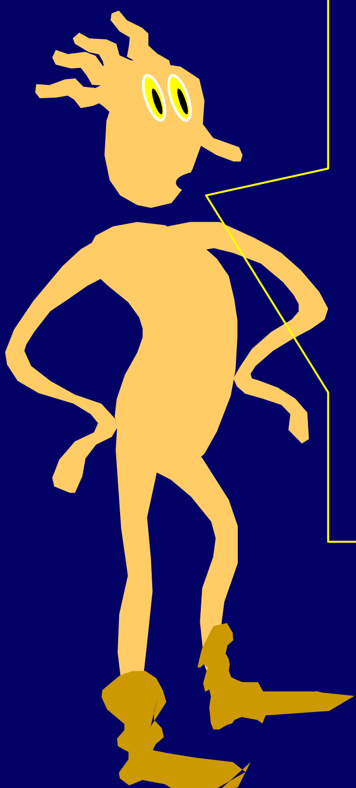
Put node i into the set that has fewer of i's neighbors so far.

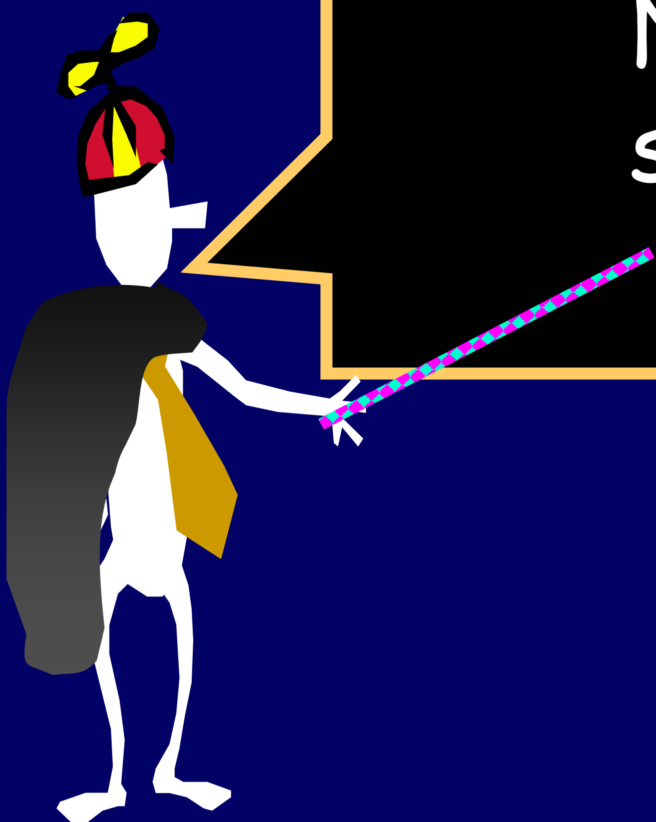
The Probabilistic Method was just useful to prove it correct.

Sometimes, though, we can't get an exact handle on these expectations.

The Probabilistic Method often gives us proofs of existence without an algorithm for finding the object.

In many cases, no efficient algorithms for finding the desired objects are known!



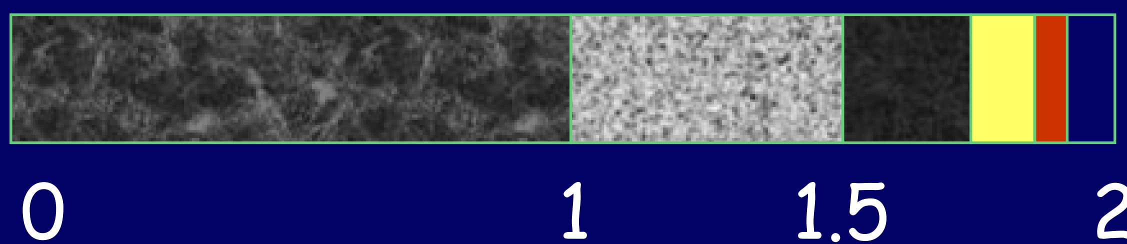


Now for something slightly different…

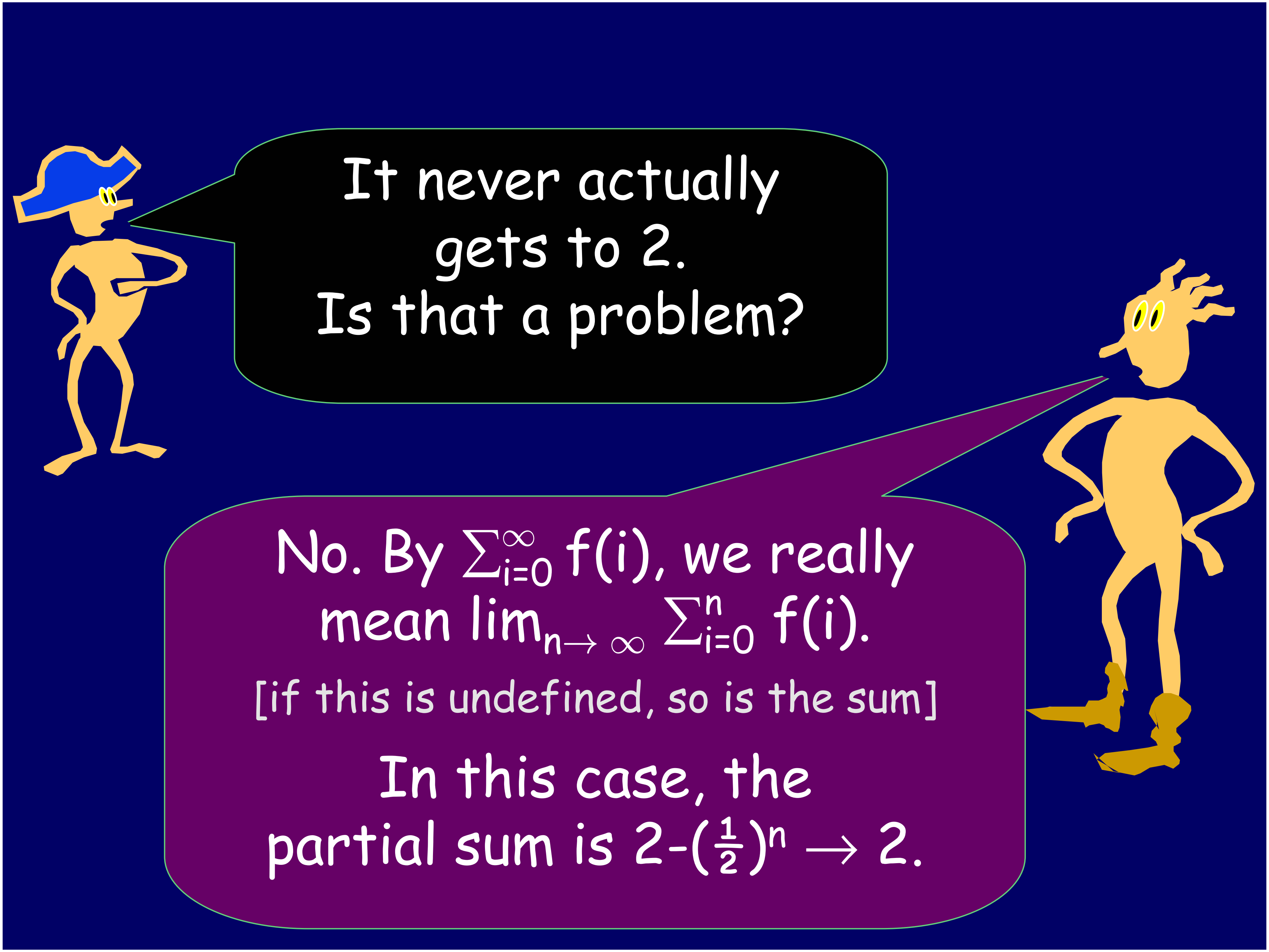
An easy question

What is $\sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{i}$?

A: 2.



It never actually
gets to 2.
Is that a problem?



It never actually gets to 2. Is that a problem?
No. By $\sum_{i=0}^{\infty} f(i)$, we really mean $\lim_{n\to\infty} \sum_{i=0}^{n} f(i)$.
[if this is undefined, so is the sum]
In this case, the partial sum is $2-(\frac{1}{2})^n \to 2$.

A related question

Suppose I flip a coin of bias p, stopping when I first get heads.

What's the chance that I:

- Flip exactly once?
  Ans: p
- Flip exactly two times?
  Ans: $(1-p)\, p$
- Flip exactly k times?
  Ans: $(1-p)^{k-1}\, p$
- Eventually stop?
  Ans: 1. (assuming $p>0$)

A related question

Pr(flip once)
+ Pr(flip 2 times)
+ Pr(flip 3 times) + ...
= 1.

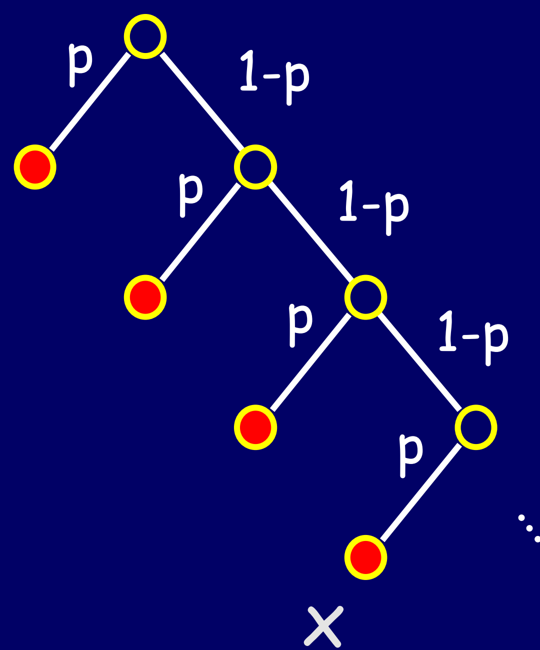
So, $p + (1-p)p + (1-p)^2 p + (1-p)^3 p + \ldots = 1$

Or, using $q = 1-p$,

$$\sum_{i=0}^{\infty} q^i = \frac{1}{1-q}.$$

$p > 0$
$\Rightarrow q < 1$

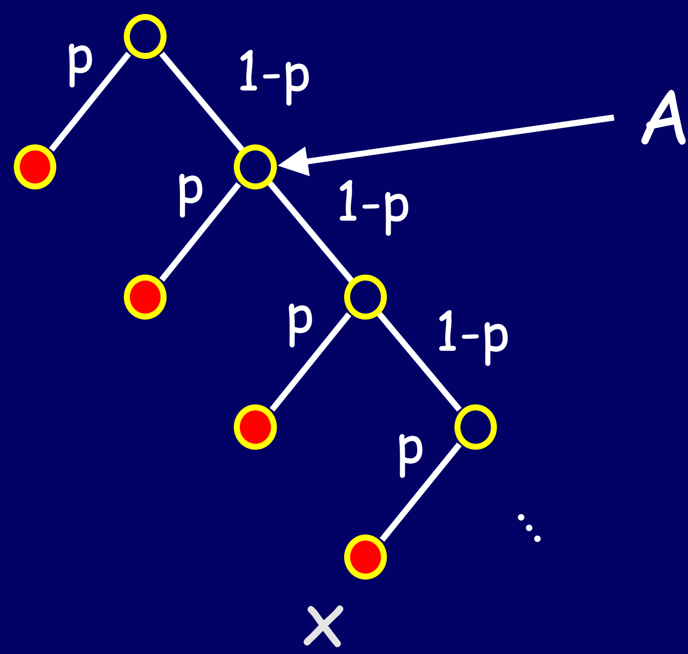
Pictorial view of coin tossing



Sample space S = leaves in this tree.
Pr(x) = product of edges on path to x.

If $p>0$, Pr[not halted by time n] $\to 0$ as $n \to \infty$.

Use to reason about expectations too



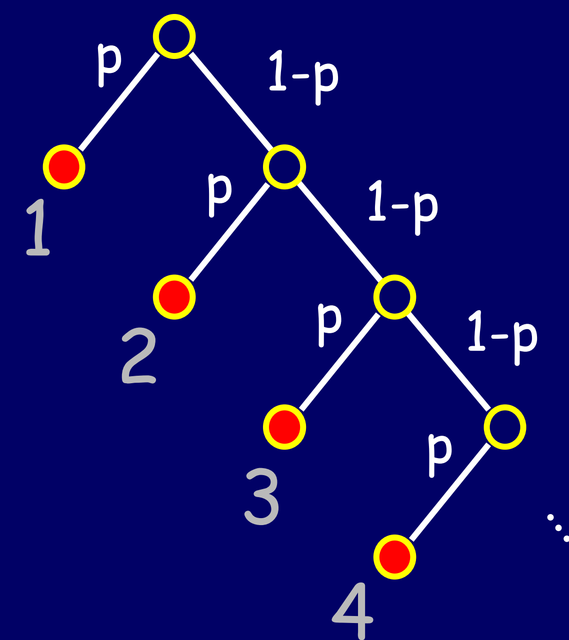
$Pr(x|A)$ = product of edges on path from A to x.

(It is as if we started the game at A.)

$E[X] = \sum_x Pr(x)\, X(x)$.

$E[X|A] = \sum_{x \in A} Pr(x|A) X(x)$.

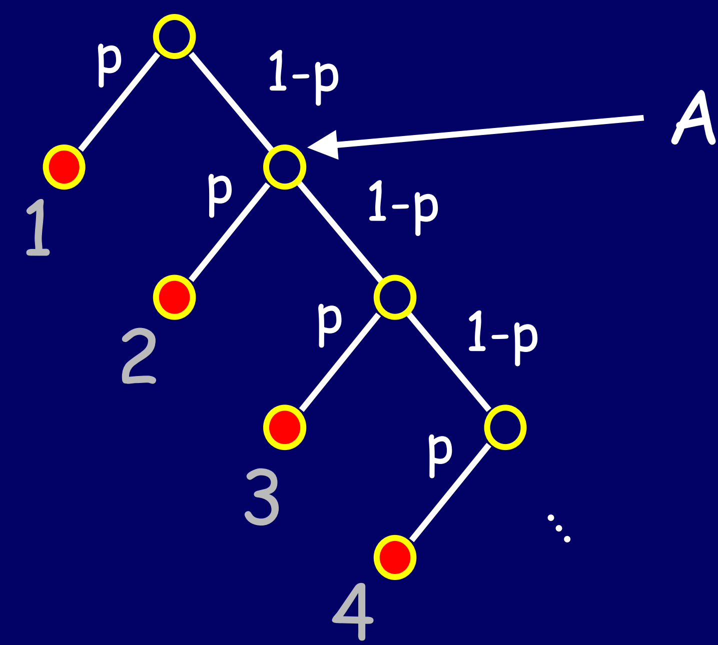
Use to reason about expectations too



Flip bias-p coin until you see heads.
What is expected number of flips?

$$\sum_k k \Pr(X = k) = \sum_k k \left[(1-p)^{k-1} p\right]$$

Use to reason about expectations too



Let X = # flips. Let A = event that 1st flip is tails.

$$E[X] = E[X|\neg A]\,Pr(\neg A) + E[\,X|A\,]\,Pr(A)$$
$$= 1^*p \;+\; (1 + E[X])^*(1\text{-}p).$$

Solves to p E[X] = p + (1-p) = 1, so E[X] = 1/p.

Infinite Probability spaces

Note:

We are using infinite probability spaces, but we really only defined things for finite spaces so far.

Infinite probability spaces can be weird.

Luckily, in CS we will almost always look at spaces that can be viewed as choice trees with

$$\Pr(\text{haven't halted by time } t) \to 0 \text{ as } t \to \infty.$$

More generally

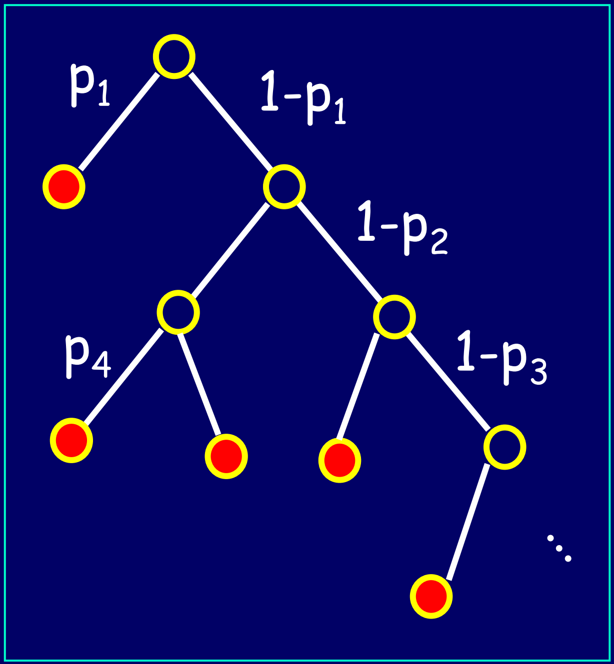
Let S be sample space we can view as leaves of a choice tree.

S_n = leaves at depth $\leq n$

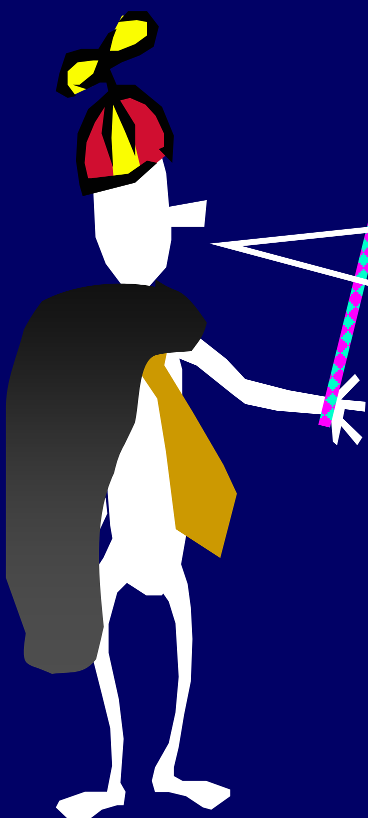
Let A be some event,
  and $A_n = A \cap S_n$

If $\lim_{n\to\infty} \Pr(S_n) = 1$, can define

$$\Pr(A) = \lim_{n\to\infty} \Pr(A_n)$$



Setting that doesn't fit our model



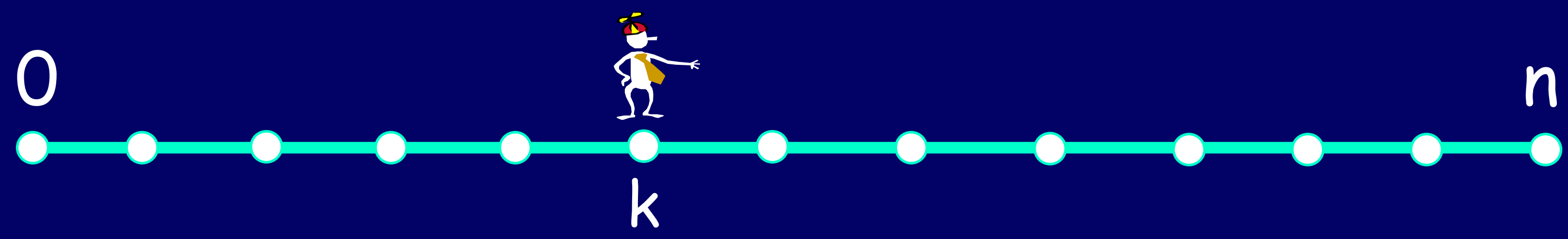
Flip a coin until
#(heads) > 2 × #(tails)

There's a reasonable chance this will <u>never</u> stop...

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:

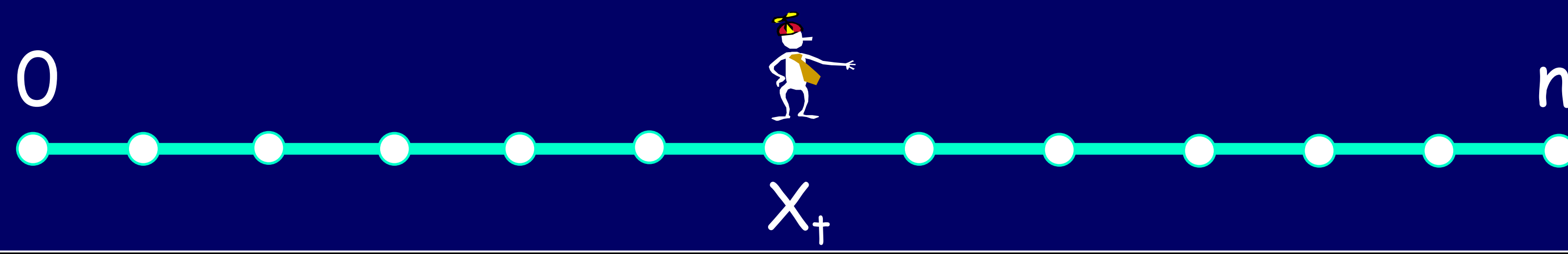
what is your expected amount of money at time t?

Let X_t be a R.V. for the amount of money 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 a RV for the change in your money at time i.)

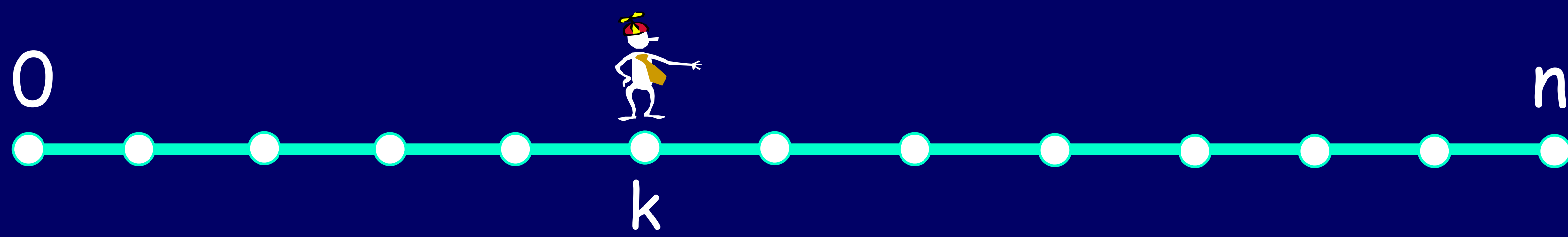
$E[\delta_i] = 0$, since $E[\delta_i | A] = 0$ for all situations A at time i.

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 ?

$E[X_t] = k.$

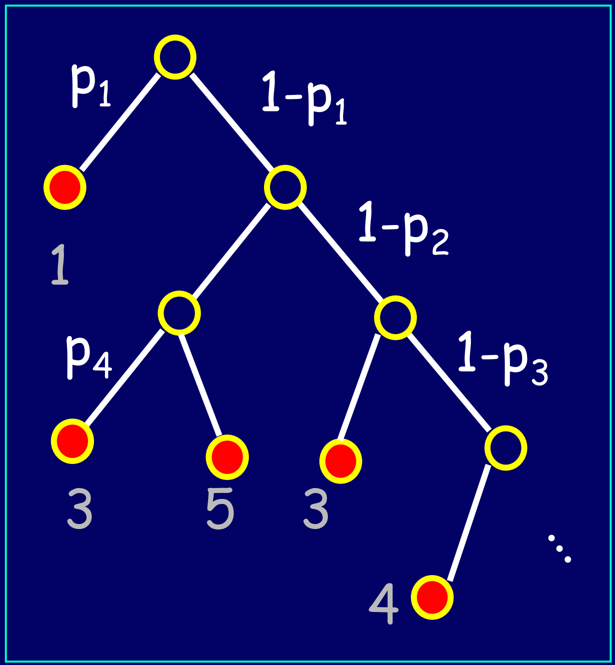
$$
\begin{aligned}
E[X_t] &= E[X_t \mid X_t = 0] \times Pr(X_t = 0) && 0 \\
&\quad + E[X_t \mid X_t = n] \times Pr(X_t = n) && + n \times Pr(X_t = n) \\
&\quad + E[X_t \mid \text{neither}] \times Pr(\text{neither}) && + (\text{something}_t \\
& && \quad \times Pr(\text{neither}))
\end{aligned}
$$

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

Hence $Pr(X_t = n) \to k/n$.

Expectations in infinite spaces

Let S be sample space we can view as leaves of a choice tree.



S_n = leaves at depth $\leq n$

Let $\lim_{n\to\infty} \Pr(S_n) = 1$

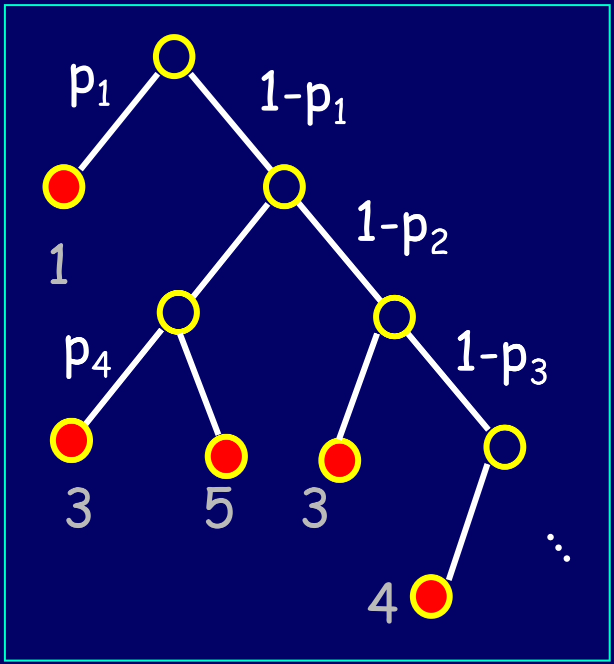
Define

$E[X] = \lim_{n\to\infty} \sum_{x \in S_n} X(x) \Pr(x)$

If this limit is undefined, then the expectation is undefined.

Expectations in infinite spaces

Let S be sample space we can view as leaves of a choice tree.



S_n = leaves at depth $\le n$

Let $\lim_{n\to\infty} Pr(S_n) = 1$

Define

$E[X] = \lim_{n\to\infty} \sum_{x \in S_n} X(x)\, Pr(x)$

Can get weird: so we want all the conditional expectations $E[X|A]$ to exist as well.

Boys and Girls

A country has the following law:

Each family must have children until they have a girl, and then they must have no more children.

What is the expected number of boys?

The expected number of girls?

References

The Probabilistic Method

Noga Alon and Joel Spencer