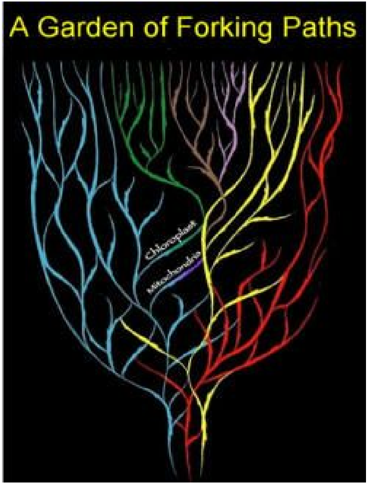
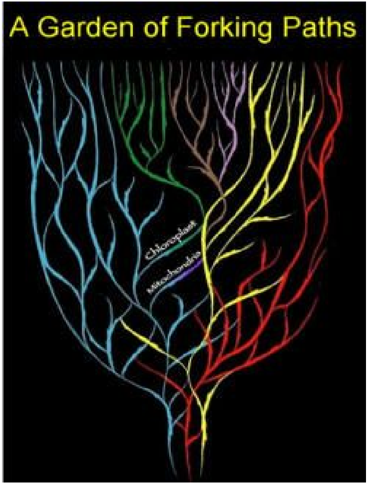
Learning Goals

- Basic definitions of finite probabilities: sample space, probability, events
- State and apply union bound.
- Define independence, and apply its properties in probability calculations
- Contention resolution with random access, and analysis of its efficiency

Borges's Garden of Forking Paths

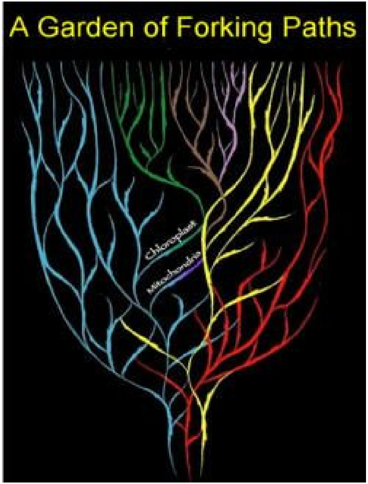


Borges's Garden of Forking Paths



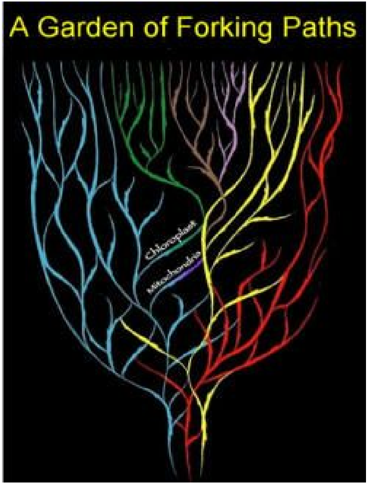
- Leaves are realizations of the world.

Borges's Garden of Forking Paths



- Leaves are realizations of the world.
- "Sample space" is the set of those realizations.

Borges's Garden of Forking Paths



- Leaves are realizations of the world.
- "Sample space" is the set of those realizations.
- A probability space is defined by weights on those realizations.

Discrete/Finite Probability Space

- Finite sample space: Ω (intuitively, the set of all realizable outcomes)

Discrete/Finite Probability Space

- Finite sample space: Ω (intuitively, the set of all realizable outcomes)
- Each point (outcome) $i \in \Omega$ has a *probability mass* $p(i) \geq 0$. We require $\sum_i p(i) = 1$.

Discrete/Finite Probability Space

- Finite sample space: Ω (intuitively, the set of all realizable outcomes)
- Each point (outcome) $i \in \Omega$ has a *probability mass* $p(i) \geq 0$. We require $\sum_i p(i) = 1$.
- An *event* $\mathcal{E}$ is a subset of Ω.

Discrete/Finite Probability Space

- Finite sample space: Ω (intuitively, the set of all realizable outcomes)
- Each point (outcome) $i \in \Omega$ has a *probability mass* $p(i) \geq 0$. We require $\sum_i p(i) = 1$.
- An *event* $\mathcal{E}$ is a subset of Ω.
- $\Pr[\mathcal{E}] = \sum_{i \in \mathcal{E}} p(i)$.

Example

- Let Ω be the set of outcomes of two rolls of a die. Then $|\Omega| = 36$.

Discrete/Finite Probability Space

- Finite sample space: Ω (intuitively, the set of all realizable outcomes)
- Each point (outcome) $i \in \Omega$ has a *probability mass* $p(i) \geq 0$. We require $\sum_i p(i) = 1$.
- An *event* $\mathcal{E}$ is a subset of Ω.
- $\Pr[\mathcal{E}] = \sum_{i \in \mathcal{E}} p(i)$.

Example

- Let Ω be the set of outcomes of two rolls of a die. Then $|\Omega| = 36$.
- If everything is fair, then each outcome has probability mass $1/36$.

Discrete/Finite Probability Space

- Finite sample space: Ω (intuitively, the set of all realizable outcomes)
- Each point (outcome) $i \in \Omega$ has a *probability mass* $p(i) \geq 0$. We require $\sum_i p(i) = 1$.
- An *event* $\mathcal{E}$ is a subset of Ω.
- $\Pr[\mathcal{E}] = \sum_{i \in \mathcal{E}} p(i)$.

Example

- Let Ω be the set of outcomes of two rolls of a die. Then $|\Omega| = 36$.
- If everything is fair, then each outcome has probability mass $1/36$.
- Let $\mathcal{E}$ be the event that the sum of the two numbers is 11, then $\mathcal{E} = \{(6,5),(5,6)\}$, so $\Pr[\mathcal{E}] = 1/18$.

Set operations on events

- Let A and B be two events of a probability space.

Set operations on events

- Let A and B be two events of a probability space.
- $\overline{A}$, the complement of A, is the event that event A does not happen, and $\Pr[\overline{A}] = 1 - \Pr[A]$.

Set operations on events

- Let A and B be two events of a probability space.
- $\overline{A}$, the complement of A, is the event that event A does not happen, and $\Pr[\overline{A}] = 1 - \Pr[A]$.
- $A \cup B$ is the event that at least one of A and B happens.

Proposition (Union Bound)

$\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$.

Set operations on events

- Let A and B be two events of a probability space.
- $\overline{A}$, the complement of A, is the event that event A does not happen, and $\Pr[\overline{A}] = 1 - \Pr[A]$.
- $A \cup B$ is the event that at least one of A and B happens.

Proposition (Union Bound)

$\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$.

- $A \cap B$ is the event that both A and B happen.

Definition

A and B are said to be *independent* if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Set operations on events

- Let A and B be two events of a probability space.
- $\overline{A}$, the complement of A, is the event that event A does not happen, and $\Pr[\overline{A}] = 1 - \Pr[A]$.
- $A \cup B$ is the event that at least one of A and B happens.

Proposition (Union Bound)

$\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$.

- $A \cap B$ is the event that both A and B happen.

Definition

A and B are said to be *independent* if $\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]$.

Exercise: If A and B are independent, then so are $\overline{A}$ and B, and so are $\overline{A}$ and $\overline{B}$.

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.
    - A potentially infinite sequence of coin flips

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.
  - A potentially infinite sequence of coin flips
  - A number from $[0, 1]$ uniformly at random

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.
  - A potentially infinite sequence of coin flips
  - A number from $[0, 1]$ uniformly at random
- The probability of a "leaf" in such a space is often 0.

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.
  - A potentially infinite sequence of coin flips
  - A number from $[0, 1]$ uniformly at random
- The probability of a "leaf" in such a space is often 0.
- One can define probability of events in fairly intuitive ways, satisfying the following axioms of probability:

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.
  - A potentially infinite sequence of coin flips
  - A number from $[0, 1]$ uniformly at random
- The probability of a "leaf" in such a space is often 0.
- One can define probability of events in fairly intuitive ways, satisfying the following axioms of probability:
  1. $\forall$ "measurable" event A, $\Pr[A] \geq 0$.
  2. $\Pr[\Omega] = 1$.
  3. for countably many disjoint events $A_1, A_2, \cdots, \Pr[\Cup_i A_i] = \sum_i \Pr[(]A_i)$.

A Word on Infinite Sample Space

- Sometimes we are interested in infinite sample spaces, e.g.
  - A potentially infinite sequence of coin flips
  - A number from $[0, 1]$ uniformly at random
- The probability of a "leaf" in such a space is often 0.
- One can define probability of events in fairly intuitive ways, satisfying the following axioms of probability:
  1. $\forall$ "measurable" event A, $\Pr[A] \geq 0$.
  2. $\Pr[\Omega] = 1$.
  3. for countably many disjoint events $A_1, A_2, \cdots, \Pr[\biguplus_i A_i] = \sum_i \Pr[(]A_i)$.
- It takes measure theory to make things rigorous. We will make use of such probability spaces in very few occasions in this course.

Contention Resolution

- Set up: one server, n tasks

Contention Resolution

- Set up: one server, n tasks
- Tasks all want to use the server for a time step (we have discrete time steps)
- At each time step, each task may request the server:
  - If exactly one task requests the server, the task gets served successfully;
  - If more than one tasks request the server, clash and no task gets served in that step (but later steps are not affected).

Contention Resolution

- Set up: one server, n tasks
- Tasks all want to use the server for a time step (we have discrete time steps)
- At each time step, each task may request the server:
  - If exactly one task requests the server, the task gets served successfully;
  - If more than one tasks request the server, clash and no task gets served in that step (but later steps are not affected).
- We would like that all tasks to get served fast.
- Trivial if the tasks can agree on some ordering and requests the service one by one.

Contention Resolution

- Set up: one server, n tasks
- Tasks all want to use the server for a time step (we have discrete time steps)
- At each time step, each task may request the server:
  - If exactly one task requests the server, the task gets served successfully;
  - If more than one tasks request the server, clash and no task gets served in that step (but later steps are not affected).
- We would like that all tasks to get served fast.
- Trivial if the tasks can agree on some ordering and requests the service one by one.
- Problem: The tasks cannot talk with each other and there is no central authority.

Contention Resolution

- Set up: one server, n tasks
- Tasks all want to use the server for a time step (we have discrete time steps)
- At each time step, each task may request the server:
  - If exactly one task requests the server, the task gets served successfully;
  - If more than one tasks request the server, clash and no task gets served in that step (but later steps are not affected).
- We would like that all tasks to get served fast.
- Trivial if the tasks can agree on some ordering and requests the service one by one.
- Problem: The tasks cannot talk with each other and there is no central authority.
- **Randomized strategy:** In each time step, each task requests with some small probability p, *independently*.

Initial analysis

- Let $A[i, t]$ denote the *event* that task i sends a request at time t. Then $\Pr[A[i, t]] = p$.

Initial analysis

- Let $A[i,t]$ denote the *event* that task i sends a request at time t. Then $\Pr[A[i,t]] = p$.
- Then $\overline{A[i,t]}$ is the event that task i does not request service at time t, and $\Pr[\overline{A[i,t]}] = 1 - p$.

Initial analysis

- Let $A[i, t]$ denote the *event* that task i sends a request at time t. Then $\Pr[A[i, t]] = p$.
- Then $\overline{A[i, t]}$ is the event that task i does not request service at time t, and $\Pr[\overline{A[i, t]}] = 1 - p$.
- Let $S[i, t]$ denote the event that task i sends a request at time t *and* gets served, then

$$\Pr\left[S[i, t]\right] = \Pr\left[A[i, t] \cap \bigcap_{j \neq i} \overline{A[j, t]}\right] = p(1-p)^{n-1}.$$

The last equality comes from independence.

Initial analysis

- Let $A[i,t]$ denote the *event* that task i sends a request at time t. Then $\Pr[A[i,t]] = p$.
- Then $\overline{A[i,t]}$ is the event that task i does not request service at time t, and $\Pr[\overline{A[i,t]}] = 1 - p$.
- Let $S[i,t]$ denote the event that task i sends a request at time t *and* gets served, then

$$\Pr\left[S[i,t]\right] = \Pr\left[A[i,t] \cap \bigcap_{j \neq i} \overline{A[j,t]}\right] = p(1-p)^{n-1}.$$

  The last equality comes from independence.
- To maximize $\Pr[S[i,t]]$, set $p = 1/n$.

Rate of success at each time step

We set p to maximize $\Pr[S[i,t]]$ to $\frac{1}{n}(1-\frac{1}{n})^{n-1}$. How good is this?

Rate of success at each time step

We set p to maximize $\Pr[S[i,t]]$ to $\frac{1}{n}(1-\frac{1}{n})^{n-1}$. How good is this?

Proposition

1. The function $(1-\frac{1}{n})^n$ converges monotonically from $\frac{1}{4}$ up to $\frac{1}{e}$ as n increases from 2.
2. The function $(1-\frac{1}{n})^{n-1}$ converges monotonically from $\frac{1}{2}$ down to $\frac{1}{e}$ as n increases from 2.

Rate of success at each time step

We set p to maximize $\Pr[S[i,t]]$ to $\frac{1}{n}(1-\frac{1}{n})^{n-1}$. How good is this?

Proposition

1. The function $(1-\frac{1}{n})^n$ converges monotonically from $\frac{1}{4}$ up to $\frac{1}{e}$ as n increases from 2.
2. The function $(1-\frac{1}{n})^{n-1}$ converges monotonically from $\frac{1}{2}$ down to $\frac{1}{e}$ as n increases from 2.

So $1/(en) \leq \Pr[S[i,t]] \leq 1/(2n)$. Therefore $\Pr[S[i,t]]$ is asymtotically $\Theta(1/n)$.

Waiting time for a particular task to succeed

- In each round, task i succeeds with probability $\Pr[S[i, t]]$. Roughly what is the waiting time for task i to succeed (for the first time)?

Waiting time for a particular task to succeed

- In each round, task i succeeds with probability $\Pr[S[i,t]]$. Roughly what is the waiting time for task i to succeed (for the first time)?
- Answers to "roughly what is X" where X is a random quantity:
  - Give the *expectation* of X (think of it as the average): later today

Waiting time for a particular task to succeed

- In each round, task i succeeds with probability $\Pr[S[i,t]]$. Roughly what is the waiting time for task i to succeed (for the first time)?
- Answers to "roughly what is X" where X is a random quantity:
  - Give the *expectation* of X (think of it as the average): later today
  - Give a range $[a, b]$, and show that X is in $[a, b]$ with "high probability": today

Waiting time for a particular task to succeed

- In each round, task i succeeds with probability $\Pr[S[i, t]]$. Roughly what is the waiting time for task i to succeed (for the first time)?
- Answers to "roughly what is X" where X is a random quantity:
  - Give the *expectation* of X (think of it as the average): later today
  - Give a range $[a, b]$, and show that X is in $[a, b]$ with "high probability": today
  - Remark: often, the two give answers that are close. Usually, the random quantity *concentrates* around its expectation. *Tail bounds* a.k.a. *Concentration inequalities* are used to show how fast this happens.

Waiting time for a particular task to succeed

- In each round, task i succeeds with probability $\Pr[S[i,t]]$. Roughly what is the waiting time for task i to succeed (for the first time)?
- Answers to "roughly what is X" where X is a random quantity:
  - Give the *expectation* of X (think of it as the average): later today
  - Give a range $[a,b]$, and show that X is in $[a,b]$ with "high probability": today
  - Remark: often, the two give answers that are close. Usually, the random quantity *concentrates* around its expectation. *Tail bounds* a.k.a. *Concentration inequalities* are used to show how fast this happens.
- Probability with which task i does not succeed in the first t steps:

$$\Pr\left[\cap_{r=1}^{t}\overline{S[i,r]}\right] = \prod_{r=1}^{t}[1-\Pr\left[S[i,r]\right]] = \left[1-\frac{1}{n}\left(1-\frac{1}{n}\right)^{n-1}\right]^{t}.$$

Waiting time for a particular task to succeed

- Probability that a task fails in the first t steps: $[1 - \frac{1}{n}(1 - \frac{1}{n})^{n-1}]^t$.

Waiting time for a particular task to succeed

- Probability that a task fails in the first t steps: $[1 - \frac{1}{n}(1 - \frac{1}{n})^{n-1}]^t$.
- We'd like to upper bound this probability:

$$\Pr\left[\cap_{r=1}^{t}\overline{S[i,r]}\right] \leq \left[1 - \frac{1}{en}\right]^t = \left[1 - \frac{1}{en}\right]^{en\cdot\frac{t}{en}} \leq e^{-t/en}.$$

Waiting time for a particular task to succeed

- Probability that a task fails in the first t steps: $[1 - \frac{1}{n}(1 - \frac{1}{n})^{n-1}]^t$.
- We'd like to upper bound this probability:

$$\Pr\left[\cap_{r=1}^{t}\overline{S[i, r]}\right] \leq \left[1 - \frac{1}{en}\right]^t = \left[1 - \frac{1}{en}\right]^{en \cdot \frac{t}{en}} \leq e^{-t/en}.$$

- Setting t to be $enc \ln n$ for some $c > 0$, the probability of failure for the first t steps is at most n^{-c}, which vanishes as n grows.

Waiting time for a particular task to succeed

- Probability that a task fails in the first t steps: $[1 - \frac{1}{n}(1 - \frac{1}{n})^{n-1}]^t$.
- We'd like to upper bound this probability:

$$\Pr\left[\cap_{r=1}^{t}\overline{S[i,r]}\right] \le \left[1 - \frac{1}{en}\right]^t = \left[1 - \frac{1}{en}\right]^{en\cdot\frac{t}{en}} \le e^{-t/en}.$$

- Setting t to be $enc \ln n$ for some $c > 0$, the probability of failure for the first t steps is at most n^{-c}, which vanishes as n grows.
- Big picture (useful rough estimations): if we have a biased coin that gives Heads with probability $1/k$:
  - In about k independent tosses, one "expects" to see a Heads;
  - However, with constant probability, a Heads doesn't show in k tosses;
  - But if one tosses the coin $\Theta(k \log k)$ times, the probability that no Heads shows up quickly tends to 0.

Waiting time for all tasks to succeed

- Let $F[i,t]$ denote the event that task i fails in the first t steps, we have shown $\Pr[F[i,t]] \leq e^{-t/en} \leq n^{-c}$ for $t = \lceil en \cdot c \ln n \rceil$.

Waiting time for all tasks to succeed

- Let $F[i,t]$ denote the event that task i fails in the first t steps, we have shown $\Pr[F[i,t]] \leq e^{-t/en} \leq n^{-c}$ for $t = \lceil en \cdot c \ln n \rceil$.
- The event that *some* task keeps failing in the first t steps is then $\cup_{i=1}^{n} F[i,t]$.

Waiting time for all tasks to succeed

- Let $F[i,t]$ denote the event that task i fails in the first t steps, we have shown $\Pr[F[i,t]] \leq e^{-t/en} \leq n^{-c}$ for $t = \lceil en \cdot c \ln n \rceil$.
- The event that *some* task keeps failing in the first t steps is then $\cup_{i=1}^{n} F[i,t]$.

By the union bound, we have

$$\Pr\left[\cup_{i=1}^{n} F[i,t]\right] \leq \sum_{i=1}^{n} e^{-t/en} = ne^{-\frac{t}{en}}.$$

So for $t = \lceil 2en \ln n \rceil$, this is at most $\frac{1}{n}$.

Birthday Paradox

Birthday Paradox

- We currently have 35 students in this class. Let's assume their birthdays are independently distributed uniformly throughout the year, say, from 1 to 365.

Birthday Paradox

- We currently have 35 students in this class. Let's assume their birthdays are independently distributed uniformly throughout the year, say, from 1 to 365.
- If you were to bet, would you bet that some pair of students have the same birthday or not?

Birthday Paradox

- We currently have 35 students in this class. Let's assume their birthdays are independently distributed uniformly throughout the year, say, from 1 to 365.
- If you were to bet, would you bet that some pair of students have the same birthday or not?
- The probability that no two among n students have the same birthday is $\prod_{i=1}^{n-1}(1 - \frac{i}{365})$.

Birthday Paradox

- We currently have 35 students in this class. Let's assume their birthdays are independently distributed uniformly throughout the year, say, from 1 to 365.
- If you were to bet, would you bet that some pair of students have the same birthday or not?
- The probability that no two among n students have the same birthday is $\prod_{i=1}^{n-1}(1 - \frac{i}{365})$.
- A useful upper bound: for $x \in (0,1), 1 - x < e^{-x}$. So the above probability is at most $\prod_{i=1}^{n-1} e^{-i/365} = e^{-n(n-1)/730}$.

Birthday Paradox

- We currently have 35 students in this class. Let's assume their birthdays are independently distributed uniformly throughout the year, say, from 1 to 365.
- If you were to bet, would you bet that some pair of students have the same birthday or not?
- The probability that no two among n students have the same birthday is $\prod_{i=1}^{n-1}(1 - \frac{i}{365})$.
- A useful upper bound: for $x \in (0, 1), 1 - x < e^{-x}$. So the above probability is at most $\prod_{i=1}^{n-1} e^{-i/365} = e^{-n(n-1)/730}$.
- As long as $e^{-n(n-1)/730} < \frac{1}{2}$, i.e., $n \geq 23$, you should bet that some pair of students have the same birthday.