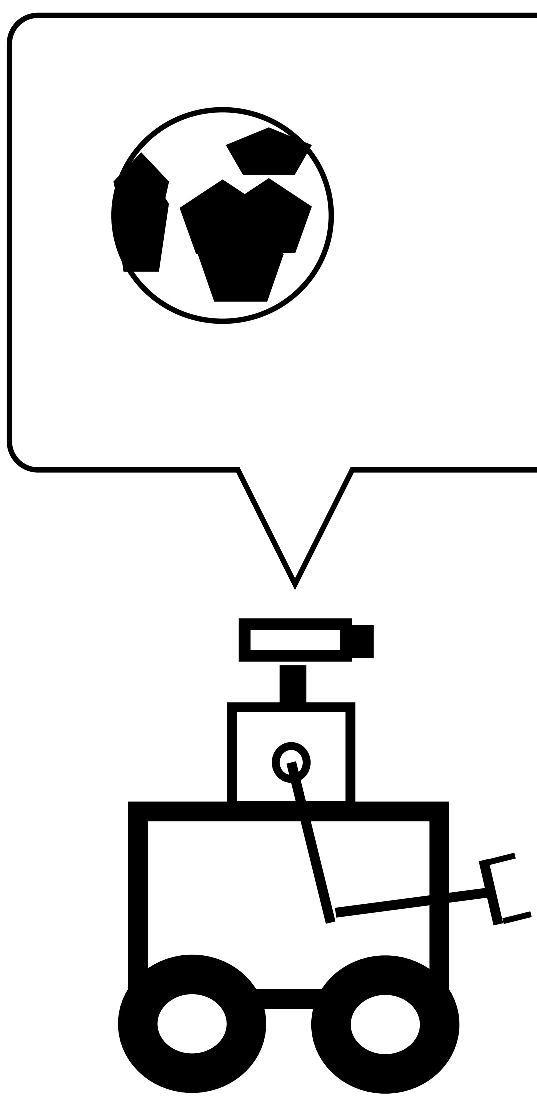
Uncertain Knowledge and Bayes' Rule

George Konidaris
gdk@cs.brown.edu

Fall 2021

Knowledge



Logic

Logical representations are based on:

- *Facts* about the world.
- Either **true** or **false**.
- *We may not know which.*
- Can be combined with logical connectives.

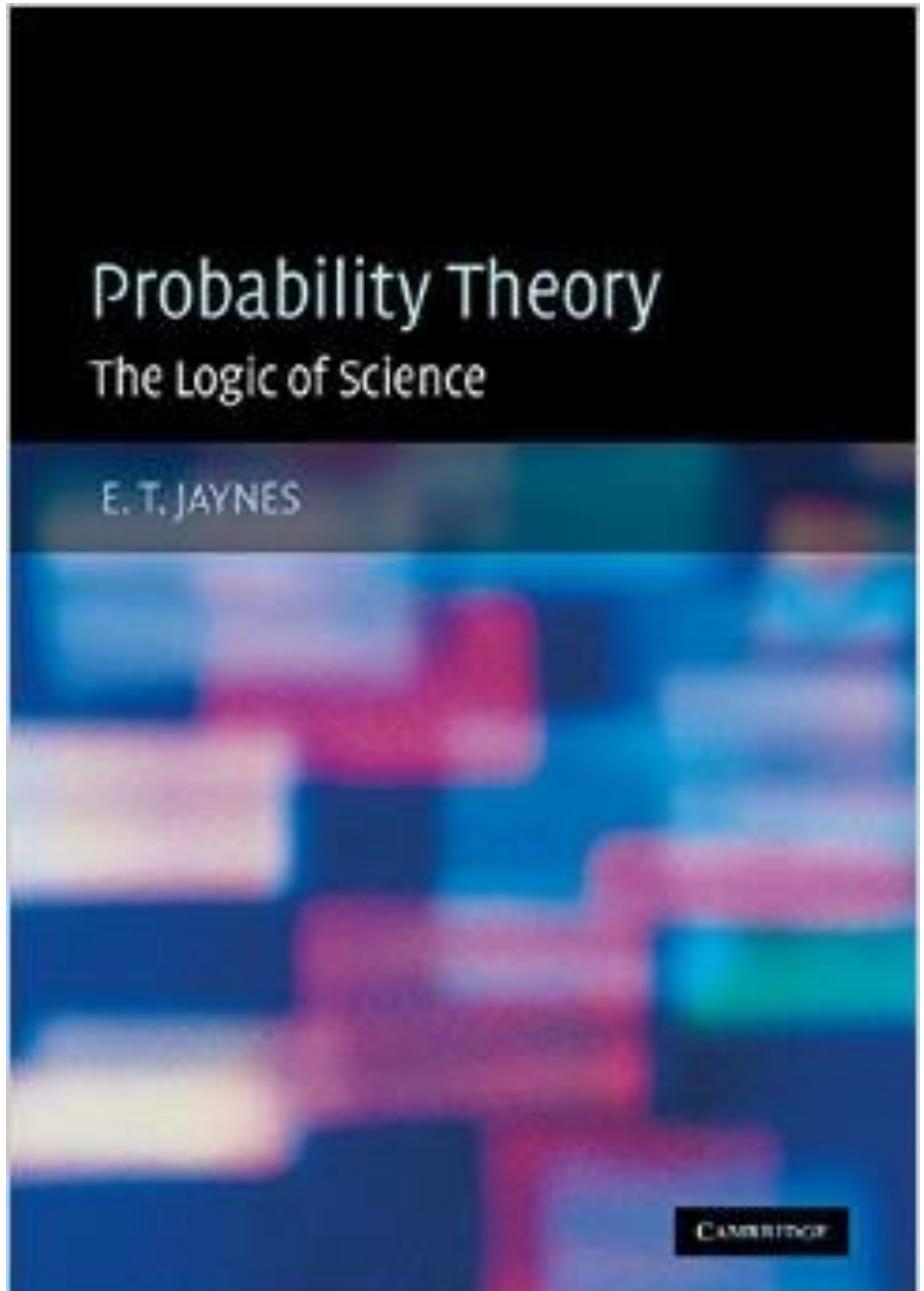
Logic inference is based on:

- What we can conclude *with certainty*.

Logic is Insufficient

The world is not deterministic.
There is no such thing as a fact.
Generalization is hard.
Sensors and actuators are noisy.
Plans fail.
Models are not perfect.
Learned models are *especially* imperfect.

$$\forall x, Fruit(x) \implies Tasty(x)$$


Probability Theory
The Logic of Science
E. T. JAYNES

Probabilities

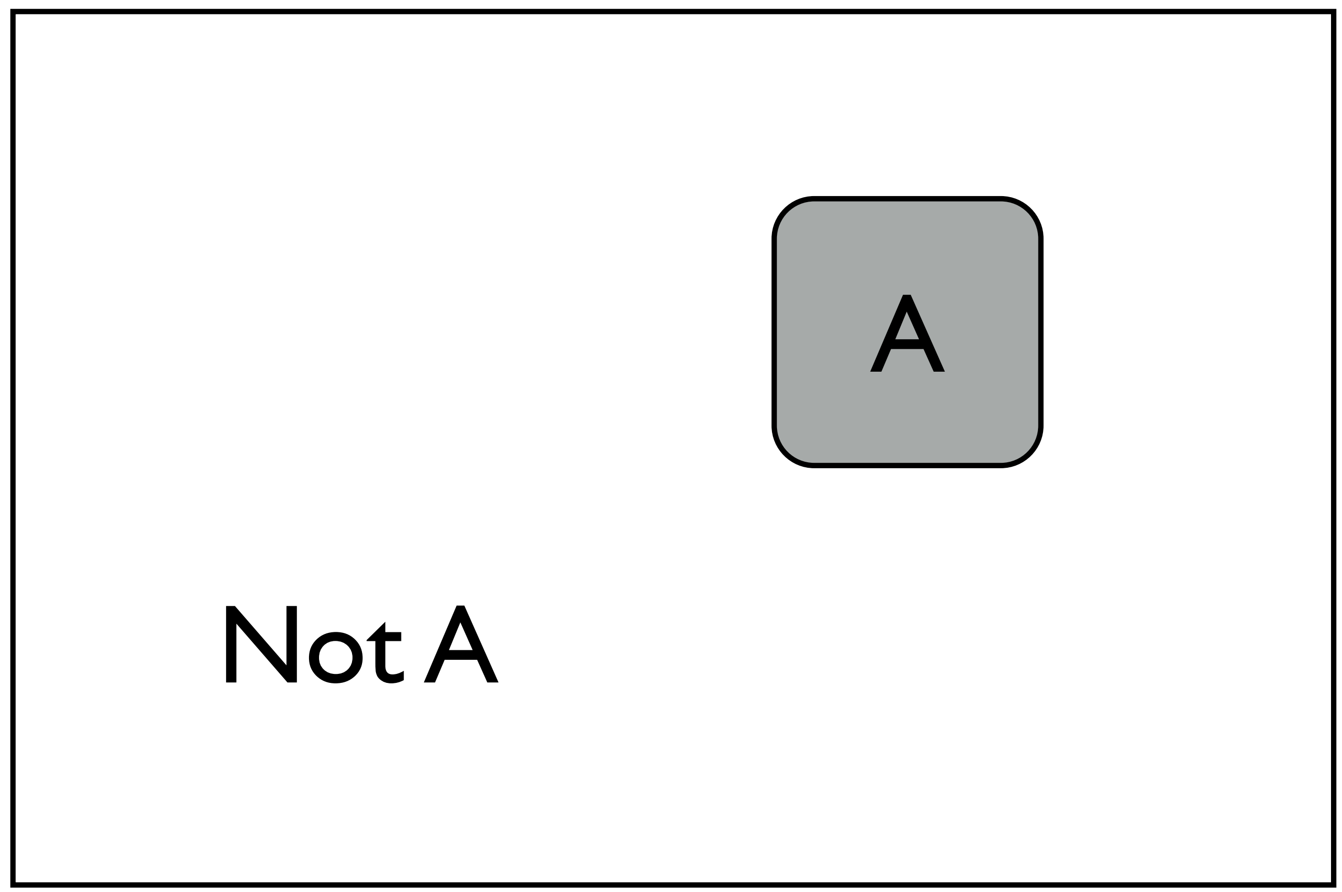
Powerful tool for reasoning about uncertainty.

Can prove that a person who holds a system of beliefs inconsistent with probability theory can be fooled.

But, we're not necessarily using them the way you would expect.

Relative Frequencies

Defined over *events*.



P(A): probability random event falls in *A*, rather than *Not A*.
Works well for dice and coin flips!

Relative Frequencies

But this feels limiting.

What is the probability that the Red Sox win this year's World Series?

- Meaningful question to ask.
- Can't count frequencies (except naively).
- Only really happens once.

In general, *all events only happen once*.

Probabilities and Beliefs

Suppose I flip a coin and hide outcome.

- What is P(Heads)?

This is a statement about *a belief*, not *the world*.
(the world is in exactly one state, with prob. 1)

Assigning truth values to probabilities is tricky - must reference speaker's *state of knowledge*.

Frequentists: probabilities come from relative frequencies.
Subjectivists: probabilities are degrees of belief.

For Our Purposes

No two events are identical, or completely unique.

Use probabilities as beliefs, but allow data (relative frequencies) to influence these beliefs.

In AI: probabilities reflect degrees of belief, given observed evidence.

We use *Bayes' Rule* to combine prior beliefs with new data.

Examples

X: RV indicating winner of Red Sox vs. Yankees game.

d(X) = {Red Sox, Yankees, tie}.

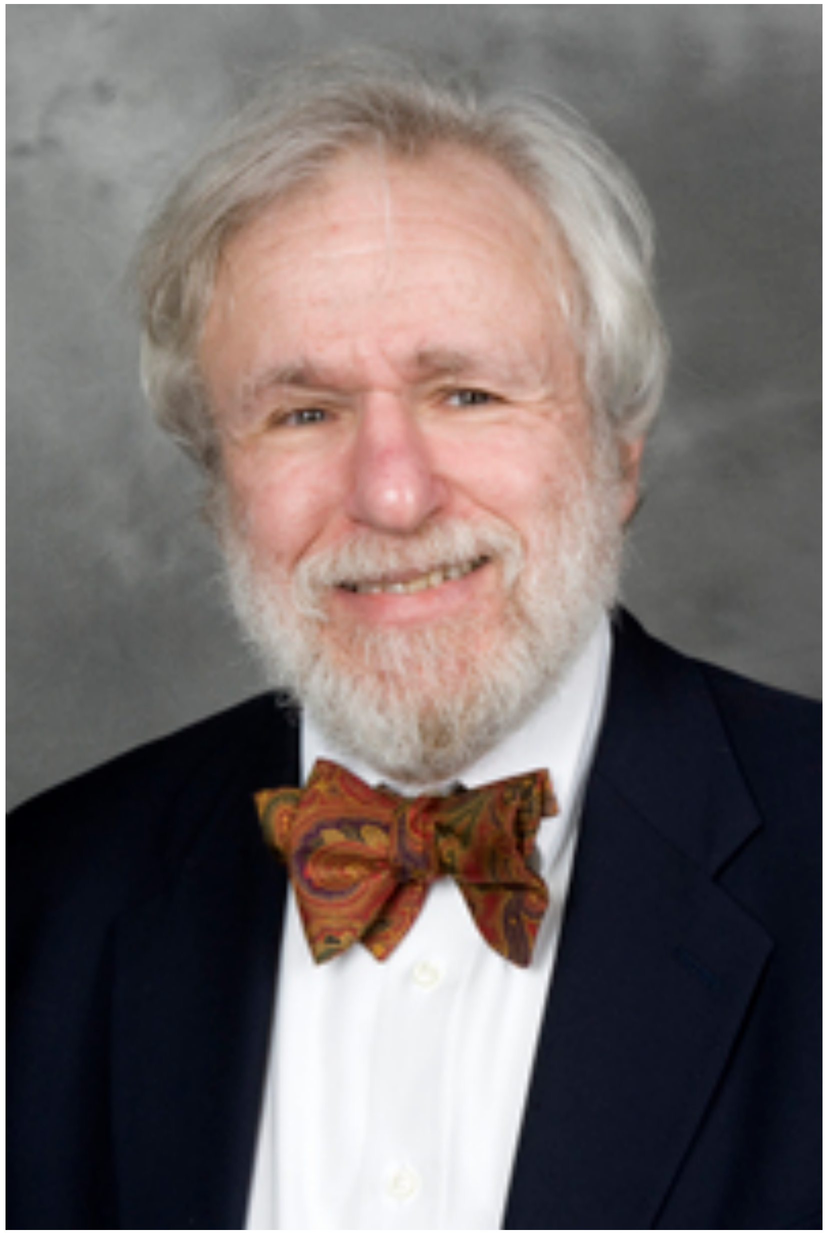
A probability is associated with each *event* in the domain:

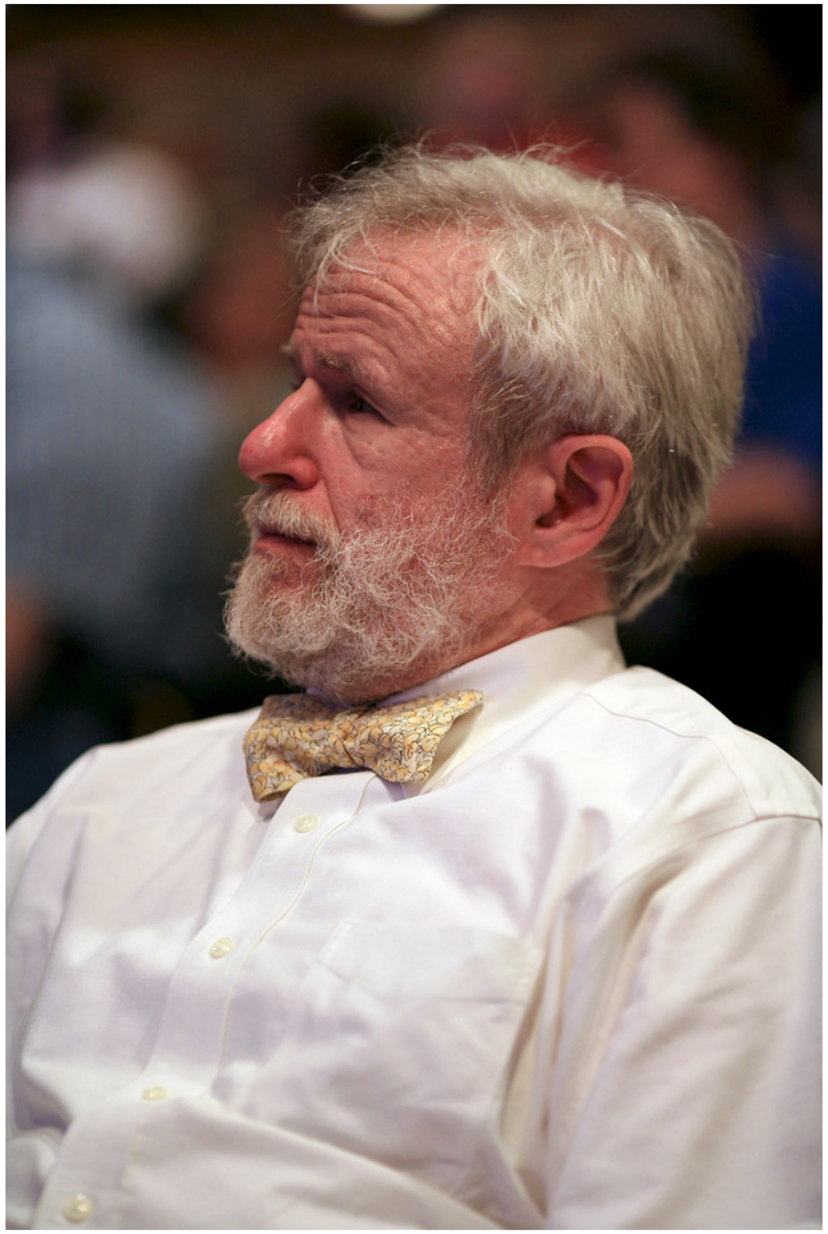
- P(X = Red Sox) = 0.8
- P(X = Yankees) = 0.19
- P(X = tie) = 0.01

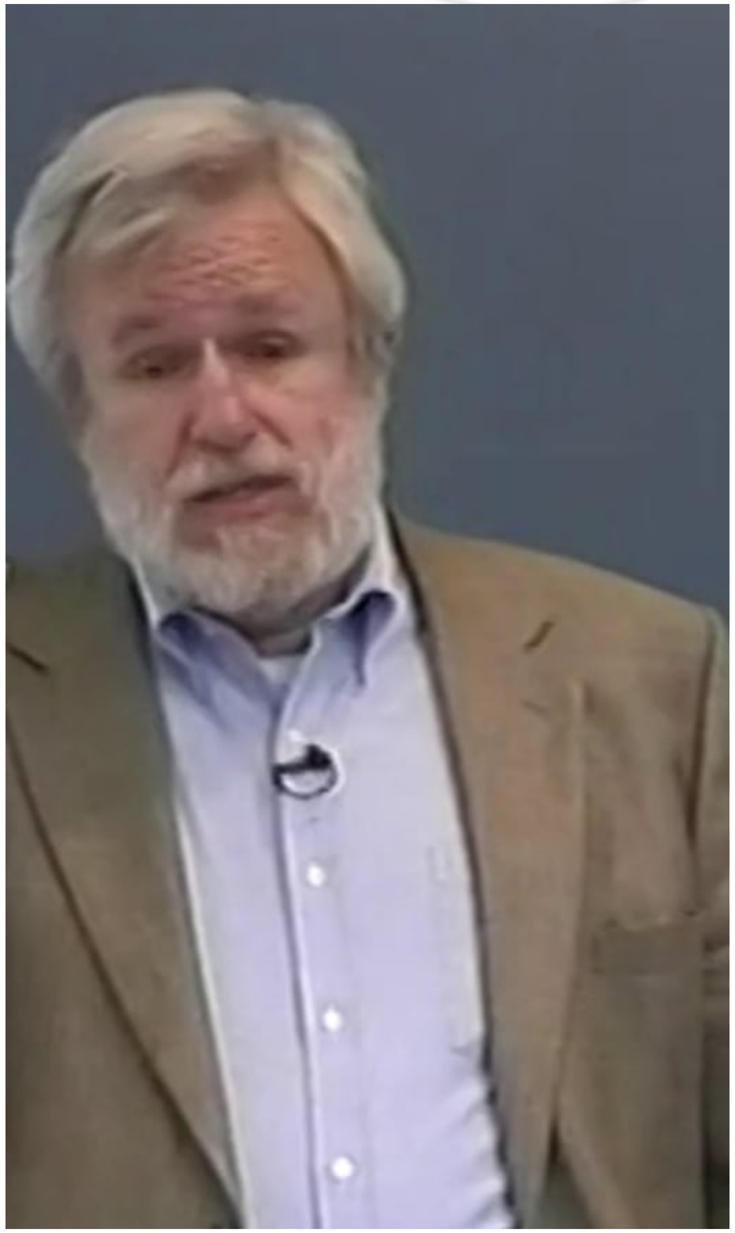
Note: probabilities over *the entire event space* must sum to 1.

Example

What is the probability that Eugene Charniak will wear a red bowtie tomorrow?

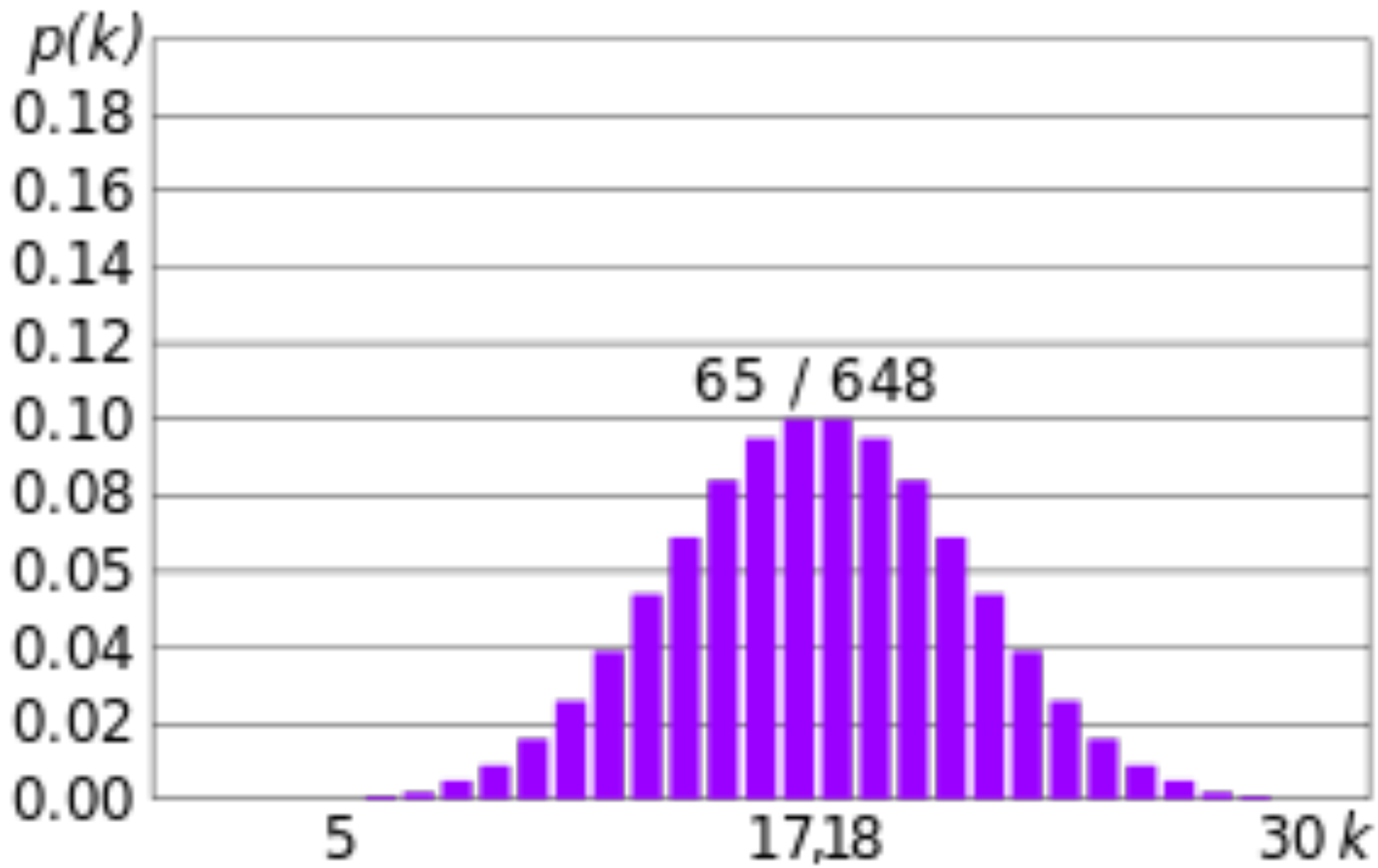






Example

How many students are sitting on the Quiet Green right now?



Joint Probability Distributions

What to do when several variables are involved?

Think about *atomic events*.

- Complete assignment of all variables.
- All possible events.
- Mutually exclusive.

RVs: Raining, Cold (both boolean):

Raining	**Cold**	**Prob.**
True	True	0.3
True	False	0.1
False	True	0.4
False	False	0.2

joint distribution

Note: *still adds up to 1*.

Joint Probability Distributions

Some analogies …

$$X \wedge Y$$

X	**Y**	P
True	True	1
True	False	0
False	True	0
False	False	0

$$X \vee Y$$

X	**Y**	P
True	True	0.33
True	False	0.33
False	True	0.33
False	False	0

$$\neg X$$

X	P
True	0
False	1

Joint Probability Distribution

Probabilities to all possible atomic events *(grows fast)*

Raining	Cold	Prob.
True	True	0.3
True	False	0.1
False	True	0.4
False	False	0.2

Can define individual probabilities in terms of JPD:
P(Raining) = P(Raining, Cold) + P(Raining, not Cold) = 0.4.

$$P(a) = \sum_{e_i \in e(a)} P(e_i)$$

Joint Probability Distribution

Simplistic probabilistic knowledge base:

- Variables of interest $X_1, \ldots, X_n$.
- JPD over $X_1, \ldots, X_n$.
- *Expresses all possible statistical information about relationships between the variables of interest.*

Inference:

- Queries over subsets of $X_1, \ldots, X_n$
- E.g., $P(X_3)$
- E.g., $P(X_3 \mid X_1)$

Conditional Probabilities

What if you have a joint probability, and you *acquire new data?*

My iPhone tells me that its cold.

What is the probability that it is raining?

Raining	**Cold**	**Prob.**
True	True	0.3
True	False	0.1
False	True	0.4
False	False	0.2

Write this as:

- P(Raining | Cold)

Conditioning

Written as:

- P(X | Y)

Here, *X* is ***uncertain***, but *Y* is ***known (fixed, given)***.

Ways to think about this:

- *X* is belief, *Y* is evidence affecting belief.
- *X* is belief, *Y* is hypothetical.
- *X* is unobserved, *Y* is observed.

Soft version of *implies*:

- $Y \implies X \approx P(X|Y) = 1$

Conditional Probabilities

We can write:

$$P(a|b) = \frac{P(a \text{ and } b)}{P(b)}$$

This tells us the probability of *a* **given only knowledge b**.

This is a probability w.r.t a **state of knowledge.**

- P(Disease | Symptom)
- P(Raining | Cold)
- P(Red Sox win | injury)

Conditional Probabilities

P(Raining | Cold)
= P(Raining and Cold)
/ P(Cold)

Raining	Cold	Prob.
True	True	0.3
True	False	0.1
False	True	0.4
False	False	0.2

… P(Cold) = 0.7
… P(Raining and Cold) = 0.3

P(Raining | Cold) ~= 0.43.

Note!
P(Raining | Cold) + P(not Raining | Cold) = 1!

Joint Distributions Are Everything

All you (statistically) need to know about $X_1 \ldots X_n$.

Classification

- $P(X_1 \mid X_2 \ldots X_n)$ ← things you know

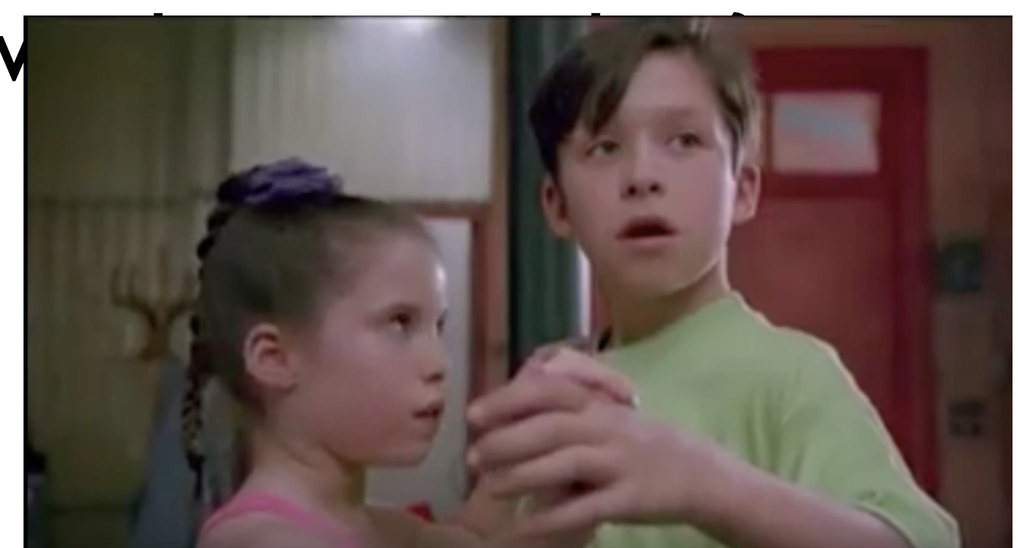
← thing you want to know

Co-occurrence

- $P(X_a, X_b)$ ← how likely are these tw

Rare event detection

- $P(X_1, \ldots, X_n)$ ←



Joint Probability Distributions

Joint probability tables …

- Grow very fast.
- Need to sum out the other variables.
- Might require lots of data.
- NOT a function of P(A) and P(B).

Independence

Critical property! But rare.

If A and B are independent:

- P(A and B) = P(A)P(B)
- P(A or B) = P(A) + P(B) - P(A)P(B)

Independence: two events don't effect each other.

- Red Sox winning world series, Andy Murray winning Wimbledon.
- Two successive, fair, coin flips.
- It is raining, and winning the lottery.
- Poker hand and date.

Independence

Are *Raining* and *Cold* independent?

Raining	**Cold**	**Prob.**
True	True	0.3
True	False	0.1
False	True	0.4
False	False	0.2

P(Raining = True) = 0.4
P(Cold = True) = 0.7

P(Raining = True, Cold = True) = ?

Independence

If independent, can break JPD into separate tables.

Raining	Prob.
True	0.6
False	0.4

Cold	Prob.
True	0.75
False	0.25

x

Raining	Cold	Prob.
True	True	0.45
True	False	0.15
False	True	0.3
False	False	0.1

Independence is Critical

Much of probabilistic knowledge representation and machine learning is concerned with identifying and leveraging independence and mutual exclusivity.

Independence is also rare. Is there a weaker type of structure we might be able to exploit?

Conditional Independence

A and B are **conditionally independent given C** if:

- P(A | B, C) = P(A | C)
- P(A, B | C) = P(A | C) P(B | C)

(recall independence: P(A, B) = P(A)P(B))

This means that, ***if we know C***, we can treat A and B ***as if they were independent***.

A and B might not be independent otherwise!

Example

Consider 3 RVs:

- Temperature
- Humidity
- Season

Temperature and humidity are not independent.

But, they might be, given the season: *the season explains both*, and they become independent of each other.

Bayes' Rule

Special piece of conditioning magic.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

If we have conditional P(B | A) and we **receive new data** for B, we can compute **new distribution for A.** (Don't need joint.)

As evidence comes in, revise belief.

Bayes

sensor model

prior

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

evidence

Bayes' Rule Example

Suppose:

- P(disease) = 0.001
- P(test | disease) = 0.99
- P(test | no disease) = 0.05

What is P(disease | test)?

$$P(d|t) = \frac{P(t|d)P(d)}{P(t)} = \frac{0.99 \times 0.001}{P(t)} \simeq 0.0194$$

$$\begin{aligned} P(t) &= P(t|d)P(d) + P(t|\neg d)P(\neg d) \\ &= 0.99 \times 0.001 + 0.05 \times 0.999 = 0.05094 \end{aligned}$$

Not always symmetric!
Not always intuitive!

Bayes' Rule Example

Suppose:

- P(UFO) = 0.0001
- P(Digits of Pi | UFO) = 0.95
- P(Digits of Pi | not UFO) = 0.001

What is P(UFO | Digits of Pi)?

$P(U|\pi) =$ **~0.087**

$P(\neg U|\pi) =$ **~0.913**

$$P(U|\pi) = \frac{0.95 \times 0.0001}{P(\pi)} \qquad P(\neg U|\pi) = \frac{0.001 \times 0.9999}{P(\pi)}$$

$$\frac{0.001 \times 0.9999}{P(\pi)} + \frac{0.95 \times 0.0001}{P(\pi)} = 1$$

$$\boxed{P(\pi) = 0.0010949}$$

Bayesian Knowledge Bases

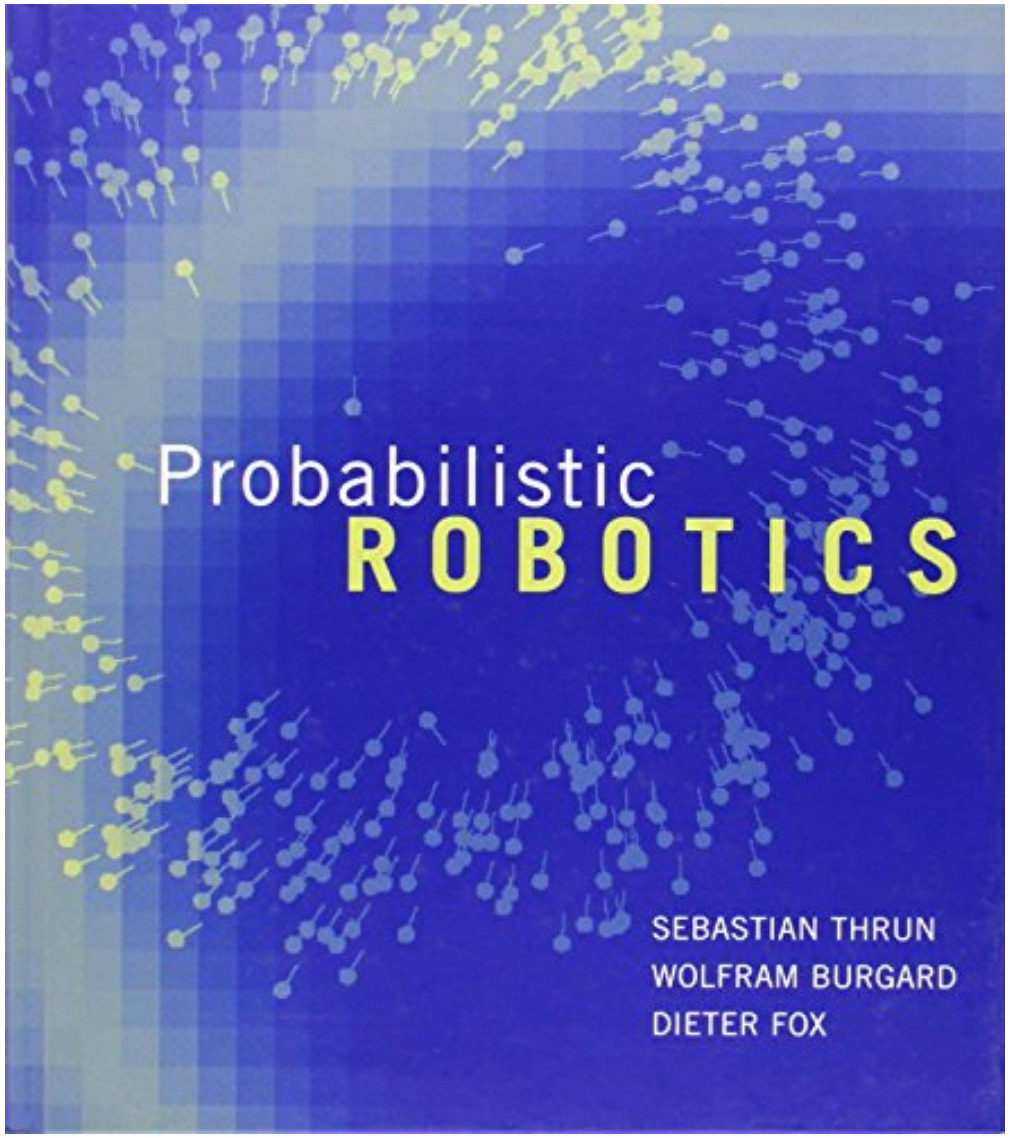
List of conditional and marginal probabilities …

- $P(X_1) = 0.7$
- $P(X_2) = 0.6$.
- $P(X_3 \mid X_2) = 0.57$

Queries:

- $P(X_2 \mid X_1)$?
- $P(X_3)$?

Less onerous than a JPD, but you may, or may not, be able to answer some questions.



(courtesy Thrun and Haehnel)