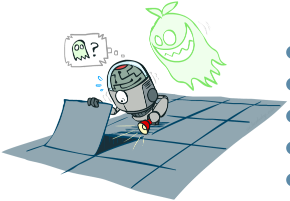
CS188

We've seen how AI methods can solve problems in:

- Search
- Constraint Satisfaction Problems
- Games
- Markov Decision Problems
- Reinforcement Learning

Next up: Part II: Uncertainty and Learning!

Our Status in CS188



We’re done with Part I Search and Planning!

Part II: Probabilistic Reasoning

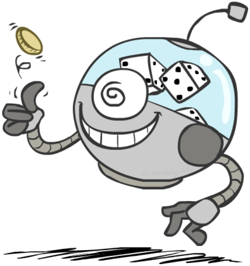
- Diagnosis
- Speech recognition
- Tracking objects
- Robot mapping
- Genetics
- Error correcting codes
- . . . lots more!

Part III: Machine Learning

CS 188: Artificial Intelligence

Probability

Today

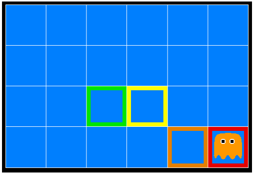


Probability

- Random Variables
- Joint and Marginal Distributions i
- Conditional Distribution
- Product Rule, Chain Rule, Bayes' Rule
- Inference
- Independence

You'll need all this stuff A LOT for the next few weeks, so make sure you get this!

Inference in Ghostbusters



A ghost is in the grid somewhere

Sensor readings tell how close a square is to the ghost

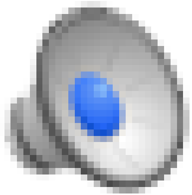
- On the ghost: red
- 1 or 2 away: orange
- 3 or 4 away: yellow
- 5+ away: green

Sensors are noisy, but we know $P(Color|Distance)$

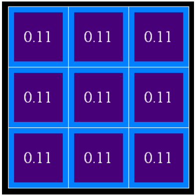
$P(red|3)$	$P(orange|3)$	$P(yellow|3)$	$P(green|3)$
0.05	0.15	0.5	0.3

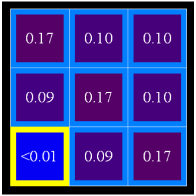
[Demo: Ghostbuster – no probability (L12D1)]

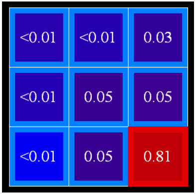
Video of Demo Ghostbuster – No probability



Uncertainty





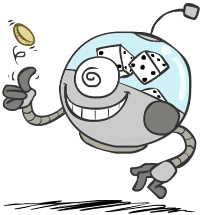


General situation:

- **Observed variables (evidence):** Agent knows certain things about the state of the world (e.g., sensor readings or symptoms)
- **Unobserved variables:** Agent needs to reason about other aspects (e.g. where an object is or what disease is present)
- **Model:** Agent knows something about how the known variables relate to the unknown variables

Probabilistic reasoning gives us a framework for managing our beliefs and knowledge

Random Variables



A random variable is some aspect of the world about which we (may) have uncertainty

- $R =$ Is it raining?
- $T =$ Is it hot or cold?
- $D =$ How long will it take to drive to work?
- $L =$ Where is the ghost?

We denote random variables with capital letters

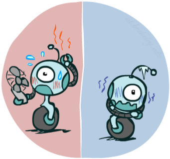
Like CSP, variables (random) have domains

- $R \in \{true, false\}$ (often write as $\{+r, -r\}$)
- $T \in \{hot, cold\}$
- $D \in [0, \infty]$
- L in possible locations, maybe $\{(0,0),(0,1),\ldots\}$

Probability Distributions

Associate a probability with each value

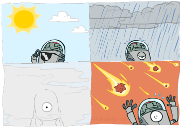
- Temperature:



$P(T)$

T	P
hot	0.5
cold	0.5

- Weather:



$P(W)$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Probability Distributions

Unobserved random variables have distributions:

T	P
hot	0.5
cold	0.5

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Shorthand notation:

$P(hot) = P(T = hot)$
$P(cold) = P(T = cold)$
$P(rain) = P(W = rain)$
$\cdots$

If domains don't overlap.

A probability (lower case value) is a single number:

$P(W = rain) = 0.1.$

A distribution is a TABLE of probabilities of values:

Must have: $\forall x, P(X = x) \geq 0$,
and $\sum_x P(X = x) = 1$.

Joint Distributions

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Set of random variables: $X_1, \ldots, X_n$
Joint Distribtuion:

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$
or $P(x_1, x_2, \ldots, x_n)$.

Same table:

W×T	hot	cold
sun	0.4	0.2
rain	0.1	0.3

- Must obey:

$$P(x_1, x_2, \ldots, x_n) \geq 0$$

$$\sum_{x_1, x_2, \ldots, x_n} P(x_1, x_2, \ldots, x_n) = 1$$

Size of distribution if n variables with domain sizes d? d^n

- For all but the smallest distributions, impractical to write out!

Probabilistic Models

A probabilistic model is a joint distribution over a set of random variables

Probabilistic models:
- (Random) variables with domains
- Assignments are called outcomes
- Joint distributions: frequency of assignments (outcomes)
- Normalized: sum to 1.0
- Ideally: only certain variables directly interact

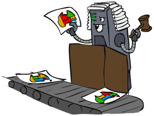
Constraint satisfaction problems:
- Variables with domains
- Constraints: state whether assignments are possible
- Ideally: only certain variables directly interact

Distribution over T,W

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Constraints over T,W

T	W	P
hot	sun	T
hot	rain	F
cold	sun	F
cold	rain	T



Events

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

An event is a set E of outcomes:

$P(E) = \sum_{(x_1,\ldots,x_n)\in E} P(x_1,\ldots,x_n)$.

From a joint distribution, we can calculate the probability of any event

- Probability that it's hot AND sunny?
- Probability that it's hot?
- Probability that it's hot OR sunny?

Typically, the events we care about are partial assignments:
examples: $P(T = hot)$. $P(W = sun)$.

$P(hot) = P(hot, sun) + P(hot, rain) = .5$

Quiz: Events

http://bit.ly/cs188prob

$P(X, Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

$P(+x, +y)$? .2

$P(+x)$? 0.2 + 0.3 = 0.5

$P(-y\ OR\ +x's)$?

P(+x,-y) + P(+x,+y) + P(+x,-y) = 0.6

Marginal Distributions

Marginal distributions are sub-tables which eliminate variables

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Marginal for Temparature.

T	P
hot	0.5
cold	0.5

Marginal for Weather.

W	P
rain	0.4
sun	0.6

Marginalization (summing out): Combine collapsed rows by adding.

Same idea

W $\times$ T	hot	cold	M(W)
sun	0.4	0.2	0.6
rain	0.1	0.3	0.4
M(T)	0.5	0.5	

Quiz: Marginal Distributions

http://bit.ly/cs188prob

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

P(X)

X	P
+x	0.5
-x	0.5

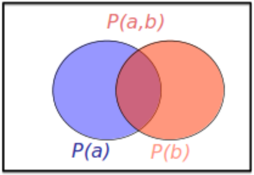
P(Y)

Y	P
+y	0.6
-y	0.4

Conditional Probabilities

A simple relation between joint and conditional probabilities

- In fact, this is taken as the definition of a conditional probability



The probability of event a given event b.

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

Probability of a given b.

Natural? Yes!

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W=s|T=c) = \frac{P(w=s,T=c)}{P(T=c)}.$$

$$P(T=c) = P(W=s,T=c) + P(W=r,T=c)$$
$$= 0.2 + 0.3 = 0.5$$

$$P(W=s|T=c) = \frac{P(w=s,T=c)}{P(T=c)} = \frac{.2}{.5} = 2/5.$$

Quiz: Conditional Probabilities

http://bit.ly/cs188prob

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

$P(+x|+y)$? $\frac{P(+x,+y)}{P(+y)} = \frac{.2}{.6} = 1/3$

$P(-x|+y)$? $= 1 - P(+x|+y) = \frac{2}{3}$.

$P(-y|+x)$? $= \frac{P(-y,+x)}{P(+x)} = \frac{.3}{.5} = 3/5$

Conditional Distributions

Conditional distributions are probability distributions over some variables given fixed values of others

Joint Distribution

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Conditional Distributions

$P(W|T = hot)$

W	P
sun	0.8
cold	0.2

$P(W|T = cold)$

W	P
sun	0.4
cold	0.6

Normalization Trick

Joint Distribution

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$$P(W = s|T = c) = \frac{P(w{=}s, T{=}c)}{P(T{=}c)}$$
$$= \frac{.2}{P(T{=}c)}.$$

$$P(T = c) = P(W = s, T = c) + P(W = r, T = c)$$
$$= 0.2 + 0.3 = 0.5$$

$$P(W = r|T = h) = \frac{P(w{=}r, T{=}h)}{P(T{=}c)}$$
$$= \frac{.3}{P(T{=}c)}.$$

$$P(T = c) = P(W = s, T = c) + P(W = r, T = c)$$

$$= 0.2 + 0.3 = 0.5$$

$P(W|T = cold)$

W	P
sun	0.4
cold	0.6

Why does normalization work?

Answer: Work it out!

Will discuss on Monday,

Have a nice weekend!