AI Programming

CS662-2013S-15

Probability Theory

David Galles

Department of Computer Science
University of San Francisco

15-0: Uncertainty

- In many interesting agent environments, *uncertainty* plays a central role.
- Actions may have nondeterministic effects.
  - Shooting an arrow at a target, retrieving a web page, moving
- Agents may not know the true state of the world.
  - Incomplete sensors, dynamic environment
- Relations between facts may not be deterministic.
  - Sometimes it rains when it's cloudy.
  - Sometimes I play tennis when it's humid.
- Rational agents will need to deal with uncertainty.

15-1: Logic and Uncertainty

- We've already seen how to use logic to deal with uncertainty.
  - $Studies(Bart) \vee WatchesTV(Bart)$
  - $Hungry(Homer) \Rightarrow Eats(Homer, HotDog) \vee Eats(Homer, Pie)$
  - $\exists x Hungry(x)$
- Unfortunately, the logical approach has some drawbacks.

15-2: Weaknesses with logic

- Qualifying all possible outcomes.
  - “If I leave now, I’ll be on time, unless there’s an earthquake, or I run out of gas, or there’s an accident ...”
- We may not know all possible outcomes.
  - “If a patient has a toothache, she may have a cavity, or may have gum disease, or maybe something else we don’t know about.”
- We have no way to talk about the likelihood of events.
  - “It’s possible that I’ll get hit by lightning today.”

15-3: Qualitative vs. Quantitative

- Logic gives us a *qualitative* approach to uncertainty.
  - We can say that one event is more common than another, or that something is a possibility.
  - Useful in cases where we don’t have statistics, or we want to reason more abstractly.
- Probability allows us to reason *quantitatively*
  - We assign concrete values to the chance of an event occurring and derive new concrete values based on observations.

15-4: Uncertainty and Rationality

- Recall our definition of rationality:
  - A rational agent is one that acts to maximize its performance measure.
- How do we define this in an uncertain world?
- We will say that an agent has a *utility* for different outcomes, and that those outcomes have a *probability* of occurring.
- An agent can then consider each of the possible outcomes, their utility, and the probability of that outcome occurring, and choose the action that produces the highest *expected* (or average) utility.
- The theory of combining preferences over outcomes with the probability of an outcome's occurrence is called *decision theory*.
  - Talk a bit more about decision theory after we've covered probability theory.

15-5: Basic Probability

- A probability signifies a *belief* that a proposition is true.
  - P(BartStudied) = 0.01
  - P(Hungry(Homer)) = 0.99
- The proposition itself is true or false - we just don't know which.
- This is different than saying the sentence is partially true.
  - "Bart is short" - this is *sort of* true, since "short" is a vague term.
- An agent's *belief state* is a representation of the probability of the value of each proposition of interest.

Notation

- A *Random Variable* (or just *variable*) is a variable whose value can be described using probabilities
  - Use Upper Case for variables – X, Y, Z, etc.
- Random Variables can have discrete or continuous values (for now, we will assume discrete values)
  - use lower case for values of variables – x, y, x_1, x_2, etc.
- $P(X = x)$ is the probability that variable X has the value x
  - Can also be written as $P(x)$

15-7: Notation

- If variable X can have the values $x_1, x_2, \ldots x_n$, then the expression $P(X)$ stands for a vector which contains $P(X = x_k)$, for all values x_k of X

$$P(X) = [P(X = x_1), P(X = x_2), \ldots P(X = x_n)]$$

- If D is a variable that represents the value of a fair die, then

$$P(D) = [1/6, 1/6, 1/6, 1/6, 1/6, 1/6]$$

15-8: Notation

- Variable W, represents Weather, which can have values sunny, cloudy, rain, or snow.
  - P(W = sunny) = 0.7
  - P(W = cloudy) = 0.2
  - P(W = rain) = 0.08
  - P(W = snow) = 0.02
- $P(W) = [0.7, 0.2, 0.08, 0.02]$

15-9: Notation – AND

$$P(x, y) = P(X = x \wedge Y = y)$$

- Given two fair dice D_1 and D_2:
  $P(D_1 = 3, D_2 = 4) = 1/36$
- $P(X, Y)$ represents the set of $P(x, y)$ for all values x of X and y of Y. Thus, $P(D_1, D_2)$ represents 36 different values.

15-10: Notation – Binary Variables

- If X has two values (false and true), we can represent:
  - $P(X = false)$ as $P(\neg x)$, and
  - $P(X = true)$ as $P(x)$

15-11: Conditional Probability

- $P(x|y)$ = Probability that $X = x$ given that all we know is $Y = y$
- $P(cavity|toothache) = 0.8$
- $P(Cavity|Toothache)$ represents 4 values:

$$\begin{bmatrix} P(\neg cavity|\neg toothache) & P(cavity|\neg toothache) \\ P(\neg cavity|toothache) & P(cavity|toothache) \end{bmatrix}$$

15-12: Conditional Probability

- We can define conditional probabilities in terms of unconditional probabilities.

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

Whenever $P(b) > 0$

- $P(a,b) = P(a|b)P(b) = P(b|a)P(a)$
- $P(A,B) = P(A|B)P(B)$ means $P(a,b) = P(a|b)P(b)$ for all values a, b

15-13: Axioms of Probability

- $0 \leq P(a) \leq 1$
- $P(true) = 1, P(false) = 0$
- $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Everything follows from these three axioms

For instance, prove $P(x) = 1 - P(\neg x)$

15-14: Axioms of Probability

- $0 \leq P(a) \leq 1$
- $P(true) = 1$, $P(false) = 0$
- $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

$$
\begin{aligned}
P(x \vee \neg x) &= P(x) + P(\neg x) - P(x \wedge \neg x) \\
1 &= P(x) + P(\neg x) - 0 \\
1 - P(\neg x) &= P(x) \\
P(x) &= 1 - P(\neg x)
\end{aligned}
$$

15-15: Joint Probability

- Probability for all possible values of all possible variables

cavity	toothache	0.04
cavity	¬toothache	0.06
¬cavity	toothache	0.01
¬cavity	¬toothache	0.89

- From the joint, we can calculate anything

Joint Probability

- Probability for all possible values of all possible variables

cavity	toothache	0.04
cavity	¬toothache	0.06
¬cavity	toothache	0.01
¬cavity	¬toothache	0.89

- From the joint, we can calculate anything
  - $P(\text{cavity})$ = 0.04 + 0.06 = 0.01
  - $P(\text{cavity} \vee \text{toothache})$ = 0.04 + 0.06 + 0.01 = 0.11
  - $P(\text{cavity}|\text{toothache}) = P(c, t)/P(t)$ = 0.04 / (0.04 + 0.01) = 0.80

15-17: Joint Probability

sunny	windy	playTennis	0.1
sunny	windy	¬ playTennis	0.1
sunny	¬ windy	playTennis	0.3
sunny	¬ windy	¬ playTennis	0.05
¬ sunny	windy	playTennis	0.05
¬ sunny	windy	¬ playTennis	0.2
¬ sunny	¬ windy	playTennis	0.1
¬ sunny	¬ windy	¬ playTennis	0.1

- $P(sunny)$?
- $P(playTennis|\neg windy)$?
- $P(playTennis)$?
- $P(playTennis|sunny \wedge \neg windy)$?
    - (also written $P(playTennis|sunny, \neg windy)$)

Joint Probability

- Joint can tell us everything
- Calculate the joint, read off what you want to know
- This will *not* work!

15-19: Joint Probability

- Joint can tell us everything
- Calculate the joint, read off what you want to know
- This will *not* work!
  - x different variables, each of which has v values
  - Size of joint = v^x
  - 50 variables, each has 7 values, $1.8 * 10^{42}$ table entires

15-20: Conditional Probability

- Working with the joint is impractical
- Work with conditional probabilities instead
- Manipulate conditional probabilities based on definition:

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

(when $P(B)$ is always > 0)

15-21: Conditional Probability

- Just as we can define conditional probability given probability of a conjunction (AND), we can define the probability of a conjunction given a conditional probability

$$P(A, B) = P(A|B)P(B)$$

Conditional Probability

Example:

- $P(cloudy) = 0.25$
- $P(rain) = 0.25$
- $P(cloudy \wedge rain) = 0.15$
- $P(cloudy \wedge \neg rain) = 0.1$
- $P(\neg cloudy \wedge rain) = 0.1$
- $P(\neg cloudy \wedge \neg rain) = 0.65$
    - Initially, $P(Rain) = 0.25$. Once we see that it's cloudy, $P(rain|cloudy) = P\frac{(rain \wedge cloudy)}{P(cloudy)} = \frac{0.15}{0.25} = 0.6$

15-23: Independence

- In some cases, we can simplify matters by noticing that one variable has no effect on another.
- For example, what if we add a fourth variable $DayOfWeek$ to our Rain calculation?
- Since the day of the week will not affect the probability of rain, we can assert $P(rain|cloudy, monday) = P(rain|cloudy, tuesday)... = P(rain|cloudy)$
- We say that $DayOfWeek$ and $Rain$ are independent.
- We can then split the larger joint probability distribution into separate subtables.
- Independence will help us divide the domain into separate pieces.

15-24: Bayes' Rule

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

- Also known as Bayes' theorem (or Bayes' law).

Bayes' Rule

$$
\begin{aligned}
P(B|A) &= \frac{P(A \wedge B)}{P(A)} \\
&= \frac{P(A|B)P(B)}{P(A)}
\end{aligned}
$$

Generalize Bayes Rule, with additional evidence E:

$$
\begin{aligned}
P(B|A \wedge E) &= \frac{P(A \wedge B|E)}{P(A|E)} \\
&= \frac{P(A|B \wedge E)P(B|E)}{P(A|E)}
\end{aligned}
$$

15-26: Bayes' theorem example

- Say we know:
  - Meningitis causes a stiff neck in 50% of patients.
  - $P(stiffNeck|meningitis) = 0.5$
  - Prior probability of meningitis is 1/50000.
  - $P(meningitis) = 0.00002$
  - Prior probability of a stiff neck is 1/20
  - $P(stiffNeck) = 0.05$
- A patient comes to use with a stiff neck. What is the probability she has meningitis?
- $P(meningitis|stiffNeck) = \frac{P(stiffNeck|meningitis)P(meningitis)}{P(stiffNeck)} = \frac{0.5 \times 0.00002}{0.05} = 0.0002$

Using Bayes Rule

- Rare disease, strikes one in every 10,000
- Test for the disease that is 95% accurate:
  - $P(t|d) = 0.95$
  - $P(\neg t|\neg d) = 0.95$
- Someone tests positive for the disease, what is the probability that they have it?
  - $P(d|t) = ?$

15-28: Using Bayes Rule

- $P(d) = 0.0001$
- $P(t|d) = 0.95$
- $P(\neg t|\neg d) = 0.95$

$$P(d|t) = P(t|d)P(d)/P(t)$$

15-29: Using Bayes Rule

- $P(d) = 0.0001$
- $P(t|d) = 0.95$ (and hence $P(\neg t|d) = 0.05$)
- $P(\neg t|\neg d) = 0.95$ (and hence $P(t|\neg d) = 0.05$)

$$P(d|t) = P(t|d)P(d)/P(t)$$

15-30: Using Bayes Rule

- $P(d) = 0.0001$
- $P(t|d) = 0.95$
- $P(\neg t|\neg d) = 0.95$

$$\begin{aligned}
P(d|t) &= P(t|d)P(d)/P(t) \\
&= 0.95 * 0.0001/(P(t|d)P(d) + P(t|\neg d)P(\neg d)) \\
&= 0.95 * 0.0001/(0.95 * 0.0001 + 0.05 * 0.9999) \\
&= 0.0019
\end{aligned}$$

Using Bayes Rule

- This is somewhat counter-intuitive
  - Test is 95% accurate
  - Test is positive
  - Only a 0.19% chance of having the disease!
  - Why?

Using Bayes Rule

- Note that for:
  - $P(a|b) = P(b|a)P(a)/P(b)$
- We needed $P(b)$, which was a little bit of a pain to calculate
- We can often get away with *not* calculating it!

Using Bayes Rule

- $P(a|b) = \alpha P(b|a)P(a)$
- α is a normalizing constant
  - Calculate $P(b|a)P(a)$, $P(b|\neg a)P(\neg a)$
  - $\alpha = \frac{1}{P(b|a)P(a)+P(b|\neg a)P(\neg a)}$
  - No magic here: $\alpha = \frac{1}{P(b)}$

Conditional Independence

- Variable A is conditionally independent of variable B, if $P(A|B) = P(A)$
- Notation: $(A \perp\!\!\!\perp B)$
  - D – roll of a fair die $(d_1 \ldots d_6)$
  - C – value of a coin flip (h or t)
  - $P(D|C) = P(D)$
    $P(C|D) = P(C)$
- $(A \perp\!\!\!\perp B) \Leftrightarrow (B \perp\!\!\!\perp A)$
- $P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B)$

15-35: Conditional Independence

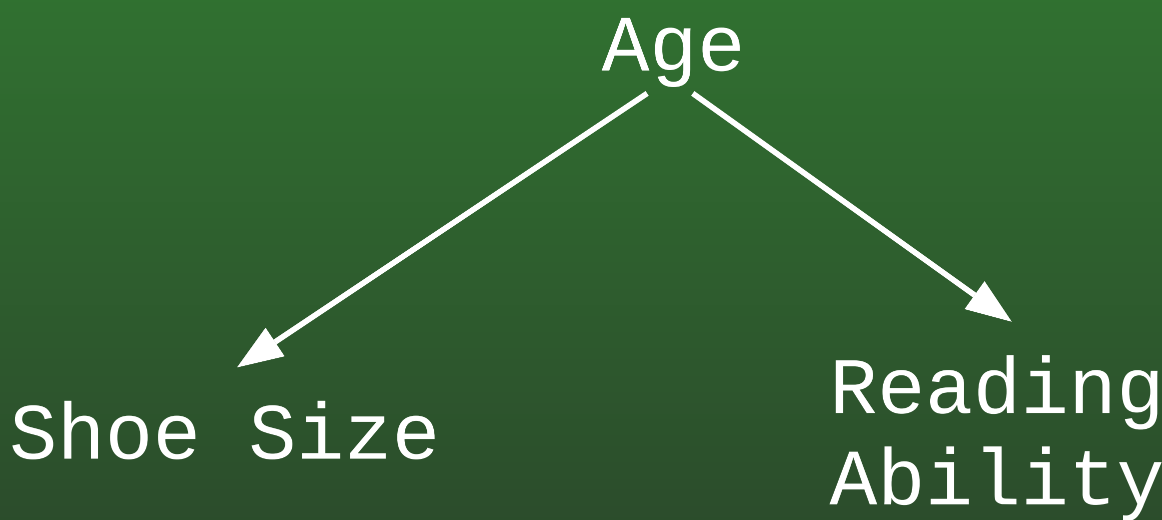
- If A and B are independent, then $P(a, b) = P(a)P(b)$
  - (Also used as a definition of conditional independence – two definitions are equivalent)
- $P(a, b) = P(a|b)P(b) = P(a)P(b)$

Conditional Independence

- At an elementary school, reading scores and shoe sizes are correlated. Why?

Conditional Independence

- At an elementary school, reading scores and shoe sizes are correlated. Why?



    - $P(R|S) \neq P(R)$
    - $P(R|S, A) = P(R|A)$
- Notation: $(R \perp\!\!\!\perp S\,|A)$

15-38: Monte Hall Problem

From the game show “Let’s make a Deal”

- Pick one of three doors. Fabulous prize behind one door, goats behind other 2 doors
- Monty opens one of the doors you did not pick, shows a goat
- Monty then offers you the chance to switch doors, to the other unopened door
- Should you switch?

Monte Hall Problem

Problem Clarification:

- Prize location selected randomly
- Monty always opens a door, allows contestants to switch
- When Monty has a choice about which door to open, he chooses randomly.

Variables

Prize: $P = p_A, p_B, p_C$

Choose: $C = c_A, c_B, c_C$

Monty: $M = m_A, m_B, m_C$

15-40: Monte Hall Problem

Without loss of generality, assume:

- Choose door A
- Monty opens door B

$P(p_A | c_A, m_B) =?$

Monte Hall Problem

Without loss of generality, assume:

- Choose door A
- Monty opens door B

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$

Monte Hall Problem

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$

- $P(m_B|c_A, p_A) = ?$

15-43: Monte Hall Problem

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$

- $P(m_B|c_A, p_A) = 1/2$
- $P(p_A|c_A) =$?

15-44: Monte Hall Problem

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$

- $P(m_B|c_A, p_A) = 1/2$
- $P(p_A|c_A) = 1/3$
- $P(m_B|c_A) = ?$

15-45: Monte Hall Problem

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$

- $P(m_B|c_A, p_A) = 1/2$
- $P(p_A|c_A) = 1/3$
- $P(m_B|c_A) = P(m_b|c_A, p_A)P(p_A) + P(m_b|c_A, p_B)P(p_B) + P(m_b|c_A, p_C)P(p_C)$

Monte Hall Problem

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$

- $P(m_B|c_A, p_A) = 1/2$
- $P(p_A|c_A) = 1/3$
- $P(m_B|c_A) = P(m_b|c_A, p_A)P(p_A)+$
  $P(m_b|c_A, p_B)P(p_B)+$
  $P(m_b|c_A, p_C)P(p_C)$
  - $P(p_A) = P(p_B) = P(p_C) = 1/3$
  - $P(m_b|c_A, p_A) = 1/2$
  - $P(m_b|c_A, p_B) = 0$ Won't open prize door
  - $P(m_b|c_A, p_C) = 1$ Monty has no choice

15-47: Monte Hall Problem

$$P(p_A|c_A, m_B) = P(m_B|c_A, p_A)\frac{P(p_A|c_A)}{P(m_B|c_A)}$$
$$= 1/3$$

- $P(m_B|c_A, p_A) = 1/2$
- $P(p_A|c_A) = 1/3$
- $P(m_B|c_A) = P(m_b|c_A, p_A)P(p_A)+$
  $P(m_b|c_A, p_B)P(p_B)+$
  $P(m_b|c_A, p_C)P(p_C) = 1/2$
  - $P(p_A) = P(p_B) = P(p_C) = 1/3$
  - $P(m_b|c_A, p_A) = 1/2$
  - $P(m_b|c_A, p_B) = 0$ Won't open prize door
  - $P(m_b|c_A, p_C) = 1$ Monty has no choice

Monte Hall Problem

$$P(p_C|c_A, m_B) = P(m_B|c_A, p_C)\frac{P(p_C|c_A)}{P(m_B|c_A)}$$
$$= 2/3$$

- $P(m_B|c_A, p_C) = 1$
- $P(p_C|c_A) = 1/3$
- $P(m_B|c_A) = P(m_b|c_A, p_A)P(p_A) + P(m_b|c_A, p_B)P(p_B) + P(m_b|c_A, p_C)P(p_C) = 1/2$
  - $P(p_A) = P(p_B) = P(p_C) = 1/3$
  - $P(m_b|c_A, p_A) = 1/2$
  - $P(m_b|c_A, p_B) = 0$ Won't open prize door
  - $P(m_b|c_A, p_C) = 1$ Monty has no choice

Rare Disease Redux

- Rare disease, strikes one in every 10,000
- Two tests, one 95% accurate, other 90% accurate:
  - $P(t1|d) = 0.95$, $P(\neg t1|\neg d) = 0.95$
  - $P(t2|d) = 0.90$, $P(\neg t2|\neg d) = 0.90$
- Tests use independent mechanisms to detect disease, and are conditionally independent, given disease state:
  - $P(t1|d, t2) = P(t1|d)$
- Both test are positive, what is the probability of disease?
  - $P(d|t1, t2) = ?$

Rare Disease Redux

$$
\begin{aligned}
P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

15-51: Rare Disease Redux

$$
\begin{aligned}
P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

$$
P(t1|t2) = ?
$$

15-52: Rare Disease Redux

$$\begin{aligned} P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\ &= P(t1|d)\frac{P(d|t2)}{P(t1|t2)} \end{aligned}$$

$$\begin{aligned} P(t1|t2) &= P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2) \\ &= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2) \end{aligned}$$

Rare Disease Redux

$$P(d|t1, t2) = P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)}$$
$$= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}$$

$$P(t1|t2) = P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2)$$
$$= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2)$$

$$P(d|t2) = ?$$

15-54: Rare Disease Redux

$$
\begin{aligned}
P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

$$
\begin{aligned}
P(t1|t2) &= P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2) \\
&= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2)
\end{aligned}
$$

$$
P(d|t2) = P(t2|d)\frac{P(d)}{P(t2)}
$$

15-55: Rare Disease Redux

$$
\begin{aligned}
P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

$$
\begin{aligned}
P(t1|t2) &= P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2) \\
&= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2)
\end{aligned}
$$

$$
P(d|t2) = P(t2|d)\frac{P(d)}{P(t2)}
$$

$$
P(t2) = \text{?}
$$

15-56: Rare Disease Redux

$$
\begin{aligned}
P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

$$
\begin{aligned}
P(t1|t2) &= P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2) \\
&= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2)
\end{aligned}
$$

$$
P(d|t2) = P(t2|d)\frac{P(d)}{P(t2)}
$$

$$
P(t2) = P(t2|d)P(d) + P(t2|\neg d)P(\neg d)
$$

Rare Disease Redux

$$
\begin{aligned}
P(d|t1,t2) &= P(t1|d,t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

$$
\begin{aligned}
P(t1|t2) &= P(t1|t2,d)P(d|t2) + P(t1|t2,\neg d)P(\neg d|t2) \\
&= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2)
\end{aligned}
$$

$$
P(d|t2) = P(t2|d)\frac{P(d)}{P(t2)}
$$

$$
\begin{aligned}
P(t2) &= P(t2|d)P(d) + P(t2|\neg d)P(\neg d) \\
&= .10008
\end{aligned}
$$

15-58: Rare Disease Redux

$$
\begin{aligned}
P(d|t1,t2) &= P(t1|d,t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)}
\end{aligned}
$$

$$
\begin{aligned}
P(t1|t2) &= P(t1|t2,d)P(d|t2) + P(t1|t2,\neg d)P(\neg d|t2) \\
&= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2)
\end{aligned}
$$

$$
\begin{aligned}
P(d|t2) &= P(t2|d)\frac{P(d)}{P(t2)} \\
&= 0.0009
\end{aligned}
$$

$$
\begin{aligned}
P(t2) &= P(t2|d)P(d) + P(t2|\neg d)P(\neg d) \\
&= .10008
\end{aligned}
$$

15-59: Rare Disease Redux

$$\begin{aligned} P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\ &= P(t1|d)\frac{P(d|t2)}{P(t1|t2)} \end{aligned}$$

$$\begin{aligned} P(t1|t2) &= P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2) \\ &= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2) \\ &= 0.05081 \end{aligned}$$

$$\begin{aligned} P(d|t2) &= P(t2|d)\frac{P(d)}{P(t2)} \\ &= 0.0009 \end{aligned}$$

$$\begin{aligned} P(t2) &= P(t2|d)P(d) + P(t2|\neg d)P(\neg d) \\ &= .10008 \end{aligned}$$

15-60: Rare Disease Redux

$$
\begin{aligned}
P(d|t1, t2) &= P(t1|d, t2)\frac{P(d|t2)}{P(t1|t2)} \\
&= P(t1|d)\frac{P(d|t2)}{P(t1|t2)} \\
&= 0.0168
\end{aligned}
$$

$$
\begin{aligned}
P(t1|t2) &= P(t1|t2, d)P(d|t2) + P(t1|t2, \neg d)P(\neg d|t2) \\
&= P(t1|d)P(d|t2) + P(t1|\neg d)P(\neg d|t2) \\
&= 0.05081
\end{aligned}
$$

$$
\begin{aligned}
P(d|t2) &= P(t2|d)\frac{P(d)}{P(t2)} \\
&= 0.0009
\end{aligned}
$$

$$
\begin{aligned}
P(t2) &= P(t2|d)P(d) + P(t2|\neg d)P(\neg d) \\
&= .10008
\end{aligned}
$$

Probabilistic Reasoning

- Given:
    - Set of conditional probabilities ($P(t1|d)$, etc)
    - Set of prior probabilities ($P(d)$)
    - Conditional independence information ($P(t1|d, t2) = P(t1|d)$)
- We can calculate any quantity that we like

15-62: Probabilistic Reasoning

- Given:
  - Set of conditional probabilities ($P(t1|d)$, etc)
  - Set of prior probabilities ($P(d)$)
  - Conditional independence information ($P(t1|d, t2) = P(t1|d)$)
- We can calculate any quantity that we like
- Problems:
  - Hard to know exactly what data we need
  - Even given sufficient data, calculations can be complex – especially dealing with conditional independence

15-63: Probabilistic Reasoning

- Next time:
  - Make some simplifiying assumptions that are not theoretically sound, but give good results in practice
    - Naive Bayes
    - Works very well for spam filtering
  - Use conditional independence to create a probabilistic alternative to rule-bases systems
    - Bayesian Networks