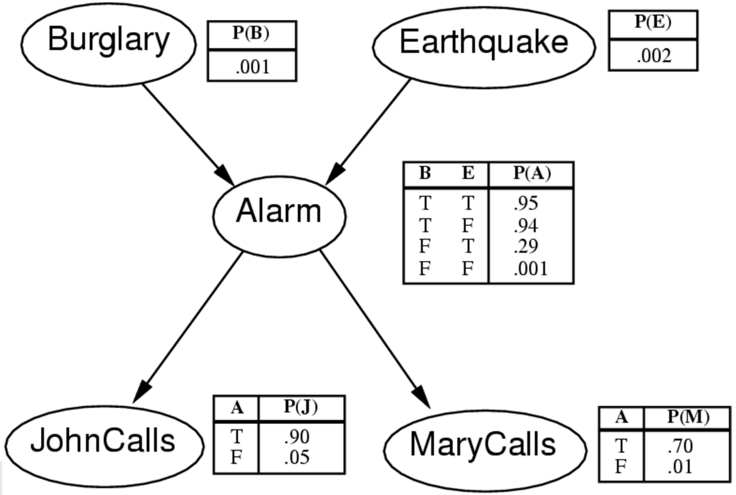
CS 482/682 – AI: Making Simple Decisions

Fall 2021 – Chapter 16

A Basic Belief Network



A conditional probability table gives the likelihood of a particular combination of values

You have a new home alarm that responds

- Accurately to burglaries
- Occasionally responds to earthquakes

When the alarm rings, your neighbors call you at work

- John always calls, but sometimes confuses the telephone for the alarm
- Mary sometimes misses the alarm, but only calls when the alarm actually rings

Making Decisions

- Probability theory describes what an agent should believe on the basis of evidence
- Utility theory describes what an agent wants
- Decision theory puts the two together to describe what an agent should do

Utility

- Utility is a measure of an agent's preferences
- Utility functions transform a world state into a utility (a real number)
- Examples (might not result in optimal behavior):
  - Evel Knievel agent (stunt performer and entertainer)
    - U(S) = 1 if you survive the jump, 0 otherwise
  - Checkers :
    - U(S) = NumberYourPieces – NumberOpponentPieces
  - Shopping Agent
    - U(S) = +1 for each item on the list
- In practical problems, the utility function can be much more complicated
  - Damage to self, energy consumed, objectives achieved, time...

What is the Utility of an Action?

- In a deterministic environment, we can simply compute the future world state and then apply the utility function to that

- In a non-deterministic environment, what do we do?
  - Compute the expected utility (the expected value of the utility) of taking a particular action

Expected Utility

- Consider a non-deterministic action *A* that has possible outcome states: $Result_i\ (A)$

- Compute the probability of ending up in each of these outcome states given that you perform action *A*: $P(Result_i\ (A) \mid Do\ (A))$

- Weight the utility of each possible state by its probability and find the expected utility:

$$EU(A) = \Sigma_i P(\ Result_i(A) \mid Do(A)) \text{ x } U(Result_i(A))$$

What action should you choose?

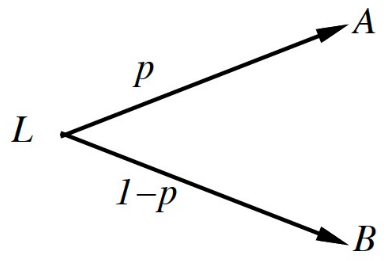
- Principle of Maximum Expected Utility:
  - Choose the action that maximizes the expected utility

- Is this always the best policy?
  - Expected value is an average
  - Why not minimize your potential loss?
  - Why not minimize the deviation?

The Basis of Utility Theory

Possible scenarios are called lotteries to emphasize the idea that the different attainable outcomes are like different prizes, and that the outcome is determined by chance

L=[p, A ; 1-p, B] indicates that there is a *p* chance of obtaining result *A* and a *1-p* chance of obtaining result *B*



We use notations to compare utilities:

$A \succ B$	A preferred to B
$A \sim B$	indifference between A and B
$A \succsim B$	B not preferred to A

Axioms of Utility Theory

- Ordering

  $(A \succ B) \vee (B \succ A) \vee (A \sim B)$

- Transitivity

  $(A \succ B) \wedge (B \succ C) \Rightarrow (A \succ C)$

- Continuity
  - If some state B is between A and C in preference, then there is some probability *p* for which the rational agent will be indifferent between getting B for sure and the lottery that yields A with probability *p* and C with probability *1 – p*

  $A \succ B \succ C \Rightarrow \exists p\ [p, A; 1\text{-}p, C] \sim B$

Axioms of Utility Theory

- Decomposition
    - Compound lotteries can be reduced to simpler lotteries using the laws of probability
- Substitutability
    - if an agent is indifferent between two lotteries, A and B, then the agent is indifferent between two more complex lotteries in which B is substituted for A

    $$A \sim B \Rightarrow [p, A; 1\text{-}p, C] \sim [p, B; 1\text{-}p, C]$$
- Monotonicity
    - If an agent prefers A to B, then the agent will prefer the lottery that has a higher chance of A

    $$A \succ B \Rightarrow (p > q \Leftrightarrow [p, A; 1\text{-}p, B] \succ [q, A; 1\text{-}q, B])$$

Utility Theory

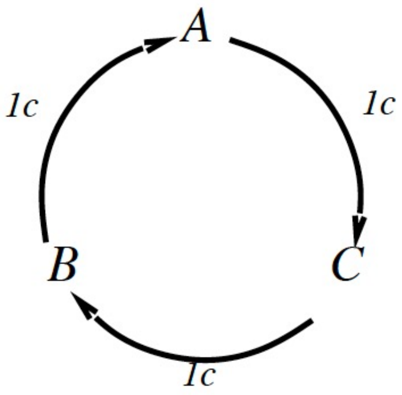
Violating the constraints leads to self-evident irrationality

For example: an agent with intransitive preferences can be induced to give away all its money

If $B \succ C$, then an agent who has C would pay (say) 1 cent to get B

If $A \succ B$, then an agent who has B would pay (say) 1 cent to get A

If $C \succ A$, then an agent who has A would pay (say) 1 cent to get C



Utility Theory

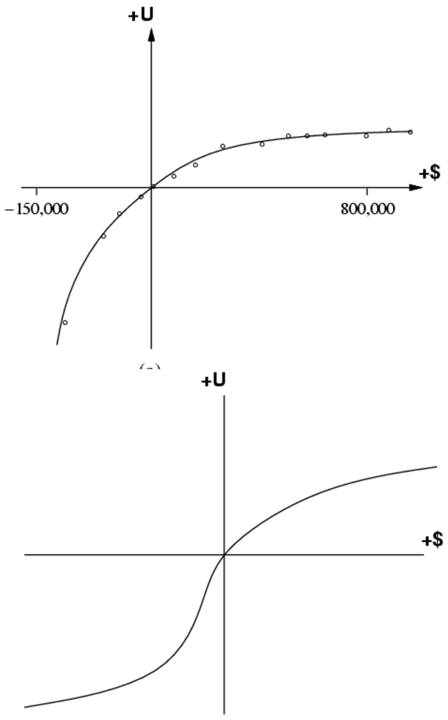
Theorem (Ramsey, 1931; von Neumann and Morgenstern, 1944):
Given preferences satisfying the constraints
there exists a real-valued function U such that

$$U(A) \geq U(B) \quad \Leftrightarrow \quad A \succsim B$$
$$U([p_1, S_1; \ \ldots \ ; \ p_n, S_n]) = \Sigma_i \ p_i U(S_i)$$

MEU principle:

Choose the action that maximizes expected utility

The Utility of Money



(b)

- Utility represents preferences... it is not necessarily logical
- Choices between
  - $5
  - a coin flip for either $10 or nothing
    - Expected value of each is $5
  - $1,000,000
  - A coin flip for either $3,000,000 or nothing

Characterizing the Utility of Money

Bernoulli's St. Petersburg paradox

- A fair coin is tossed until a head comes up.
- If the first head appears on the n^{th} toss, then you win 2^n dollars
- How much should you be willing to pay to enter this game?
- Decision theorists advise us to apply the principle of maximizing expected value.
    - According to this principle, the value of an uncertain prospect is the sum total obtained by multiplying the value of each possible outcome with its probability and then adding up all the terms

Characterizing the Utility of Money

Bernoulli's St. Petersburg paradox

- The expected monetary value of the St. Petersburg game is:
  - $\frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 4 + \frac{1}{8} \cdot 8 + \cdots = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot 2^n = \infty$
  - Some would say that the sum approaches infinity

- The "paradox" consists in the fact that our best theory of rational choice seems to entail that it would be rational to pay any finite fee for a single opportunity to play the St. Petersburg game, even though it is almost certain that the player will win a very modest amount.

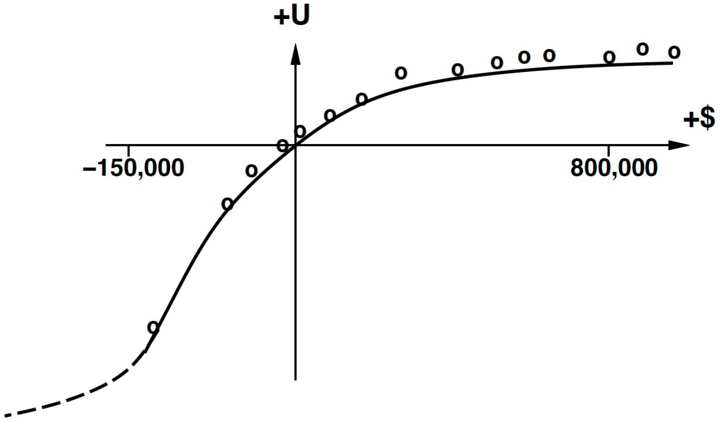
- The probability is $\frac{1}{2}$ that the player wins no more than \$2, and $\frac{3}{4}$ that they win no more than \$4

Characterizing the Utility of Money

Given a lottery L with expected monetary value $EMV(L)$, usually $U(L) < U(EMV(L))$, i.e., people are risk-averse

Utility curve: for what probability p am I indifferent between a fixed prize x and a lottery $[p, \$M;\ (1-p), \$0]$ for large M?

Typical empirical data, extrapolated with risk-prone behavior:



Risk Assessment

Utility functions are on an arbitrary scale

Two agents with utility functions U(S) and kU(s)+c for positive constants will behave in the same way

Normalize utilities on a scale of 0 to 1

Apply utility theory to many domains, including human life

- QALY: quality-adjusted life year (year in good health)
- Micromort: a one in 1,000,000 chance of death
  - A micromort is worth about $20 in 1980 dollars
  - $59.41 in 2017 dollars
  - US gov't: micromort is valued at $2 - $6, depending on the agency

More about micromort

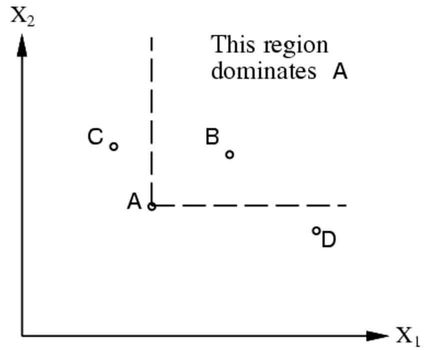
- Micromort = micro- + mortality
- A microprobability is a one-in-a million chance of some event; thus, a micromort is the microprobability of death
- The concept was introduced by Ronald A. Howard, who pioneered the modern practice of decision analysis
- An application of micromorts is measuring the value that humans place on risk. For example, a person can consider the amount of money they would be willing to pay to avoid a one-in-a-million chance of death (or conversely, the amount of money they would receive to accept a one-in-a-million chance of death).
- People are less inclined to spend money after a certain point to increase their safety.

More about micromort

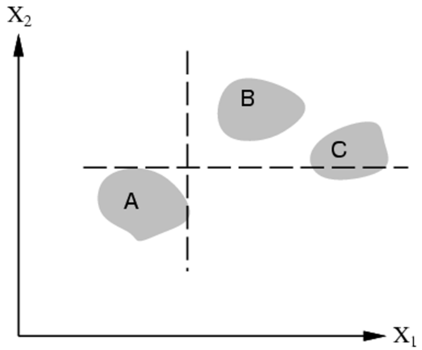
Death from	Context	Time period	N deaths	N population	Micromorts per unit of exposure	
All causes	England and Wales	2012	499,331	56,567,000	24 per day 8,800 per year	
All causes	Canada	2011	242,074	33,476,688	20 per day 7,200 per year	
All causes	US	2010	2,468,435	308,500,000	22 per day 8,000 per year	
Non-natural cause	England and Wales	2012	17,462	56,567,000	0.8 per day 300 per year	
Non-natural cause	US	2010	180,000	308,500,000	1.6 per day 580 per year	
Non-natural cause (excluding suicide)	England and Wales	2012	12,955	56,567,000	0.6 per day 230 per year	
Non-natural cause (excluding suicide)	US	2010	142,000	308,500,000	1.3 per day 460 per year	
All causes – first day of life	England and Wales	2007			430 per first day of life	
All causes – first year of life	US	2013			16.7 per day 6100 per year	B
Murder/homicide	England and Wales	2012/13	551	56,567,000	10 per year	
Homicide	Canada	2011	527	33,476,688	15 per year	
Murder and non-negligent manslaughter	US	2012	14,173	292,000,000	48 per year	

Multi-attribute Utility Functions: Finding the better Utility

Deterministic

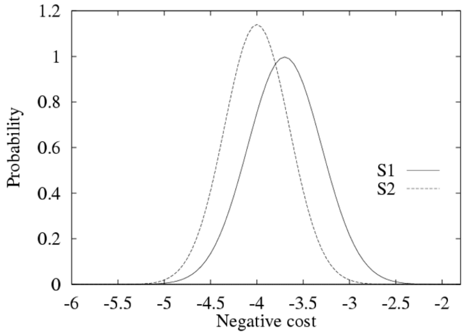


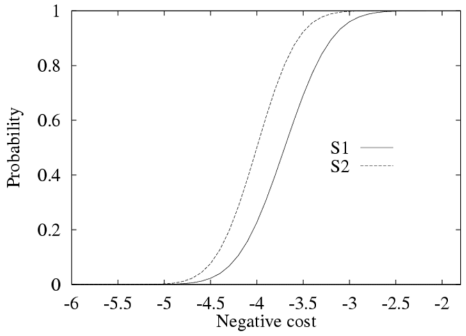
Non-Deterministic



- Assume that two attributes that contribute to the utility function
  - Deterministic environment: choice B dominates A, but A is not dominated by C or D
  - Non-Deterministic environment: choice B dominates A, but A is not strictly dominated by C
    - (there is a chance that a value from C would be worse than A)

Stochastic Dominance



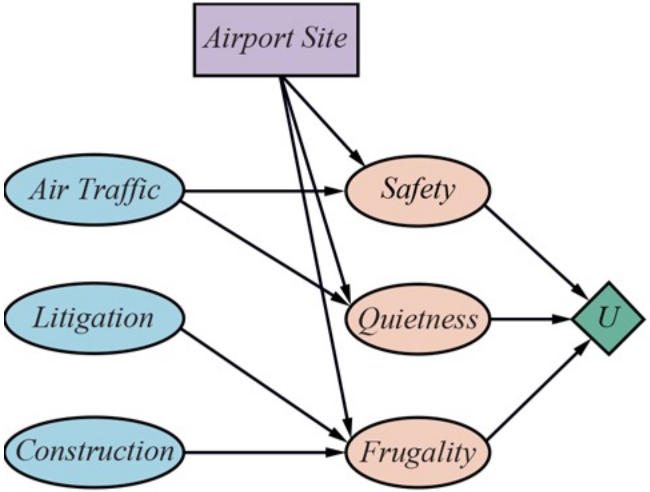


If the cumulative distribution of S1 is always greater than the cumulative distribution of S2, then S1 stochastically dominates S2

$$\forall x \int_{-\infty}^{x} p_1(x')\,dx' \geq \int_{-\infty}^{x} p_2(x')\,dx'$$

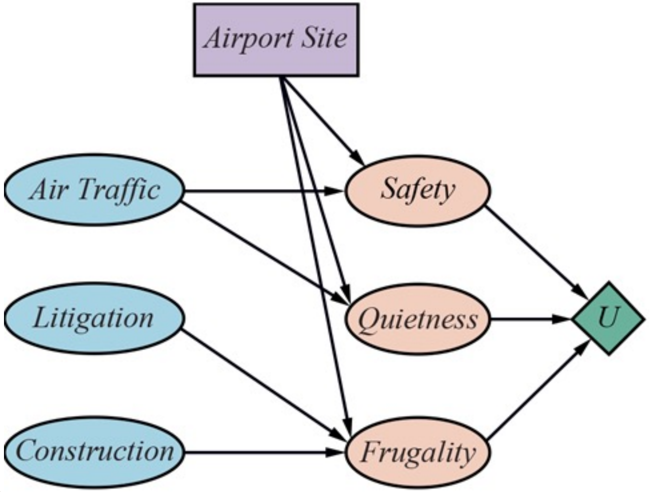
- Example: if cost is dependent on distance, then the closer site will always win even if the cost is non-deterministic

Decision Networks



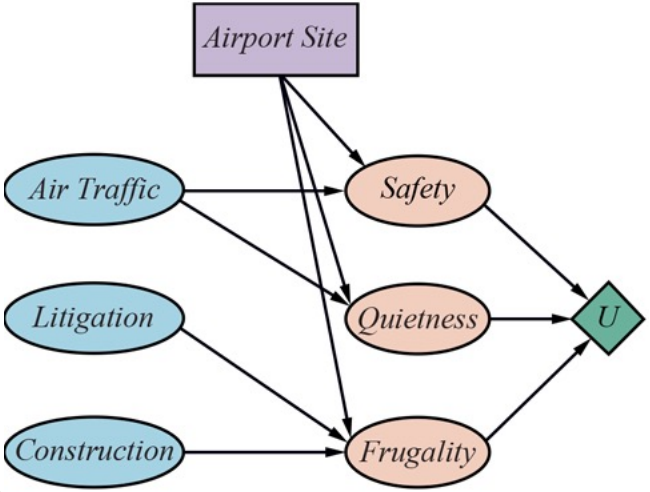
- Using utility theory, we can construct a general mechanism for making decisions
- Decision networks (also called influence diagrams) combine belief networks with additional node types for actions and utilities

Decision Networks



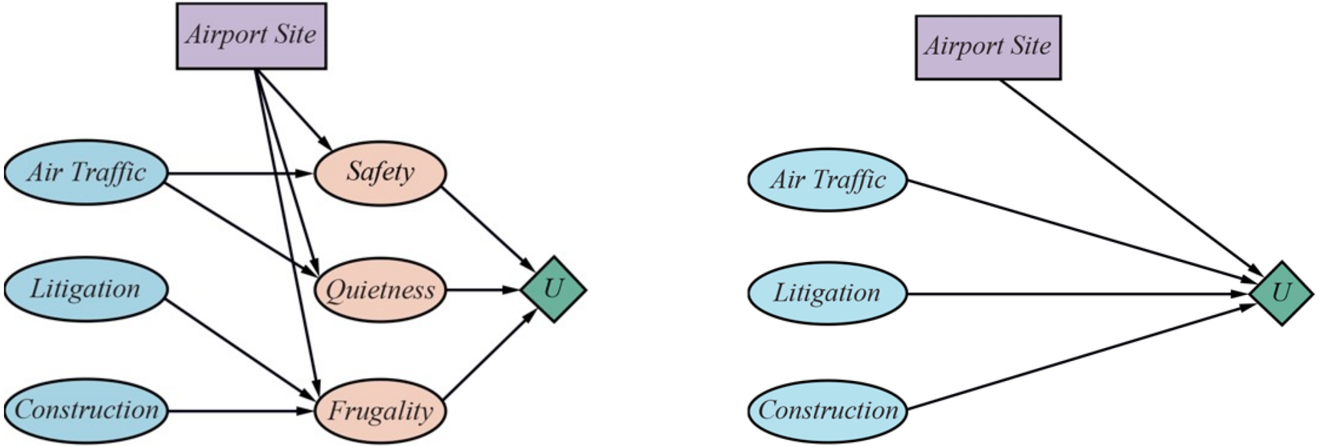
- Chance nodes: (ovals) represent random variables, just as in belief nets
  - Has a conditional probability table (CPT) that is indexed by the state of the parent nodes
  - Parent nodes can include decision nodes as well as chance nodes
- Decision nodes: (rectangles) represent a choice of actions
  - AirportSite action can take on different value for each site under consideration
  - The choice influences the safety, quietness, and frugality of the solution

Decision Networks



- Utility nodes: (diamonds) has as parents all those variables describing the outcome state that directly affect utility
  - Has an action-utility table that calculates the utility function based on the parents
  - Can be parameterized additive or linear function of the attribute values

Simplifying Decision Networks

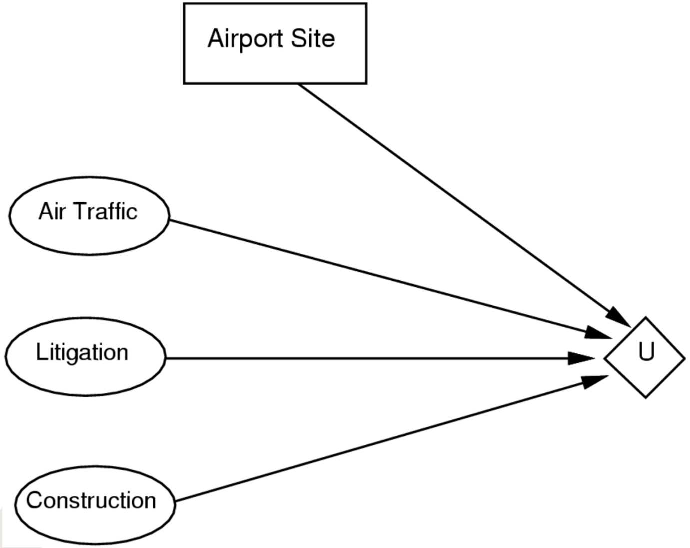


Notice that *Deaths, Noise, and Cost* are removed.

We can often remove these chance nodes that are associated with output states in order to simplify the problem

Action-utility function (Q-function in reinforcement learning)

Evaluating Decision Networks

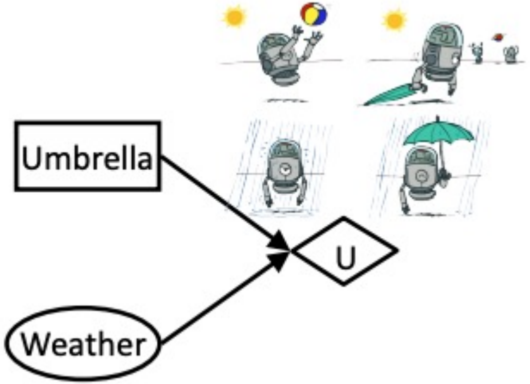


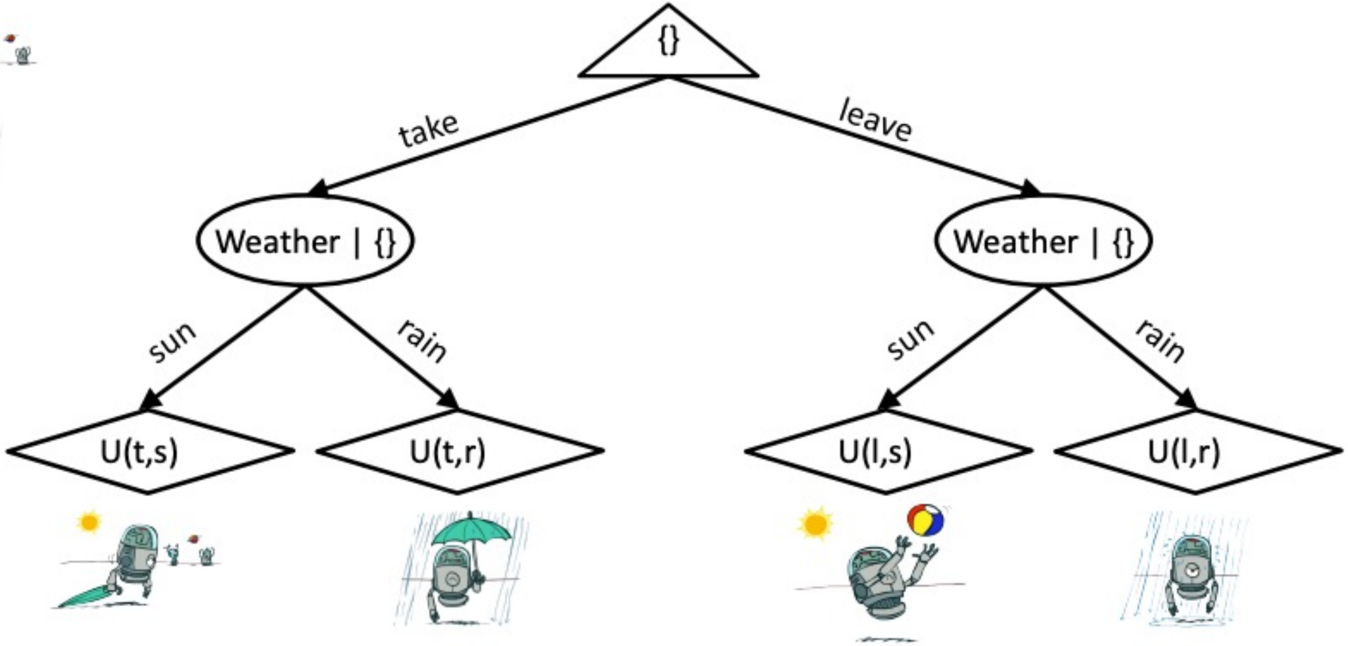
1. Set the evidence variables for the current state
2. For each possible value of the decision node:
   a. Set the decision node to that variable
   b. Calculate the posterior probabilities for the parent nodes of the utility node using probabilistic inference
   c. Calculate the resulting utility for that action
3. Return the action with the largest utility
4. We will see more in Chapter 17!

Example

1. Set the evidence variables for the current state
2. For each possible value of the decision node:
    a. Set the decision node to that variable
    b. Calculate the posterior probabilities for the parent nodes of the utility node using probabilistic inference
    c. Calculate the resulting utility for that action
3. Return the action with the largest utility

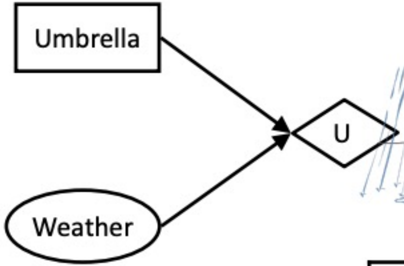
Example

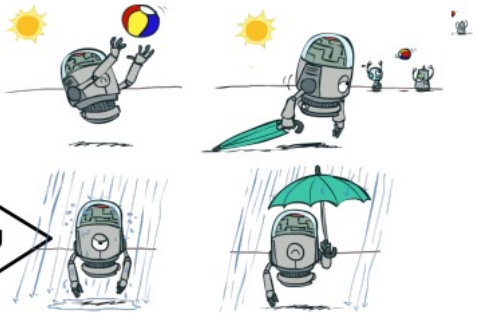


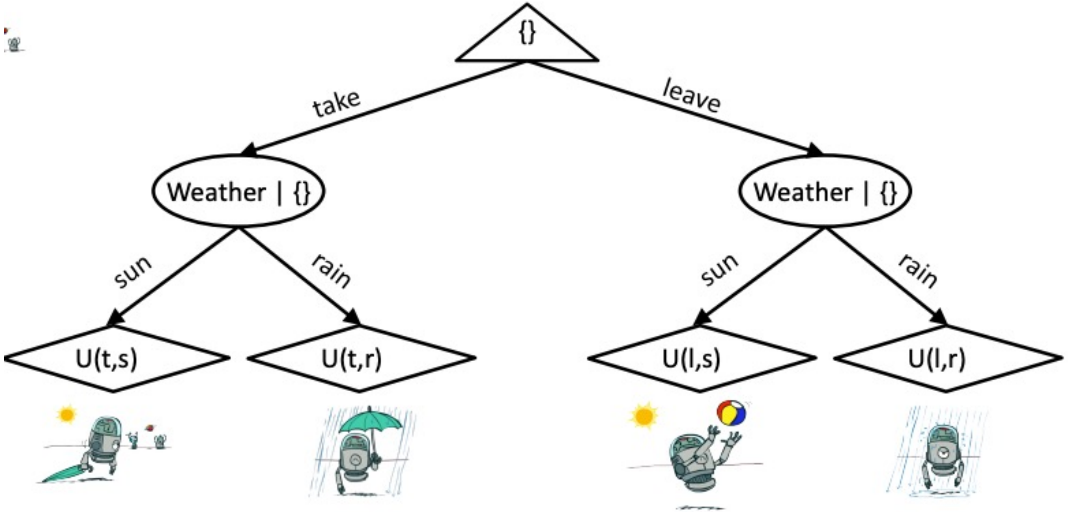


- Almost exactly like expectimax / MDPs
- What’s changed?

Example





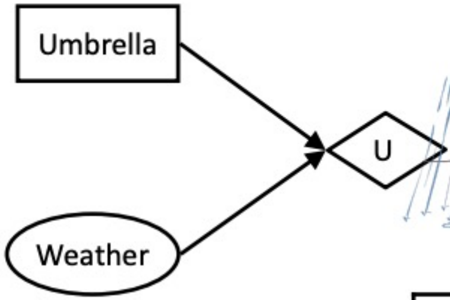


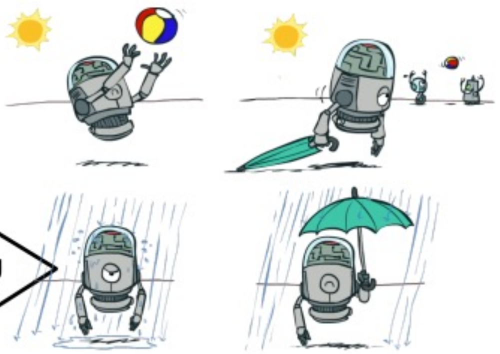
W	P(W)
sun	0.7
rain	0.3

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

- Almost exactly like expectimax / MDPs
- What's changed?

Example





W	P(W)
sun	0.7
rain	0.3

A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

Umbrella = leave

$$\mathrm{EU}(\text{leave}) = \sum_w P(w)U(\text{leave}, w)$$

$$= 0.7 \cdot 100 + 0.3 \cdot 0 = 70$$

Umbrella = take

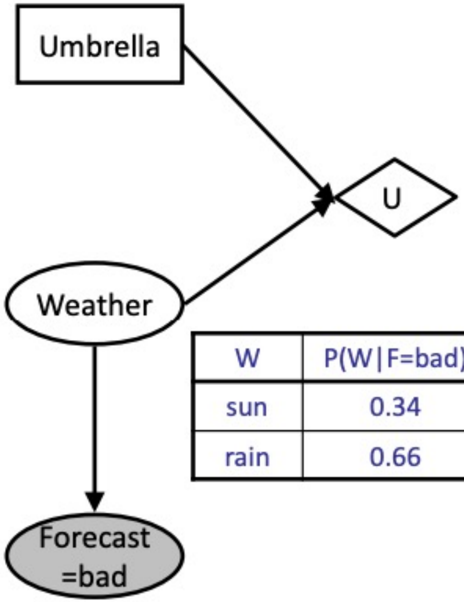
$$\mathrm{EU}(\text{take}) = \sum_w P(w)U(\text{take}, w)$$

$$= 0.7 \cdot 20 + 0.3 \cdot 70 = 35$$

Optimal decision = leave

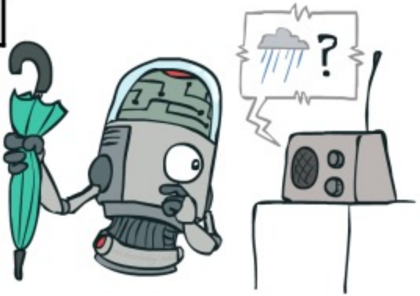
$$\mathrm{MEU}(\phi) = \max_a \mathrm{EU}(a) = 70$$

Example



A	W	U(A,W)
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70

W	P(W|F=bad)
sun	0.34
rain	0.66



Umbrella = leave

$$\mathrm{EU}(\mathrm{leave}|\mathrm{bad}) = \sum_w P(w|\mathrm{bad})U(\mathrm{leave}, w)$$

$$= 0.34 \cdot 100 + 0.66 \cdot 0 = 34$$

Umbrella = take

$$\mathrm{EU}(\mathrm{take}|\mathrm{bad}) = \sum_w P(w|\mathrm{bad})U(\mathrm{take}, w)$$

$$= 0.34 \cdot 20 + 0.66 \cdot 70 = 53$$

Optimal decision = take

$$\mathrm{MEU}(F = \mathrm{bad}) = \max_a \mathrm{EU}(a|\mathrm{bad}) = 53$$

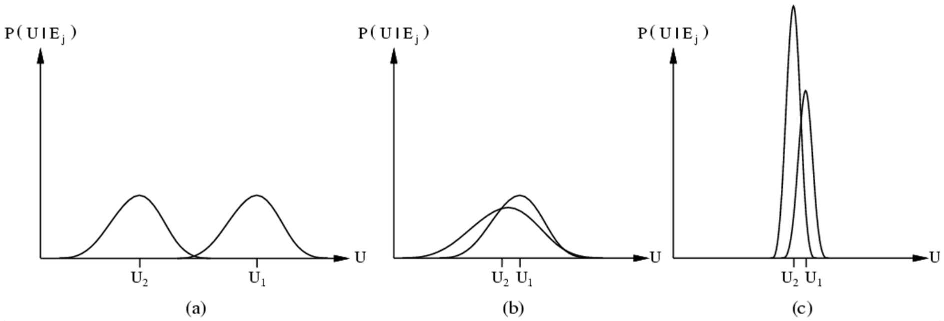
The Value of Information

We assumed that all relevant information was available to the agent when making a decision

One of the most important parts of decision making is knowing what questions to ask

Information value theory enables an agent to choose what information to acquire

The Value of Information



The value of information derives from the fact that with the information, one's course of action may change to become more appropriate to the actual situation

Choice between two roads, which may be blocked by avalanche... do we pay for satellite photos of the two areas?

a. Smooth road around the mountain or a dangerous mountain pass
b. Slightly different lengths, but we have an injured passenger
c. Two roads in summer

Quantifying the Value of Information

In general, the value of a given piece of information is defined to be the difference in expected value between best actions before and after information is obtained

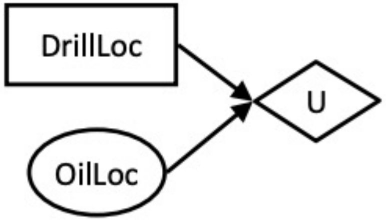
We use the phrase “value of perfect information” (VPI) to describe the value of information that will guarantee we move from the current state to the next best state

- VPI is non-negative
- VPI is not additive

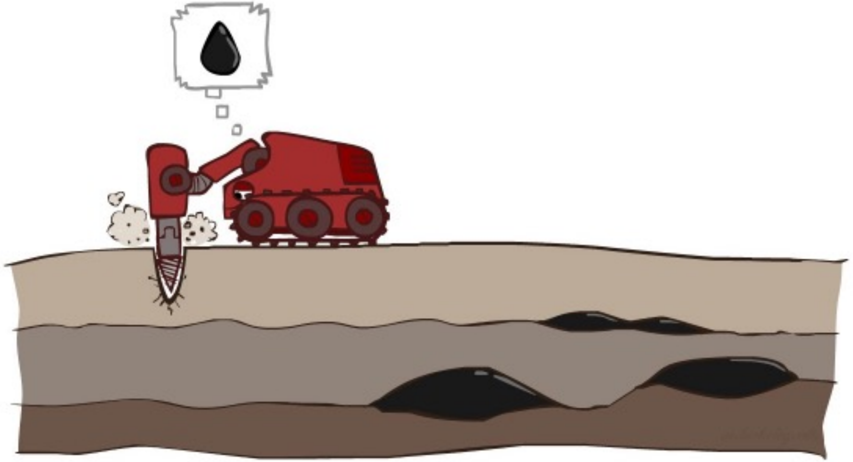
Value of perfect information (VPI)

- Idea: compute value of acquiring evidence
  - Can be done directly from decision network
- Example: buying oil drilling rights
  - Two blocks A and B, exactly one has oil, worth k
  - You can drill in one location
  - Prior probabilities 0.5 each, & mutually exclusive
  - Drilling in either A or B has EU = k/2, MEU = k/2
- Question: what's the value of information of O?
  - Value of knowing which of A or B has oil
  - Value is expected gain in MEU from new info
  - Survey may say "oil in a" or "oil in b", prob 0.5 each
  - If we know OilLoc, MEU is k (either way)
  - Gain in MEU from knowing OilLoc?
  - VPI(OilLoc) = k/2
  - Fair price of information: k/2

O	P
a	1/2
b	1/2



D	O	U
a	a	k
a	b	0
b	a	0
b	b	k



Value of perfect information (VPI)

- Assume we have evidence E=e. Value if we act now:

$$\text{MEU}(e) = \max_a \sum_s P(s|e)\, U(s,a)$$

- Assume we see that E' = e'. Value if we act then:

$$\text{MEU}(e,e') = \max_a \sum_s P(s|e,e')\, U(s,a)$$

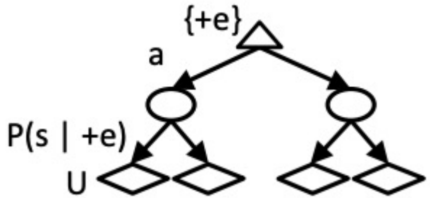
- BUT E' is a random variable whose value is unknown, so we don't know what e' will be

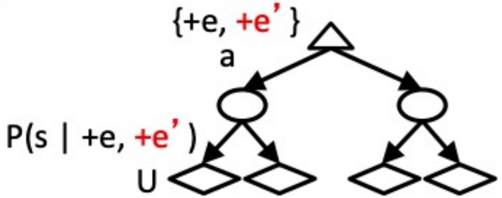
- Expected value if E' is revealed and then we act:

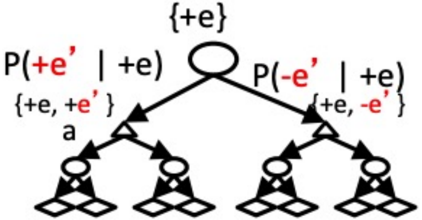
$$\text{MEU}(e,E') = \sum_{e'} P(e'|e)\text{MEU}(e,e')$$

- Value of information: how much MEU goes up by revealing E' first then acting, over acting now:

$$\text{VPI}(E'|e) = \text{MEU}(e,E') - \text{MEU}(e)$$







Value of perfect information (VPI)

MEU with no evidence

$$\text{MEU}(\emptyset) = \max_a \text{EU}(a) = 70$$

MEU if forecast is bad

$$\text{MEU}(F = \text{bad}) = \max_a \text{EU}(a|\text{bad}) = 53$$

MEU if forecast is good

$$\text{MEU}(F = \text{good}) = \max_a \text{EU}(a|\text{good}) = 95$$

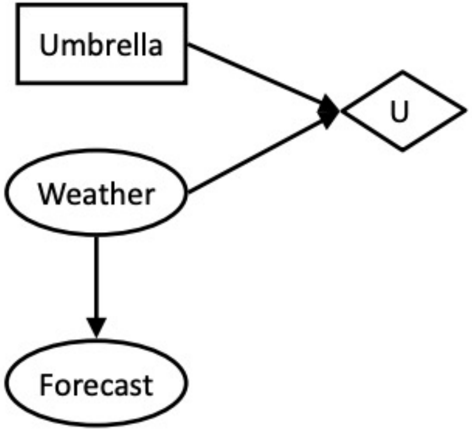
Forecast distribution

F	P(F)
good	0.59
bad	0.41

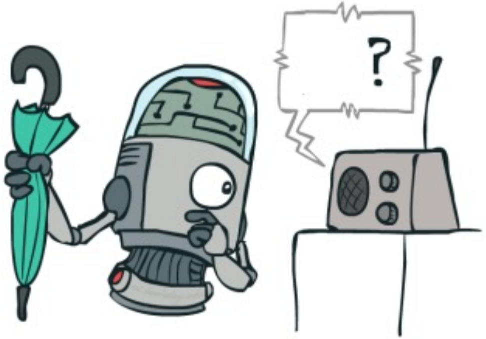
$$0.59 \cdot (95) + 0.41 \cdot (53) - 70$$

$$77.8 - 70 = 7.8$$

$$\text{VPI}(E'|e) = \left(\sum_{e'} P(e'|e)\text{MEU}(e, e') \right) - \text{MEU}(e)$$



A	W	U
leave	sun	100
leave	rain	0
take	sun	20
take	rain	70



Value of perfect information (VPI)

- Nonnegative

$$\forall E', e : \text{VPI}(E'|e) \geq 0$$

- Nonadditive

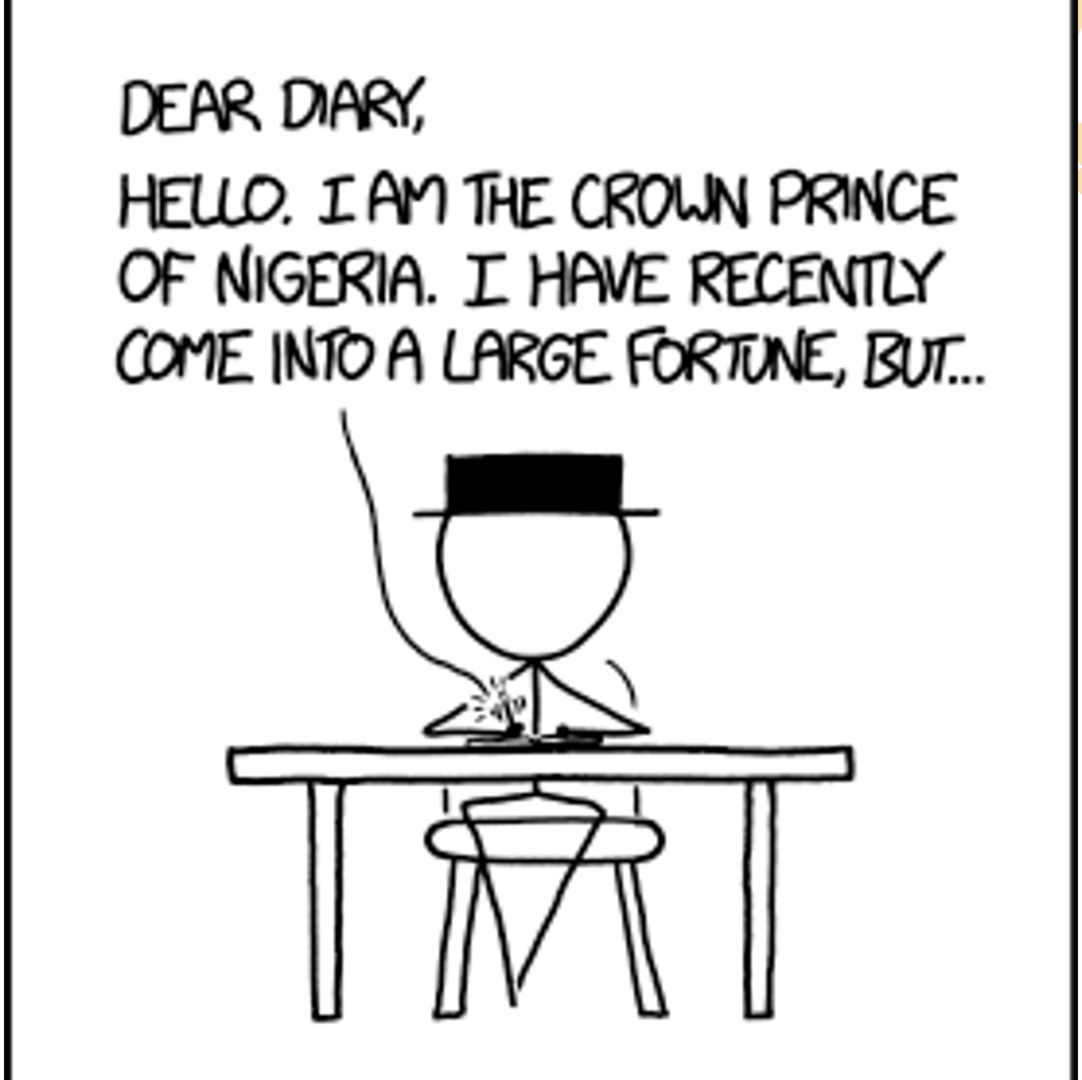
  (think of observing E_j twice)

$$\text{VPI}(E_j, E_k|e) \neq \text{VPI}(E_j|e) + \text{VPI}(E_k|e)$$

- Order-independent

$$\text{VPI}(E_j, E_k|e) = \text{VPI}(E_j|e) + \text{VPI}(E_k|e, E_j)$$
$$= \text{VPI}(E_k|e) + \text{VPI}(E_j|e, E_k)$$

SPAM Filtering


DEAR DIARY,
HELLO. I AM THE CROWN PRINCE OF NIGERIA. I HAVE RECENTLY COME INTO A LARGE FORTUNE, BUT...

What does spam look like?


SPAM
ith CHEESE
BLACK PEPPER
TERIYAKI
Classic
Glorious SPAM!
NET WT 12 OZ (340g)
Serving Suggestion

What does spam look like?

Dear Sir:

I have been requested by the Nigerian National Petroleum Company to contact you for assistance in resolving a matter. The Nigerian National Petroleum Company has recently concluded a large number of contracts for oil exploration in the sub-Sahara region. The contracts have immediately produced moneys equaling US$40,000,000. The Nigerian National Petroleum Company is desirous of oil exploration in other parts of the world, however, because of certain regulations of the Nigerian Government, it is unable to move these funds to another region.

You assistance is requested as a non-Nigerian citizen to assist the Nigerian National Petroleum Company, and also the Central Bank of Nigeria, in moving these funds out of Nigeria. If the funds can be transferred to your name, in your United States account, then you can forward the funds as directed by the Nigerian National Petroleum Company. In exchange for your accommodating services, the Nigerian National Petroleum Company would agree to allow you to retain 10%, or US$4 million of this amount.

However, to be a legitimate transferee of these moneys according to Nigerian law, you must presently be a depositor of at least US$100,000 in a Nigerian bank which is regulated by the Central Bank of Nigeria.

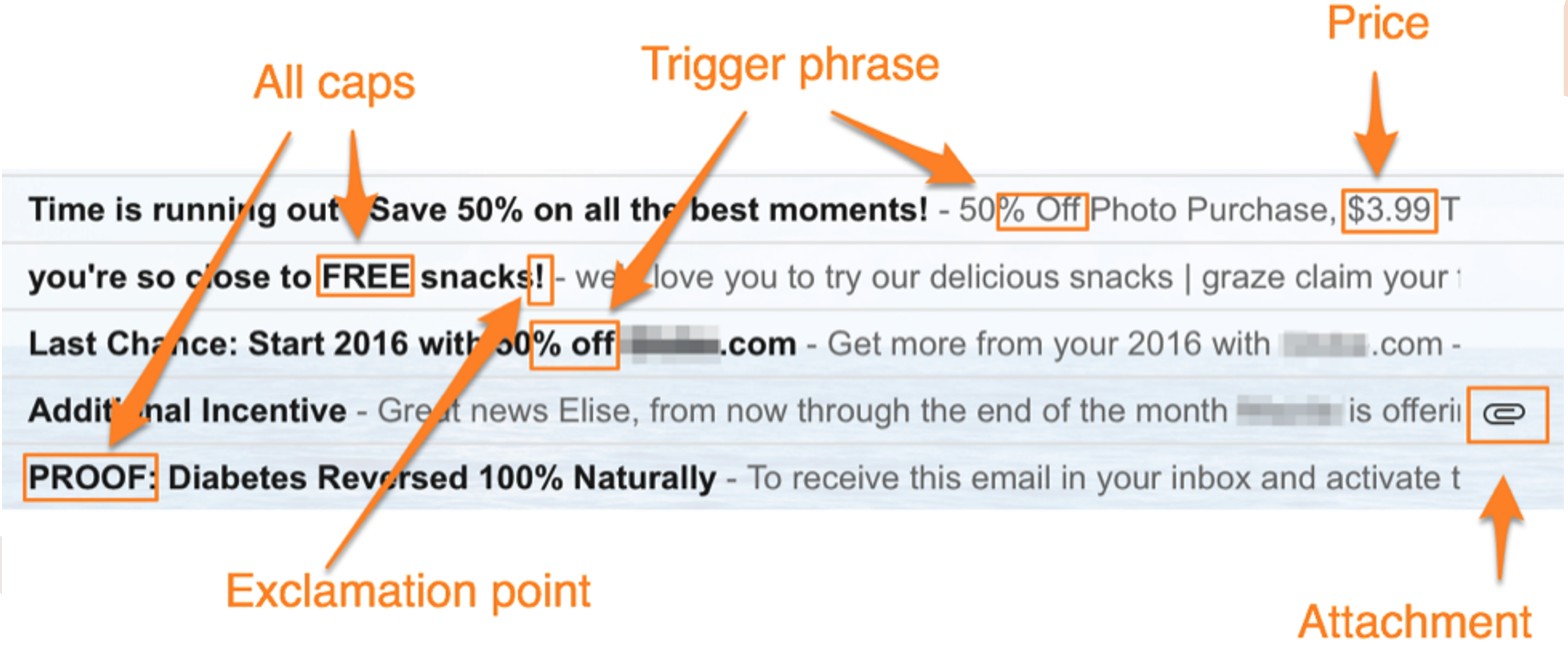
If it will be possible for you to assist us, we would be most grateful. We suggest that you meet with us in person in Lagos, and that during your visit I introduce you to the representatives of the Nigerian National Petroleum Company, as well as with certain officials of the Central Bank of Nigeria.

Please call me at your earliest convenience at 18-467-4975. Time is of the essence in this matter; very quickly the Nigerian Government will realize that the Central Bank is maintaining this amount on deposit, and attempt to levy certain depository taxes on it.

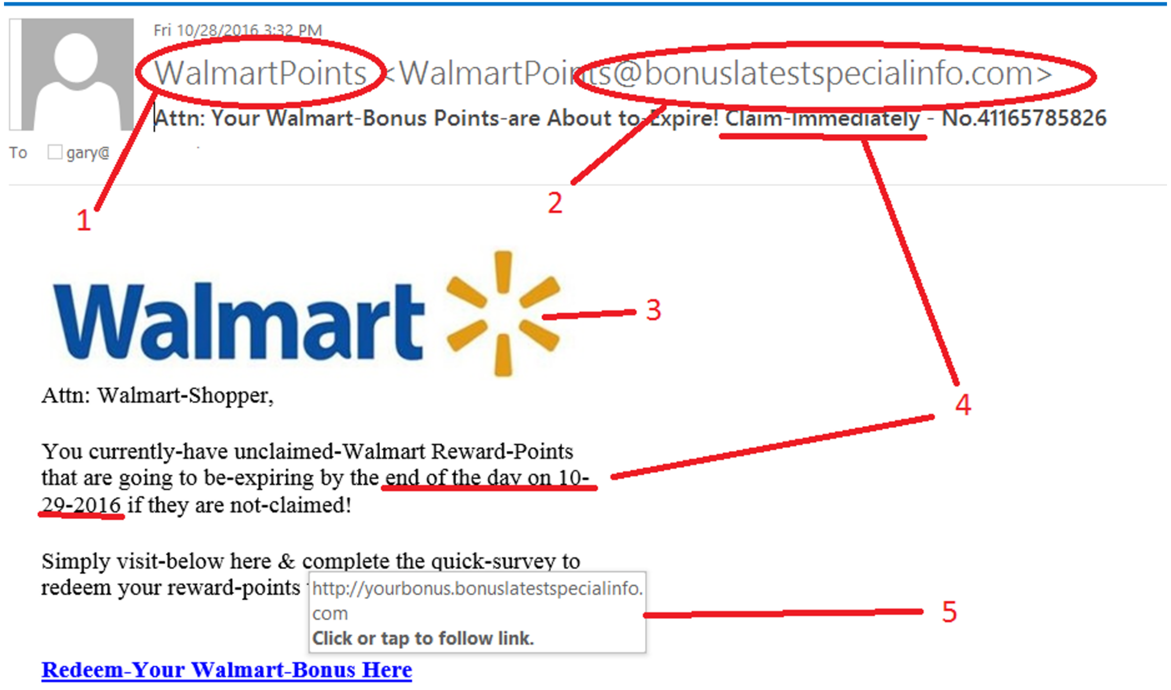
Yours truly,

Prince Alyusi Islassi

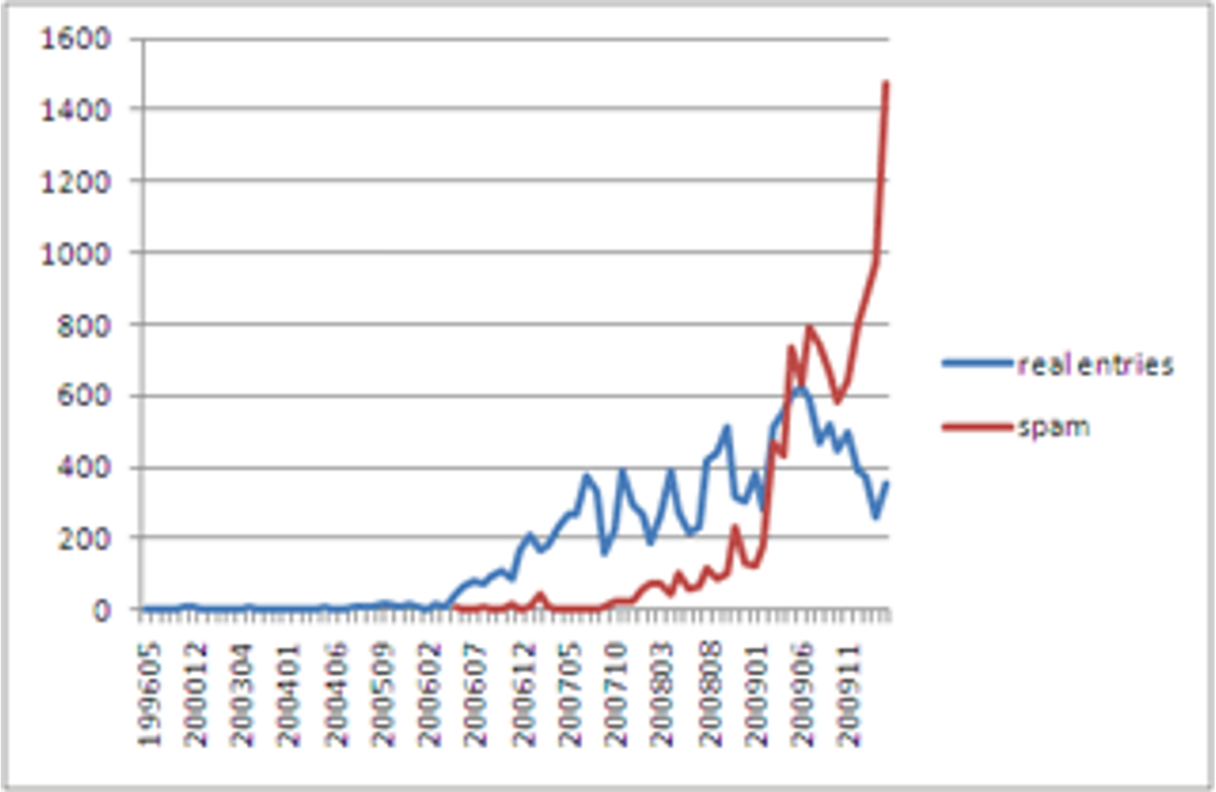
What does spam look like?



phishing


Fri 10/28/2016 3:32 PM
WalmartPoints <WalmartPoints@bonuslatestspecialinfo.com>
Attn: Your Walmart-Bonus Points-are About to-Expire! Claim-Immediately - No.41165785826
To gary@
1
2
Walmart
3
4
Attn: Walmart-Shopper,
You currently-have unclaimed-Walmart Reward-Points that are going to be-expiring by the end of the day on 10-29-2016 if they are not-claimed!
Simply visit-below here & complete the quick-survey to redeem your reward-points
http://yourbonus.bonuslatestspecialinfo.com
Click or tap to follow link.
5
Redeem-Your Walmart-Bonus Here

spam/time



1600
1400
1200
1000
800
600
400
200
0
199605
200012
200304
200401
200406
200509
200602
200607
200612
200705
200710
200803
200808
200901
200906
200911
real entries
spam

how do we stop it?

Filter

- emails have words
- certain words are likely to be spammier than others:
    - Prince
    - Nigeria
    - cash + back + now!
- filter out emails with those words?
    - what if you're from Nigeria?
    - what if you're Prince?
    - what if you want cash back now?

Likely = Bayes Rule

$p(s) = 0.5$

$p(w \mid s) = ???$

$p(w) = ???$

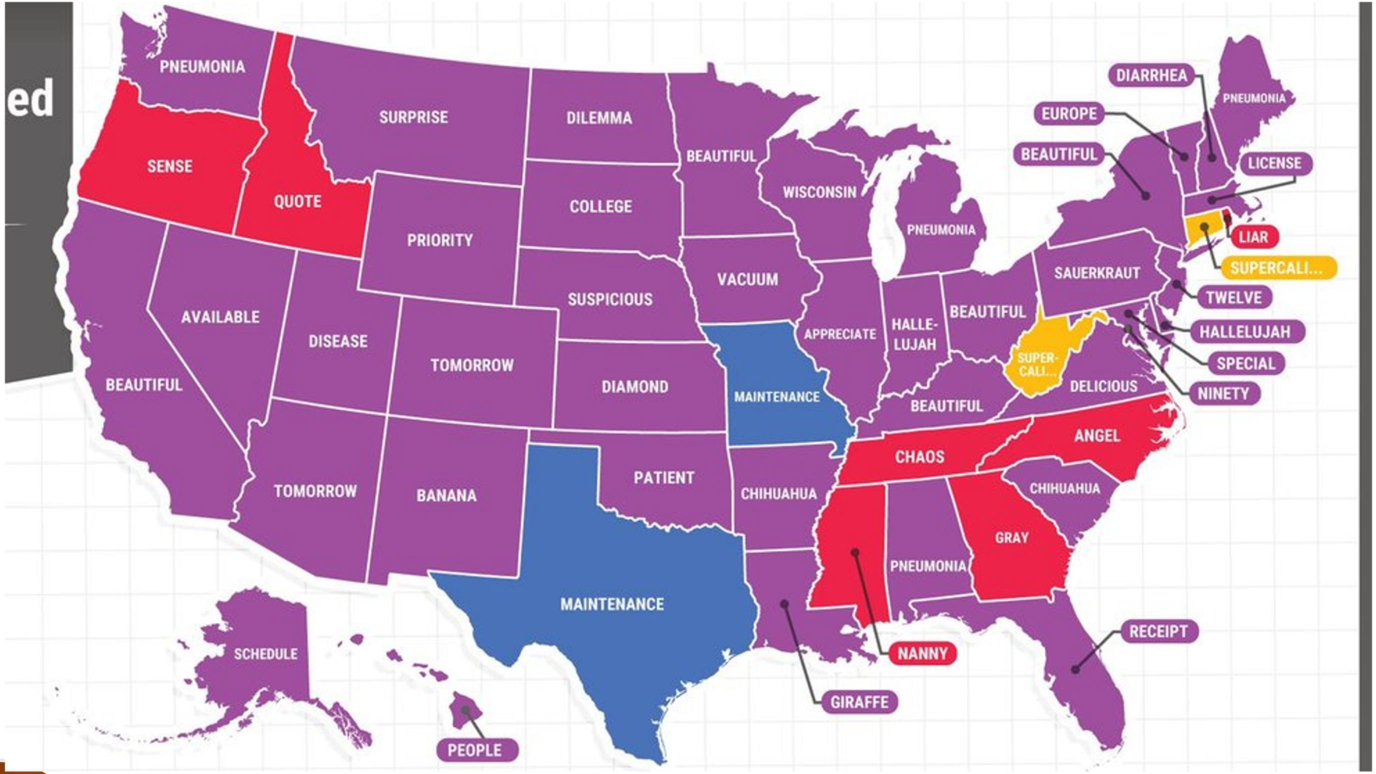
$p(s \mid w) = p(w|s)p(s)/p(w)$

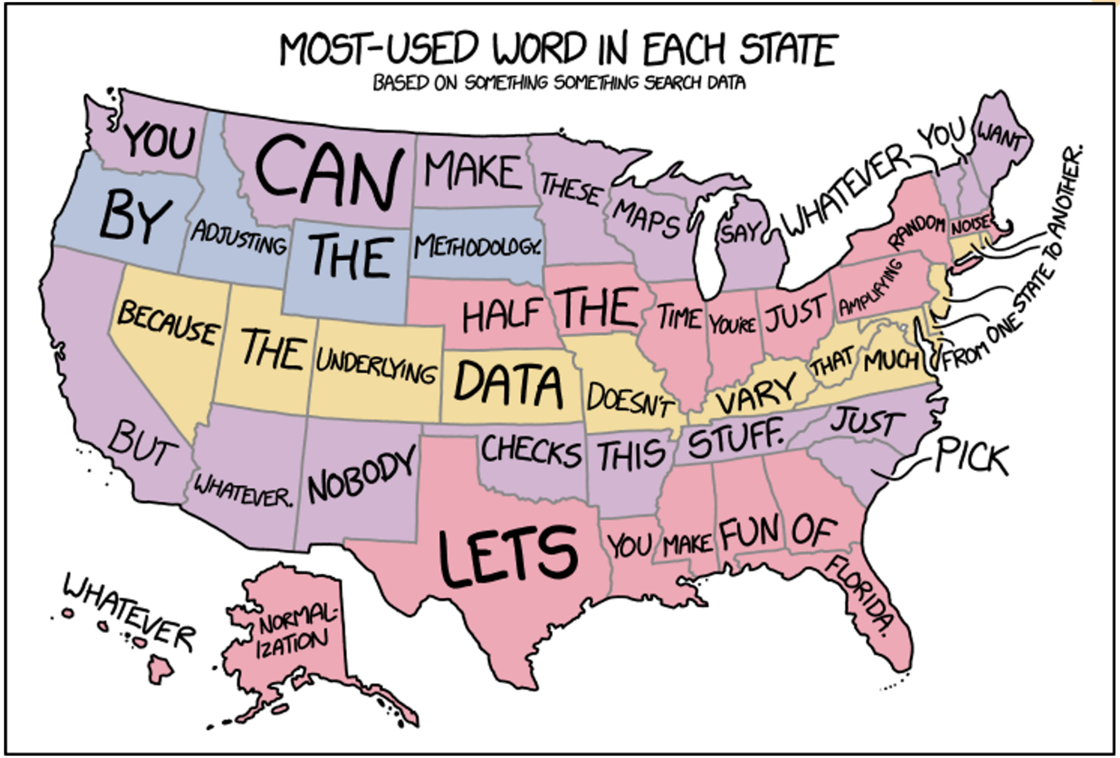
$p(-s \mid w) = p(w|-s)p(-s)/p(w)$

if($p(s|w) > p(-s|w)$) email is spam!


Can Bernie win?
Who is Ammon Bundy?
Is Ted Cruz the Zodiac Killer?
How to move to Canada?
When is NFL Draft?
Is "funner" a word?
When is "American Idol" on?
How to write a cover letter?
Who should I vote for?
Who won Survivor?
How many beers in a keg?
Did Jay-Z cheat on Beyoncé?
Will Trump win?
Is Mercury in retrograde?
Is Mr. T Dead?
When is X-Files?
Is weed bad for you?
How to get away with murder?
Is Joe Flacco elite?
When is Ramadan?
When is Cinco de Mayo?
Why are my nipples so sore?
Is Bernie Sanders vegan?
Who invented pizza?
What is Wyoming?
Why is my poop green?
What is Tinder?
Do penguins have knees?
What does O.P.P. mean?
What is satire?
Is pot legal?
What is quinoa?
Where is Waldo?
How to make meth?
Am I a psycho?
How to make a baby?
Why is Virginia for lovers?
How to open a jar?
What is the clap?
Where is my phone?
What is the minimum wage?
How to be emo?
Is Obama Muslim?
Who won the Civil War?
Who killed Tupac?
Who owns FOX News?
How to smoke salmon?
Where is the internet?
When is Wrestlemania?
Why does everyone hate Florida?
What is the meaning of life?
Powered by Estately
Source: Google Trends (2004-present)

commonly misspelled words





MOST-USED WORD IN EACH STATE
BASED ON SOMETHING SOMETHING SEARCH DATA
YOU
CAN
MAKE
THESE
MAPS
SAY
WHATEVER
YOU
WANT
BY
ADJUSTING
THE
METHODOLOGY.
HALF
THE
TIME
YOU'RE
JUST
AMPLIFYING
RANDOM
NOISE
BECAUSE
THE
UNDERLYING
DATA
DOESN'T
VARY
THAT
MUCH
FROM ONE STATE TO ANOTHER.
BUT
WHATEVER.
NOBODY
CHECKS
THIS
STUFF.
JUST
PICK
WHATEVER
NORMAL-
IZATION
LETS
YOU
MAKE
FUN
OF
FLORIDA.

Likely = Bayes Rule

$p(s) = 0.5$

$p(w \mid s) = ???$

$p(w) = ???$

$p(s \mid w) = p(w|s)p(s)/p(w)$

$p(-s \mid w) = p(w|-s)p(-s)/p(w)$

if($p(s|w) > p(-s|w)$) email is spam!