Reasoning and Decisions Over Time

March 2, 2021

Today:

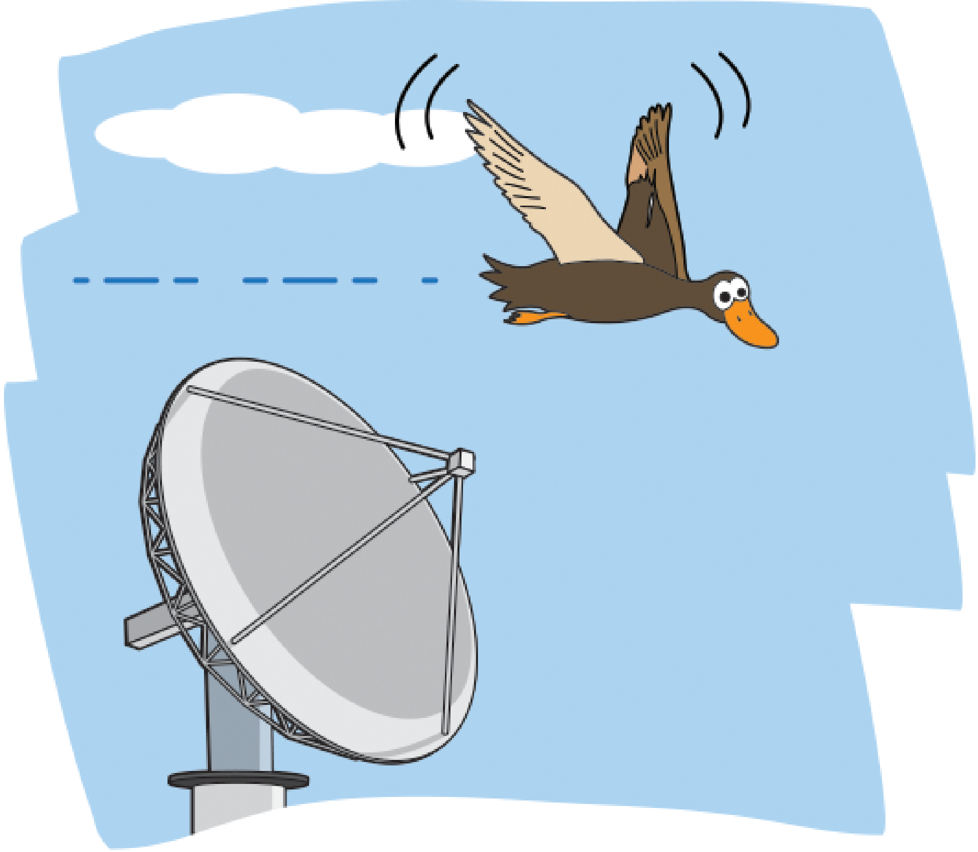
- Markov Chains
- Modeling Evidence
- Hidden Markov Models
- Dynamic Bayesian Networks
- Decision trees and networks
- Utility function

Motivating Examples

- Car diagnosis (static problem)
  - There exists uncertainty
  - We don't care about time
  - Whatever is broken remains broken during diagnosis
- Diabetes management (dynamic problem)
  - It's a dynamic problem with uncertainty
  - Variable values change over time
    - Insulin doses, Blood sugar level, food intake, physical signs etc
  - We must model time to estimate present states and predict future states of a patient

We need time...

- World is not static: transitions occur over time
- Need to introduce time into modeling
- Many useful applications:
  - Projectile/Robot tracking
  - Patient monitoring
  - Speech recognition
  - Statistical modeling: Economy, population, weather, etc



Markov Chains

- Model for sequential data (trials)

$$X_0 \rightarrow X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4 \dashrightarrow$$

- X_i : state variable
- Possible values of X_i : **States**
- Viewing the world as a series of time slices: $X_i \rightarrow X_t$
- Example: Weather

$$X_t = \{rain, sun\}$$

Markov Chains

- **Markov property**

"Future is independent of the past given the present"

$$X_0 \rightarrow X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4 \dashrightarrow$$

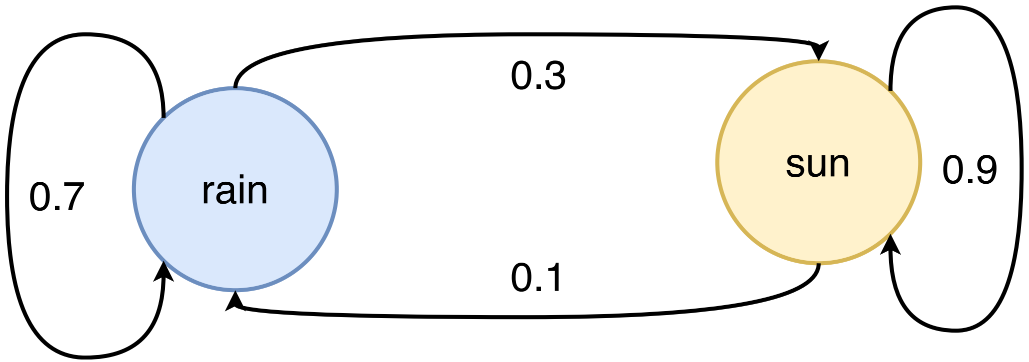
$$P(X_t | X_{0:t-1}) = P(X_t | X_{t-1})$$

- **Stationarity**: Transition probabilities stay the same

$$P(X_t = \bar{s} | X_{t-1} = s) = P(X_{t-1} = \bar{s} | X_{t-2} = s)$$

Markov Chains

- **State transition diagram**: describes all system



$$P(rain|rain) = 0.7$$
$$P(sun|rain) = 0.3$$
$$P(sun|sun) = 0.9$$
$$P(rain|sun) = 0.1$$

Example

Given the weather model and an initial distribution of $P(X_0) = (0.0, 1.0) \equiv (P_{rain}, P_{sun})$, what is the probability distribution after one step and two steps?

$$P(X_1) = \sum_{X_0} P(X_1|X_0)P(X_0) = 0.0 \times (0.7, 0.3) + 1.0 \times (0.1, 0.9) = (0.1, 0.9)$$

$$\begin{aligned} P(X_2) &= \sum_{X_1} P(X_2|X_1)P(X_1) = 0.1 \times (0.7, 0.3) + 0.9 \times (0.1, 0.9) \\ &= (0.07, 0.03) + (0.09, 0.81) = (0.16, 0.84) \end{aligned}$$

If we continue at some "point" (infinity) we will reach a state that won't be changing, this is the stationary state.

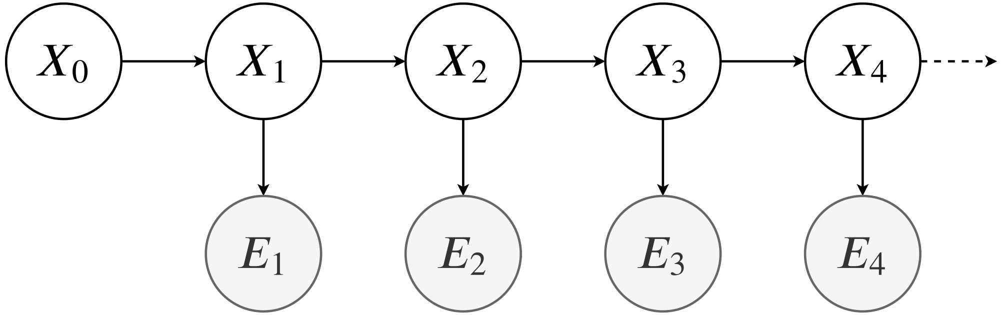
Evidence (observations)

- Markov chains not always applicable An agent may have access to evidence, e.g.
  - sonar signal of a robot
  - patient exam results
  - radar signal for projectile location
- We need a way to incorporate this evidence
- Introduce **evidence variables** E_t
- Example:

$$X_t = \{rain, sun\}$$
$$E_t = \{umbrella, no\ umbrella\}$$

Hidden Markov Models (HMMs)

Extending Markov Chains to introduce evidence variables → **Hidden Markov Models**



Assumptions:

- Evidence variables $\boldsymbol{E_t}$ are stationary
- Evidence variables $\boldsymbol{E_t}$ satisfy the Markov property, i.e. they are conditionally independent of other variables given the current state

Difference between $\boldsymbol{X_t}$ and $\boldsymbol{E_t}$: At time $\boldsymbol{t}$ all evidence variables $\boldsymbol{E_{1:t}}$ are fixed but $\boldsymbol{X_{0:t}}$ not.

HMMs

HMM definition:

$$\begin{aligned} \text{Initial Distribution} \qquad & P(X_0) \\ \text{Transition/Environment Model} \qquad & P(X_t | X_{t-1}) \\ \text{Emmisions/Sensor Model} \qquad & P(E_t | X_t) \end{aligned}$$

Joint distribution:

$$P(X_{0:T}, E_{0:T}) = P(X_0) \prod_{t=1}^{t=T} P(X_t | X_{t-1}) P(E_t | X_t)$$

HMMs

HMM applications/usage:

- **Monitoring/Filtering**: What is the current state belief based on all evidence so far?

$$P(X_t | E_{1:t})$$

- **Prediction**: What is the belief for a future state?

$$P(X_{t+k} | E_{1:t})$$

- **Hindsight/Smoothing**: What is the belief for a past state?

$$P(X_{t-k} | E_{1:t})$$

- **Most Likely Explanation**: What is the belief on the most likely sequence of states?

Reminder

- Bayes' rule

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

- Law of total probability

$$P(A) = \sum_B P(A|B)P(B)$$

Monitoring

$$
\begin{aligned}
P(X_{t+1}|E_{1:t+1}) &= P(X_{t+1}|E_{1:t}, E_{t+1}) \\
&= \alpha P(E_{t+1}|X_{t+1}, E_{1:t}) P(X_{t+1}|E_{1:t}) \quad \text{Bayes' rule} \\
&= \alpha P(E_{t+1}|X_{t+1}) P(X_{t+1}|E_{1:t}) \quad \text{Sensor Markov Assumption} \\
&= \alpha P(E_{t+1}|X_{t+1}) \sum_{X_t} P(X_{t+1}|X_t, E_{1:t}) P(X_t|E_{1:t}) \\
&= \alpha P(E_{t+1}|X_{t+1}) \sum_{X_t} P(X_{t+1}|X_t) P(X_t|E_{1:t}) \text{ Environment Markov Assumption}
\end{aligned}
$$

Recursive

Example

- $\boldsymbol{W}$ = Employee is working/ is lazy
- $\boldsymbol{R}$= Employee has produced results/ no results
- Agent is the boss who is observing the results
- Time $\boldsymbol{t}$ represents a week
- Transition model

$$P(W_t | W_{t-1} = \text{working}) = (0.8, 0.2)$$
$$P(W_t | W_{t-1} = \text{lazy}) = (0.3, 0.7)$$

- Emmissions model

$$P(R_t = \text{results} | W_t) = (0.6, 0.2)$$
$$P(R_t = \text{no results} | W_t) = (0.4, 0.8)$$

Problem

What is the probability that the employee worked in the second week although the boss didn't see any results for two weeks if $P(W_0) = (1, 0)$?

$$P(W_2 | R_1 = \text{no results}, R_2 = \text{no results}) = ?$$

Let's apply what we learned:

First we need to predict one step in the future

$$\begin{aligned} P(W_1) &= \sum_{W_0} P(W_1 | W_0) P(W_0) \\ &= (0.8, 0.2) \times 1 + (0.3, 0.7) \times 0 \\ &= (0.8, 0.2) \end{aligned}$$

Problem

Now we account for the first evidence

$$\begin{aligned} P(W_1|R_1) &= \alpha P(R_1|W_1)P(W_1) \\ &= \alpha(0.4, 0.8)(0.8, 0.2) \\ &= \alpha(0.32, 0.16) \\ &= (0.667, 0.333) \end{aligned}$$

Working employee probability drops!

Let's predict one more step into the future

$$\begin{aligned} P(W_2|R_1) &= \sum_{W_1} P(W_2|W_1)P(W_1|R_1) \\ &= (0.8, 0.2) \times 0.667 + (0.3, 0.7) \times 0.333 \\ &= (0.633, 0.367) \end{aligned}$$

Problem

Now we account for the second evidence

$$\begin{aligned} P(W_2|R_1, R_2) &= \alpha P(R_2|W_2)P(W_2|R_2) \\ &= \alpha(0.4, 0.8)(0.633, 0.367) \\ &= \alpha(0.253, 0.293) \\ &= (0.463, 0.537) \end{aligned}$$

Two consecutive non-productive weeks alert that employee may not be working!

This is called the Forward Algorithm: recursively from past to present

Hindsight

It can be shown that:

$$P(X_k|E_{1:t}) = \alpha P(E_{k+1:t}|X_k)P(X_k|E_{1:k})$$

$$P(E_{k+1:t}|X_k) = \sum_{x_{k+1}} P(E_{k+1}|X_{k+1})P(E_{k+2:t}|X_{k+1})P(X_{k+1}|X_k)$$

Also Recursive

We can compute $P(X_k|E_{1:k})$ with a forward pass from time 0 to k.

$P(E_{k+1:t}|X_k)$ can also be computed recursively goint backwards from time t to k.

Forward–backward algorithm: Combining the two recursive passes

Problem

Now that we have further evidence what is the probability that the employee worked in the first week?

$$P(W_1 | R_1 = \text{no results}, R_2 = \text{no results}) = ?$$

Remember $P(W_1 | R_1) = (0.667, 0.333)$

$$P(R_2 | W_1) = \sum_{W_2} P(R_2 | W_2) P(W_2 | W_1) = 0.4(0.8, 0.3) + 0.8(0.2, 0.7) = (0.48, 0.68)$$

Therefore

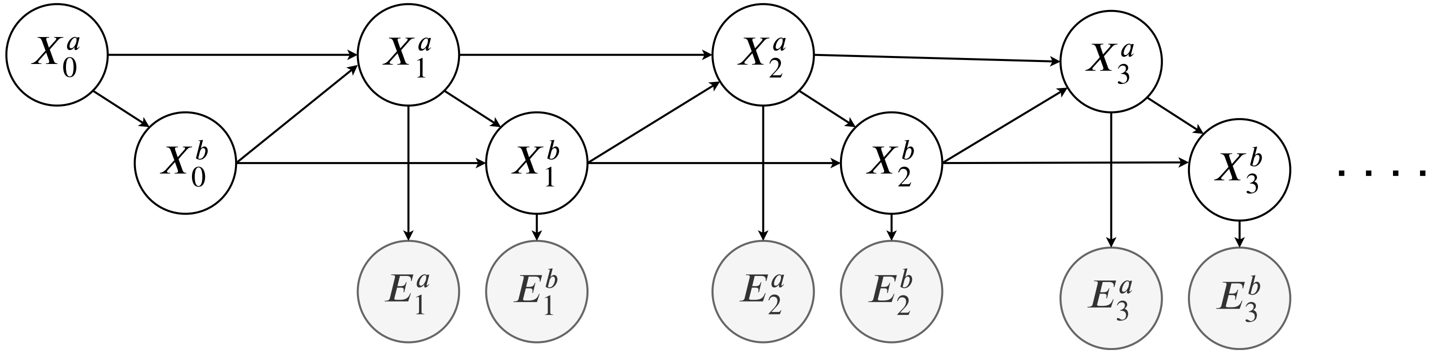
$$\begin{aligned} P(W_1 | R_1, R_2) &= \alpha P(R_2 | W_1) P(W_1 | R_1) \\ &= \alpha(0.48, 0.68)(0.633, 0.367) = \alpha(0.304, 0.249) = (0.549, 0.451) \end{aligned}$$

What's left?

- Most likely explanation: Can also be computed recursively using the Markov property to decompose the most likely path into two.
- All algorithms we encountered are efficient, i.e. run time is linear in the length of the sequence.
- Single state variable in HMMs. We are working at the level of world states.
- Issue: A world can have too many states!
- Example: An office with 20 employees
  - Current approach: W = Who is working/who is lazy
    $$size(world) = |W| = 2^{20} = 1048576$$
  - Alternative approach: W^i = Employee i is working/is lazy
    $$size(world) = |W^1| + |W^2| + \cdots + |W^{20}| = 2 + 2 + \cdots + 2 = 2 \times 20 = 40$$

Dynamic Bayesian Networks (DBNs)

- We need a generalization of HMMs to think about the world at the level of state variables and not states.
- Idea: repeat a Bayesian Network structure over time
- We get **DBNs** which are generalization of HMMs.



Continuous variables

- So far we examined state variables associated with discrete states. What if the state variables can take continuous values, e.g. location of an object, velocity of an object

- Probability distribution over a continuous variable?
  - Probability density function
  - Summations become integrals

- Computations of a general case are messy except in special cases:
  - Transition and sensor models are described by linear Gaussian distributions.
  - Linear distributions of the previous step variables only.
  - Filtering and smoothing integrals have closed form solutions.
  - Solution is known as **Kalman filter**

Continuous variables

- Kalman Filters applications:
  - Applied to navigation for the Apollo Project
  - Radar tracking of aircraft and missiles
  - Acoustic tracking of submarines
  - Visual tracking of vehicles

Project Apollo Archive

Advanced Topics

- Continuous Time
  - Modeling with differential equations
- Non-stationarity
  - Model changes over time
  - Need learning
- Approximate inference for large systems
  - Example: Particle filtering Allowing fixed number of samples; that are weighted by estimated likelihood to be in a state
  - Active research area

Decisions

- So far, we saw how to do inference on a given model, e.g. filtering, prediction, smoothing
- The results of inference are useful for decision making
- This is not enough though
- Decisions should be incorporated in modeling
  - Goal of constructing a model
  - May affect model state

Example: Finance

Suppose your favorite stock NASDIP is down-graded by a reputable analyst and it plummets from \$40 to \$10 per share. You feel this is a good buy, but there is a lot of uncertainty involved. NASDIP's quarterly earnings are about to be released and you think they will be good, which should positively influence its market value. However, you also think there is a good chance the whole market will crash, which will negatively influence NASDIP's market value. In an attempt to quantify your uncertainty, you decide there is a .25 probability the market will crash, in which case you feel NASDIP will go to \$5 by the end of the month. If the market does not crash, you feel by the end of the month NASDIP will be either at \$10 or at \$20 depending on the earnings report. You think it is twice as likely it will be at \$20 as at \$10. So you assign a .5 probability to NASDIP being at \$20 and a .25 probability to it being at \$10 at month end. Your decision now is whether to buy 100 shares of NASDIP for \$1000 or to leave the \$1000 in the bank where it will earn .005 interest in the next month.

The space of possible outcomes

Summary:

Current share value:	\$10
Funds:	\$1000
Probability of \$5 per share:	0.25
Probability of \$10 per share:	0.25
Probability of \$20 per share:	0.5
Bank interest rate:	0.005

One way to decide is to find the expected value of every decision. Value V: Here, the amount of money you have at the end of the month

Expected Values

- Buy NASDIP

$$\begin{aligned} E(V) &= .25(\$500) + .25(\$1000) + .5(\$2000) \\ &= \$1375 \end{aligned}$$

- Leave money in bank

$$\begin{aligned} E(V) &= 1 \times 1.005 \times (\$1000) \\ &= \$1005 \end{aligned}$$

The "rational choice" is to buy NASDIP.

Maximum Expected Utility

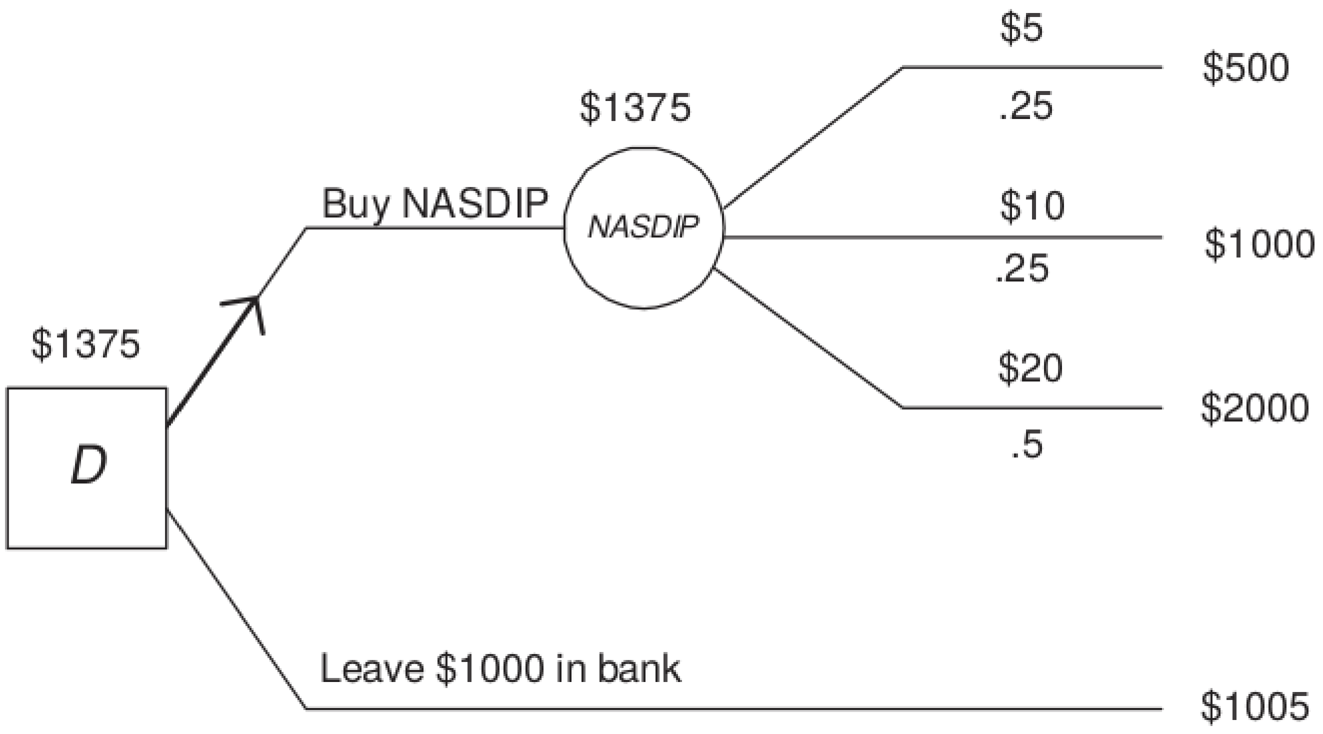
- Why was buying the stocks the "rational choice"?
- Agent's preferences are captured by a utility function $U(s)$, reflecting a desirability of a state, i.e. in our case money outcome
- Expected utility of an action given evidence $E(a|e)$ is given by

$$EU(a|e) = \sum_{s'} P(\mathrm{RESULT(a) = s'}|a, e)U(s')$$

- Principle of maximum expected utility says that a rational agent should choose the action that maximizes the agent's expected utility:

$$action = \mathrm{argmax}_a EU(a|e)$$

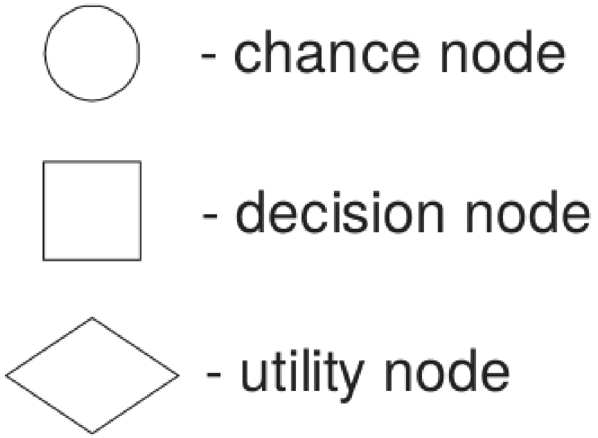
Decision Tree

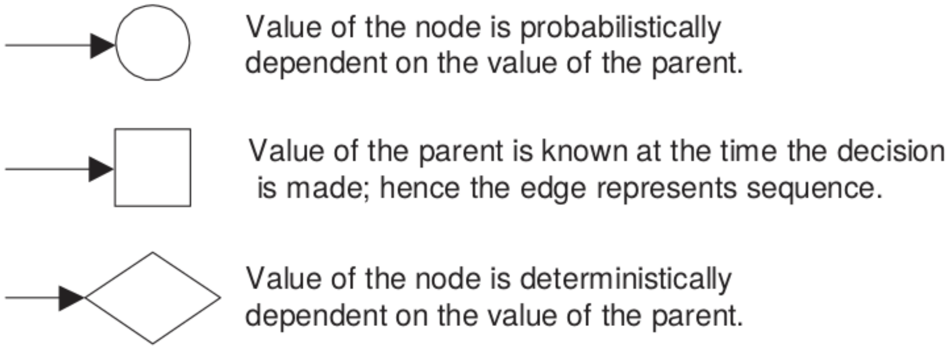


Decision Networks (Influence Diagrams)

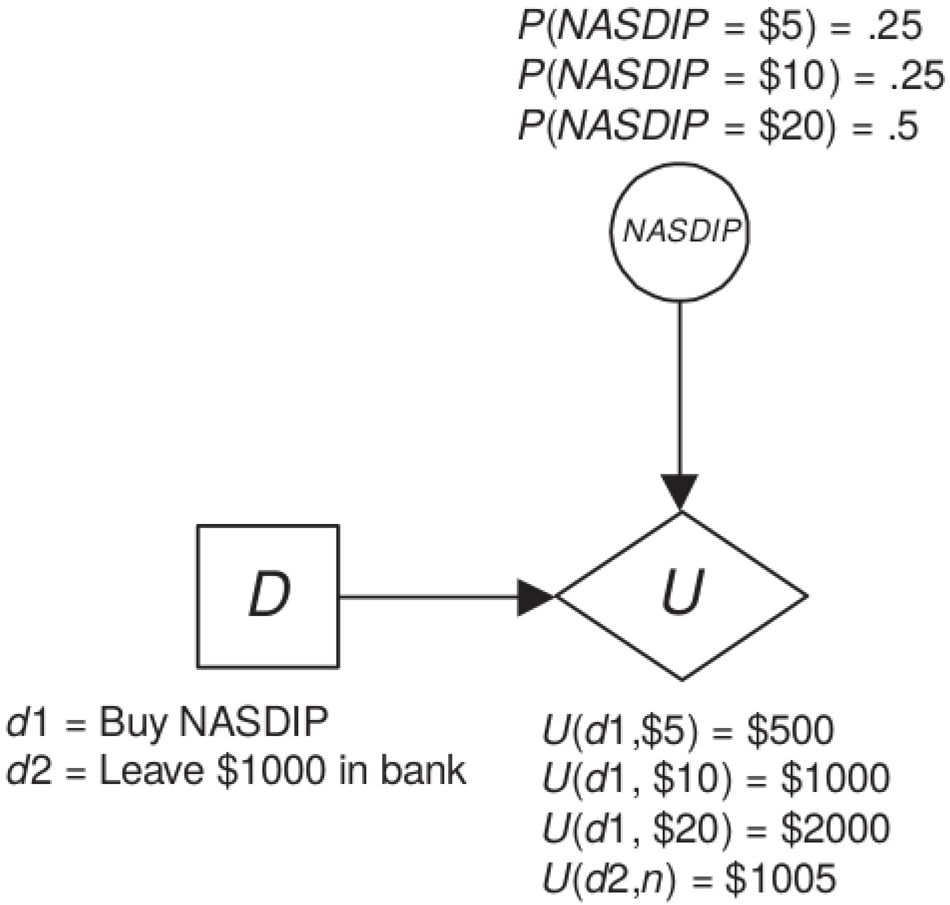
Decision trees grow exponentially with number of nodes; not an efficient representation

Instead we introduce decision networks. They are extension of BNs with extra nodes.





Decision Network



Evaluating Decision Networks or Trees

1. Set the evidence variables for the current state

2. For each possible value configuration of the decision nodes:

    - Set decision nodes to the decision values
    - Calculate the posterior probabilities for the parent nodes of the utility node
    - Calculate the resulting expected utility

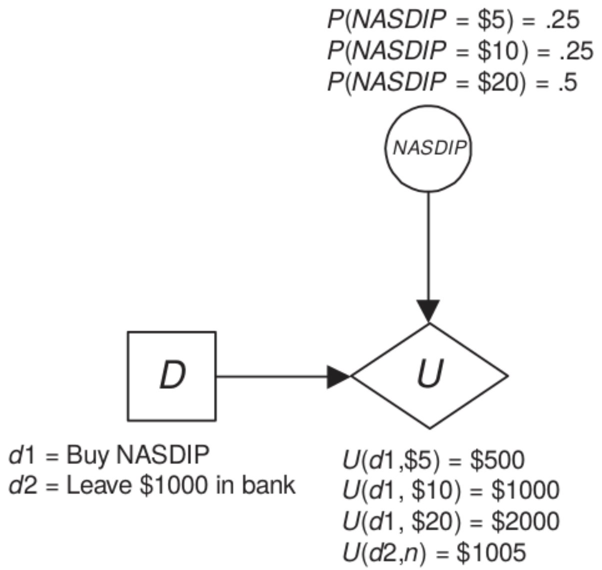
3. Return the decision configuration with highest utility

Evaluating the Example

$$
\begin{aligned}
E(U|d1) = {} & P(\$5|d1)U(d1,\$5) + P(\$10|d1)U(d1,\$10) \\
& + P(\$20|d1)U(d1,\$20) \\
& = (.25)(\$500) + (.25)(\$1000) + (.5)(\$2000) \\
& = \$1375
\end{aligned}
$$

$$
\begin{aligned}
E(U|d2) = {} & P(\$5|d2)U(d2,\$5) + P(\$10|d2)U(d2,\$10) \\
& + P(\$20|d2)U(d2,\$20) \\
& = (.25)(\$1005) + (.25)(\$1005) + (.5)(\$1005) \\
& = \$1005
\end{aligned}
$$

$$EU(D) = \max(E(U|d1), E(U|d2)) = \$1375$$

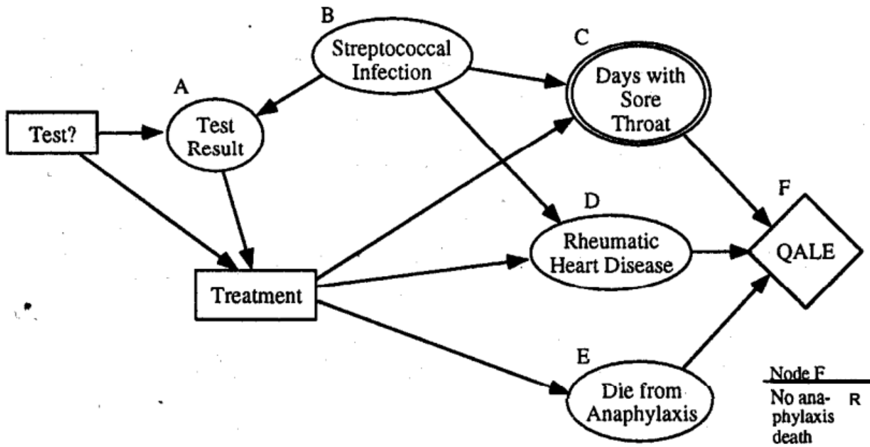


Another example: Sore–throat

W—A		Probability Test+	Test −	NA
Test	Strep present	0.90	0.10	0
	strep absent	0.30	0.70	0
No test	Strep present	0	0	1
	Strep absent	0	0	1

Node B	Probability Strep present	Strep absent
	0.60	0.40

Node C		Days with sore throat
Treat	Strep present	3
	Strep absent	4
Do not treat	Strep present	4
	Strep absent	4



Node D		Probability RHD	No RHD
Treat	Strep present	0.000013	0.999987
	Strep absent	0	0
Do not treat	Strep present	0.000063	0.999937
	Strep absent	0	0

Node E	Probability Die from anaphylaxis	Do not die from anaphylaxis
Treat	0.000003	0.999997
Do not treat	0	1

Node F			QALE
No anaphylaxis death	RHD	3 sore throat days	24.996 yrs
		4 sore throat days	24.995 yrs
	No RHD	3 sore throat days	59.996 yrs
		4 sore throat days	59.99s yrs
Die from anaphylaxis			0 yrs

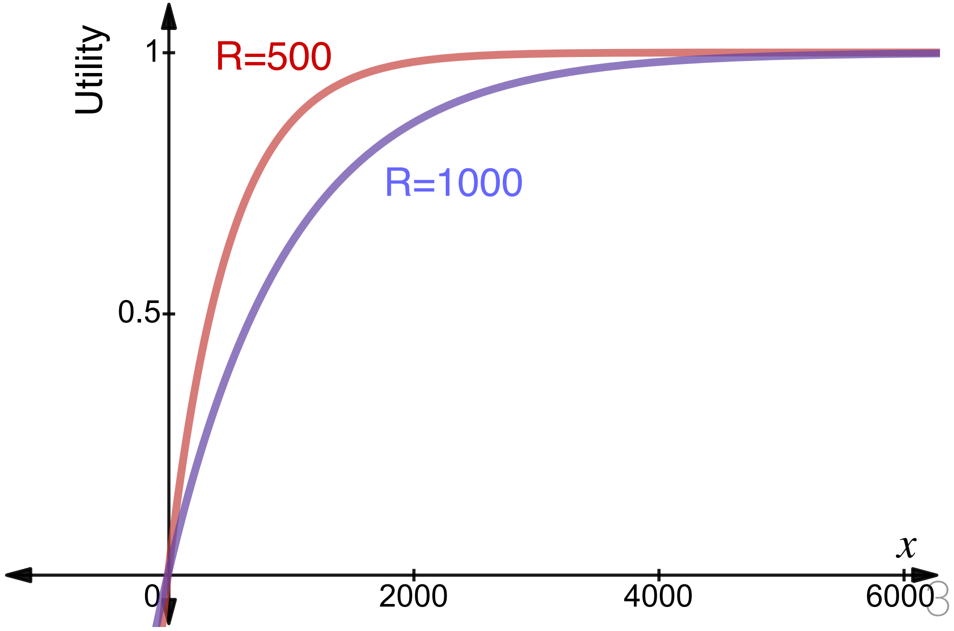
FIGURE 1. Influence diagrams for the sore-throat problem. Tables show the assessments required to specify fully the influence diagram, and are based on a decision analysis by Dippel et al.' QALE = quality-adjusted life expectancy; Strep = streptococcal infection; RHD = rheumatic heart disease; NA = test result not available.

Why did we choose to maximize utility?

- Risk–averse people might prefer the sure \$1005 over the possibility of ending up with only \$500
- There are ways to incorprorate attitude towards risk
- People tend to maximize expected value when amount is small relative to their total wealth
- We could model attitude towards risk using a utility function. With x being money,

$$U_R(x) = 1 - e^{-x/R}$$

  - R is the risk tolerance
  - smaller R, more risk averse
  - higher R, more risk seaking



Using Utility Function

- Case $R = 500$

$$\begin{aligned} EU(\text{Buy NASDIP}) &= 0.25\ U_{500}(\$500) + 0.25\ U_{500}(\$1000) + 0.5\ U_{500}(\$2000) \\ &= 0.25\left(1 - e^{-500/500}\right) + 0.25\left((1 - e^{-1000/500}\right) \\ &\qquad + 0.5\left(1 - e^{-2000/500}\right) \\ &= .86504 \end{aligned}$$

$$EU(\text{Leave \$1000 in bank}) = U_{500}(\$1005) = 1 - e^{-1005/500} = .86601.$$

Individual chooses to leave money in the bank, more risk-avert

Using Utility Function

- Case $R = 1000$

$$EU(\text{Buy NASDIP}) = 0.68873$$

$$EU(\text{Leave \$1000 in bank}) = 0.63396$$

Individual buys stocks, more risk-seeking

References

- Artificial Intelligence: A Modern Approach, Russell and Norvig (Third Edition, Chapters 15–16)
- Pieter Abbeel & Dan Klein, CS 188, UC Berkeley
  https://inst.eecs.berkeley.edu/~cs188/fa18/assets/slides/lec18/FA18_cs188_lecture18_HMMs_6pp.p
- Roland Parr, CPS 170, Duke University
  https://pdfs.semanticscholar.org/presentation/cd2e/49b2c9d3b83445fe6908ce4d580fdfd0f12f.pd