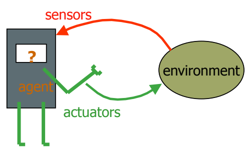
Decision Making Under Uncertainty

AI Class 10 (Ch. 15.1-15.2.1, 16.1-16.3)



Material from Marie desJardin, Lise Getoor, Jean-Claude Latombe, Daphne Koller, and Paula Matuszek

Today's Class

- Making Decisions Under Uncertainty
  - Tracking Uncertainty over Time
  - Decision Making under Uncertainty
  - Decision Theory
  - Utility

Introduction

- The world is not a well-defined place.
- Sources of uncertainty
  - Uncertain **inputs**: What's the temperature?
  - Uncertain (imprecise) **definitions**: Is Trump a good president?
  - Uncertain (unobserved) **states**: What's the top card?
- There is uncertainty in **inferences**
  - If I have a blistery, itchy rash and was gardening all weekend I **probably** have poison ivy

Sources of Uncertainty

- Uncertain **inputs**
  - Missing data
  - Noisy data
- Uncertain **knowledge**
  - >1 cause → >1 effect
  - Incomplete knowledge of causality
  - Probabilistic effects
- Uncertain **outputs**
  - All uncertain:
    - Reasoning-by-default
    - Abduction & induction
    - Incomplete deductive inference
- Result is derived correctly but wrong in real world

Probabilistic reasoning only gives **probabilistic results** (summarizes uncertainty from various sources)

Reasoning Under Uncertainty

- People constantly make decisions anyhow.
  - Very successfully!
  - How?
  - More formally: how do we **reason under uncertainty** with **inexact knowledge**?
- Step one: **understanding what we know**

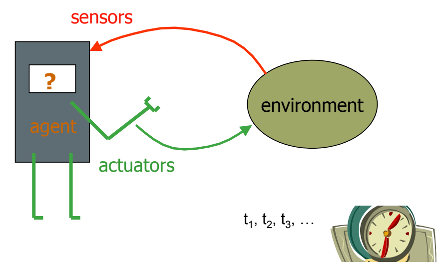
Part I: Modeling Uncertainty Over Time

States and Observations

- Agents don't have a continuous view of world
  - People don't either!
- We see things as a series of snapshots:
- **Observations**, associated with **time slices**
  - $t_1, t_2, t_3, \ldots$
- Each snapshot contains all variables, observed or not
  - $\mathbf{X}_t$ = (unobserved) state variables at time t; observation at t is $\mathbf{E}_t$
- This is **world state at time t**

7

Temporal Probabilistic Agent



8

Uncertainty and Time

- The world changes
  - Examples: diabetes management, traffic monitoring
- Tasks: **track** changes; **predict** changes
- Basic idea:
  - For each time step, copy state and evidence variables
  - Model uncertainty in **change over time** (the Δ)
  - Incorporate new observations as they arrive

9

Uncertainty and Time

- Basic idea:
  - Copy state and evidence variables for each time step
  - Model uncertainty in change over time
  - Incorporate new observations as they arrive
- $\mathbf{X}_t$ = unobserved/unobservable state variables at time t: $BloodSugar_t$, $StomachContents_t$
- $\mathbf{E}_t$ = evidence variables at time t: $MeasuredBloodSugar_t$, $PulseRate_t$, $FoodEaten_t$
- Assuming discrete time steps

10

States (more formally)

- Change is viewed as series of snapshots
  - Time slices/timesteps
  - Each describing the **state** of the world at a particular time
    - So we also refer to these as *states*
- Each time slice/timestep/state is represented as a set of random variables indexed by *t*:
  1. the set of unobservable **state variables** $\mathbf{X}_t$
  2. the set of observable **evidence variables** $\mathbf{E}_t$

11

Observations (more formally)

- Time slice (a set of random variables indexed by *t*):
  1. the set of unobservable **state variables** $\mathbf{X}_t$
  2. the set of observable **evidence variables** $\mathbf{E}_t$
- An **observation** is a set of observed variable instantiations at some timestep
- Observation at time *t*: $\mathbf{E}_t = e_t$
  - (for some values e_t)
- $\mathbf{X}_{a:b}$ denotes the set of variables from $\mathbf{X}_a$ to $\mathbf{X}_b$

12

Transition and Sensor Models

- **So how do we model change over time?**
- Transition model
  - Models how the world changes over time
  - Specifies a probability distribution…
    - Over state variables at time *t*
    - Given values at previous times

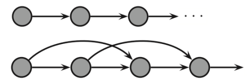
$P(\mathbf{X}_t \mid \mathbf{X}_{0:t-1})$

This can get exponentially large…

- Sensor model
  - Models how evidence (sensor data) gets its values
  - E.g.: $BloodSugar_t \rightarrow MeasuredBloodSugar_t$

Markov Assumption(s)

- **Markov Assumption**:
  - $\mathbf{X}_t$ depends on some finite (usually fixed) number of previous $\mathbf{X}_i$'s
- **First-order Markov process**: $P(\mathbf{X}_t \mid \mathbf{X}_{0:t-1}) = P(\mathbf{X}_t \mid \mathbf{X}_{t-1})$
  - k^{th} order: depends on previous *k* time steps



- **Sensor Markov assumption**: $P(\mathbf{E}_t \mid \mathbf{X}_{0:t}, \mathbf{E}_{0:t-1}) = P(\mathbf{E}_t \mid \mathbf{X}_t)$
  - Agent's *observations* depend *only* on actual current state of the world

Stationary Process

- Infinitely many possible values of *t*
  - Does each timestep need a distribution?
    - That is, do we need a distribution of what the world looks like at t_3, given t_2 AND a distribution for t_{16} given t_{15} AND …
- Assume **stationary process**:
  - Changes in the world state are governed by laws that do not themselves change over time
  - Transition model $P(\mathbf{X}_t \mid \mathbf{X}_{t-1})$ and sensor model $P(\mathbf{E}_t \mid \mathbf{X}_t)$ are time-invariant, i.e., they are the same for all *t*

Complete Joint Distribution

- Given:
  - Transition model: $P(\mathbf{X}_t \mid \mathbf{X}_{t-1})$
  - Sensor model: $P(\mathbf{E}_t \mid \mathbf{X}_t)$
  - Prior probability: $P(\mathbf{X}_0)$
- Then we can specify a **complete joint distribution** of a sequence of states:

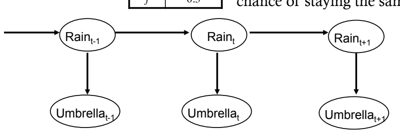
$$P(X_0, X_1, \dots, X_t, E_1, \dots, E_t) = P(X_0) \prod_{i=1}^{t} P(X_i \mid X_{i-1}) P(E_i \mid X_i)$$

- What's the joint probability of instantiations?

Example

R_{t-1}	$P(R_t \mid R_{t-1})$
t	0.7
f	0.3

Weather has a 30% chance of changing and a 70% chance of staying the same.



R_t	$P(U_t \mid R_t)$
t	0.9
f	0.2

Fully worked out HMM for rain: www2.isye.gatech.edu/~yxie77/isye6416_17/Lecture6.pdf

Inference Tasks

- **Filtering** or monitoring: $P(\mathbf{X}_t \mid e_1, \dots, e_t)$:
  - Compute the current belief state, given all evidence to date
- **Prediction**: $P(\mathbf{X}_{t+k} \mid e_1, \dots, e_t)$:
  - Compute the probability of a future state
- **Smoothing**: $P(\mathbf{X}_k \mid e_1, \dots, e_t)$:
  - Compute the probability of a past state (hindsight)
- **Most likely explanation**: $\arg\max_{x_1, \dots x_t} P(x_1, \dots, x_t \mid e_1, \dots, e_t)$
  - Given a sequence of observations, find the sequence of states that is most likely to have generated those observations

Examples

- **Filtering:** What is the probability that it is raining today, given all of the umbrella observations up through today?
- **Prediction:** What is the probability that it will rain the day after tomorrow, given all of the umbrella observations up through today?
- **Smoothing:** What is the probability that it rained yesterday, given all of the umbrella observations through today?
- **Most likely explanation:** If the umbrella appeared the first three days but not on the fourth, what is the most likely weather sequence to produce these umbrella sightings?

Filtering

- Maintain a current state estimate and update it
  - Instead of looking at all observed values in history
  - Also called **state estimation**
- Given result of filtering up to time t, agent must compute result at t+1 from new evidence $\mathbf{e}_{t+1}$:

$$P(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1}) = f(\mathbf{e}_{t+1},\ P(\mathbf{X}_t \mid \mathbf{e}_{1:t}))$$

… for some function f.

Recursive Estimation

1. Project current state forward (t → t+1)
2. Update state using new evidence $\mathbf{e}_{t+1}$

$P(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})$ as function of $\mathbf{e}_{t+1}$ and $P(\mathbf{X}_t \mid \mathbf{e}_{1:t})$:

$P(\mathbf{X}_t{+}1 \mid \mathbf{e}_{1:t+1}) = P(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t}, \mathbf{e}_{t+1})$

Recursive Estimation

- $P(\mathbf{X}_{t+1} \mid \mathbf{e}_{1:t+1})$ as a function of $\mathbf{e}_{t+1}$ and $P(\mathbf{X}_t \mid \mathbf{e}_{1:t})$:

$$P(X_{t+1} \mid e_{1:t+1}) = P(X_{t+1} \mid e_{1:t}, e_{t+1}) \quad \text{dividing up evidence}$$
$$= \alpha P(e_{t+1} \mid X_{t+1}, e_{1:t})\, P(X_{t+1} \mid e_{1:t}) \quad \text{Bayes rule}$$
$$= \alpha P(e_{t+1} \mid X_{t+1})\, P(X_{t+1} \mid e_{1:t}) \quad \text{sensor Markov assumption}$$

- $P(\mathbf{e}_{t+1} \mid \mathbf{X}_{1:t+1})$ updates with new evidence (from sensor)
- One-step prediction by conditioning on current state X:

$$= \alpha P(e_{t+1} \mid X_{t+1}) \sum_{x_t} P(X_{t+1} \mid x_t)\, P(x_t \mid e_{1:t})$$

Recursive Estimation

- One-step prediction by conditioning on current state X:

$$= \alpha P(e_{t+1} \mid X_{t+1}) \sum_{x_t} \underbrace{P(X_{t+1} \mid x_t)}_{\text{transition model}} \underbrace{P(x_t \mid e_{1:t})}_{\text{current state}}$$

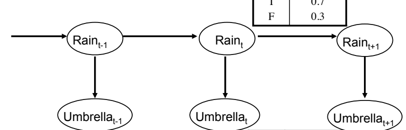
- …which is what we wanted!
- So, think of $P(\mathbf{X}_t \mid \mathbf{e}_{1:t})$ as a "message" $f_{1:t+1}$
  - Carried forward along the time steps
  - Modified at every transition, updated at every new observation
- This leads to a recursive definition:

  $f_{1:t+1} = \alpha\ \text{FORWARD}(f_{1:t}, \mathbf{e}_{t+1})$

Group Exercise: Filtering

$$P(X_{t+1} \mid e_{1:t+1}) = \alpha P(e_{t+1} \mid X_{t+1}) \sum_{X_t} P(X_{t+1} \mid X_t)\, P(X_t \mid e_{1:t})$$

R_{t-1}	$P(R_t \mid R_{t-1})$
T	0.7
F	0.3



R_t	$P(U_t \mid R_t)$
T	0.9
F	0.2

What is the probability of rain on Day 2, given a uniform prior of rain on Day 0, U_1 = true, and U_2 = true?

Part II: Decision Making Under Uncertainty

Decision Making Under Uncertainty

- Many environments have multiple possible outcomes
- Some outcomes may be good; others may be bad
- Some may be very likely; others unlikely
- **What's a poor agent to do?**

Reasoning Under Uncertainty

- How do we **reason** under uncertainty and with inexact knowledge?
  - Heuristics
    - Mimic heuristic knowledge processing methods used by experts
  - Empirical associations
    - Experiential reasoning based on limited observations
  - **Probabilities**
    - Objective (frequency counting)
    - Subjective (human experience)

Decision-Making Tools

- Decision Theory
  - Normative: how *should* agents make decisions?
  - Descriptive: how *do* agents make decisions?
- **Utility** and utility functions
  - Something's **perceived ability to satisfy needs or wants**
  - A mathematical function that ranks alternatives by utility

What is Decision Theory?

- Mathematical study of strategies for optimal decision-making
  - Options involve different risks
  - **Expectations** of gain or loss
- The study of identifying:
  - The values, uncertainties and other issues relevant to a decision
  - The resulting optimal decision for a rational agent

Decision Theory

- Combines **probability** and **utility** → Agent that makes **rational** decisions (takes rational actions)
  - On average, lead to desired outcome
- First-pass simplifications:
  - Want most desirable *immediate* outcome (episodic)
  - Nondeterministic, partially observable world
- Definition of **action**:
- An action a in state s leads to outcome s', RESULT:
  - RESULT(a) is a random variable; domain is possible outcomes
  - $P(\text{RESULT}(a) = s' \mid a, e)$)

Expected Value

- Expected Value
  - The **predicted future value** of a variable, calculated as:
  - The sum of all possible values
  - Each multiplied by the probability of its occurrence

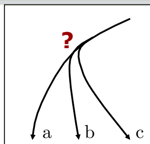
A \$1000 bet for a 20% chance to win \$10,000
[20%(\$10,000) + 80%(\$0)] = \$2000

Satisficing

- Satisficing: achieving a goal **sufficiently**
  - Achieving the goal "more" does not increase utility of resulting state
  - Portmanteau of "satisfy" and "suffice"

Win a baseball game by 1 point now, or 2 points in another inning?
Full credit for a search is ≤3K nodes visited. You're at 2K. Spend an hour making it 1K?
Do you stop the coin flipping game at 1-0, or continue playing, hoping for 2-0?
At the end of semester, you can stop with a B. Do you take the exam?
You're thirsty. Water is good. Is more water better?

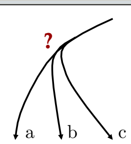
Non-deterministic vs. Probabilistic Uncertainty



{a,b,c}

→ decision that is best for worst case

Non-deterministic model



$\{a(p_a), b(p_b), c(p_c)\}$

→ decision that maximizes expected utility value

Probabilistic model

~ Adversarial search

Value Function

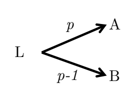
- Provides a ranking of alternatives, but not a meaningful metric scale
- Also known as an "ordinal utility function"
- Sometimes, only relative judgments (value functions) are necessary
- At other times, absolute judgments (utility functions) are required

Rational Agents

- Rationality (an overloaded word).
- A rational agent…
  - Behaves according to a **ranking over possible outcomes**
  - Which is:
    - Complete (covers all situations)
    - Consistent
    - Optimizes over strategies to best serve a desired interest
- Humans are none of these.

Preferences

- An agent chooses among:
  - *Prizes* (A, B, etc.)
  - *Lotteries* (situations with uncertain prizes and probabilities)



- Notation:
  - A > B — A preferred to B
  - A ~ B — Indifference between A and B
  - A >~ B — B not preferred to A

Rational Preferences

- Preferences of a rational agent must obey constraints
  - Transitivity $(A > B) \land (B > C) \Rightarrow (A > C)$
  - Monotonicity $(A > B) \Rightarrow [p > q \Leftrightarrow [p, A; 1 - p, B] > [q, A; 1 - q, B])$
  - Orderability $(A > B) \lor (B > A) \lor (A \sim B)$
  - Substitutability $(A \sim B) \Rightarrow [p,A; 1 - p, C] \sim [p,B; 1 - p,C]$
  - Continuity $(A > B > C \Rightarrow \exists p\ [p,A; 1-p,C] \sim B)$
- Rational preferences give behavior that **maximizes expected utility**
- Violating these constraints leads to irrationality
  - For example: an agent with intransitive preferences can be induced to give away all its money.

Expected Utility

- Goal: find best of expected outcomes
- Random variable X with:
  - n values $x_1,\ldots,x_n$
  - Distribution $(p_1,\ldots,p_n)$
- X is the state reached after doing an action A under uncertainty
  - state = some state of the world at some timestep
- Utility function U(s) is the utility of a state, i.e., **desirability**

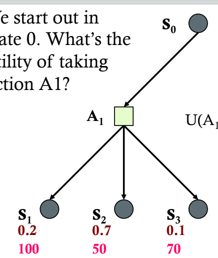
Expected Utility

- X is state reached after doing an action A under uncertainty
- U(s) is the utility of a state ← **desirability**
- EU(a|**e**): The **expected utility** of action A, given evidence, is the *average utility of outcomes* (states in S), weighted by probability an action occurs:

$$EU[A] = S_{i=1,\ldots,n}\, p(x_i|A)U(x_i)$$

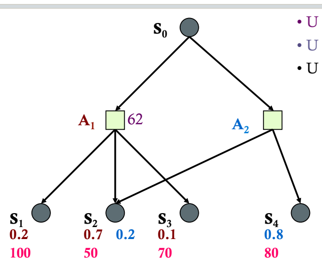
One State/One Action Example

- We start out in state 0. What's the utility of taking action A1?



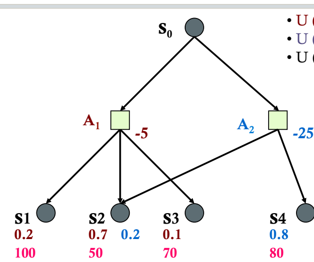
$U(A_1, S_0) = \underline{100 \times 0.2} + \underline{50 \times 0.7} + \underline{70 \times 0.1}$
$= 20 + 35 + 7$
$= 62$

One State/Two Actions Example



- $U(A_1, S_0) = 62$
- $U(A_2, S_0) = 74$
- $U(S_0) = \max_a\{U(a,S_0)\}$
  $= 74$

Introducing Action Costs



- $U(A_1, S_0) = 62 - 5 = 57$
- $U(A_2, S_0) = 74 - 25 = 49$
- $U(S_0) = \max_a\{U(a, S_0)\}$
  $= 57$

MEU Principle

- A **rational agent** should choose the action that maximizes agent's expected utility
- This is the basis of the field of **decision theory**
- The MEU principle provides a **normative criterion** for rational choice of action
- …AI is solved!

Not Quite…

- Must have a **complete** model of:
  - Actions
  - Utilities
  - States
- Even if you have a complete model, decision making is computationally **intractable**
- In fact, a truly rational agent takes into account the utility of reasoning as well (**bounded rationality**)
- Nevertheless, great progress has been made in this area
  - We are able to solve much more complex decision-theoretic problems than ever before

Money

- Money does not behave as a utility function
  - That is, people don't maximize expected value of *dollar assets.*
- People are risk-averse:
  - Given a lottery L with expected monetary value EMV(L), usually U(L) < U(EMV(L))

Want to bet $1000 for a 20% chance to win $10,000?
[20%($10,000)+80%($0)] = $2000 > [100%($1000)]

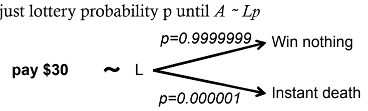
- Expected Utility Hypothesis
  - rational behavior maximizes the expectation of some function *u*… which in need not be monetary

Money Versus Utility

- Money ± Utility
  - More money is better, but not always in a linear relationship to the amount of money
- Expected Monetary Value
- Risk-averse: $U(L) < U(S_{EMV(L)})$
- Risk-seeking: $U(L) > U(S_{EMV(L)})$
- Risk-neutral: $U(L) = U(S_{EMV(L)})$

Maximizing Expected Utility

- Utilities map **states** to **real numbers**. Which numbers?
  - People are very bad at mapping their own preferences
- Standard approach to assessment of human utilities:
  - Compare a state *A* to a standard lottery *Lp* that has
    - "best possible prize" $u_{\top}$ with probability *p*
    - "worst possible catastrophe" $u_{\perp}$ with probability $(1-p)$
  - adjust lottery probability p until $A \sim Lp$



Actual Utility Scales

- Micromorts: one-millionth chance of death
  - Useful for:
  - Russian roulette
  - Paying to reduce product risks, etc.
- QALYs: quality-adjusted life years
  - Useful for:
  - Medical decisions involving substantial risk