Announcements

- W2 is due today (lecture or drop box)
- P2 is out and due on 2/18

Expectimax Search Trees

- What if we don’t know what the result of an action will be? E.g.,
  - In solitaire, next card is unknown
  - In minesweeper, mine locations
  - In pacman, the ghosts act randomly
- Can do expectimax search
  - Chance nodes, like min nodes, except the outcome is uncertain
  - Calculate expected utilities
  - Max nodes as in minimax search
  - Chance nodes take average (expectation) of value of children
- Later, we’ll learn how to formalize the underlying problem as a Markov Decision Process

Maximum Expected Utility

- Why should we average utilities? Why not minimax?
- Principle of maximum expected utility: an agent should choose the action which maximizes its expected utility, given its knowledge
- General principle for decision making
- Often taken as the definition of rationality
- We’ll see this idea over and over in this course!
- Let’s decompress this definition…
  - Probability -- Expectation -- Utility

Reminder: Probabilities

- A random variable represents an event whose outcome is unknown
- A probability distribution is an assignment of weights to outcomes
- Example: traffic on freeway?
  - Random variable: T = amount of traffic
  - Outcomes: T in {none, light, heavy}
  - Distribution: P(T=none) = 0.25, P(T=light) = 0.55, P(T=heavy) = 0.20
- Some laws of probability (more later):
  - Probabilities are always non-negative
  - Probabilities over all possible outcomes sum to one
- As we get more evidence, probabilities may change:
  - P(T=heavy) = 0.20, P(T=heavy | Hour=8am) = 0.60
  - We’ll talk about methods for reasoning and updating probabilities later

What are Probabilities?

- Objectivist / frequentist answer:
  - Averages over repeated experiments
  - E.g. empirically estimating P(rain) from historical observation
  - Assertion about how future experiments will go (in the limit)
  - New evidence changes the reference class
  - Makes one think of inherently random events, like rolling dice
- Subjectivist / Bayesian answer:
  - Degrees of belief about unobserved variables
  - E.g. an agent’s belief that it’s raining, given the temperature
  - E.g. pacman’s belief that the ghost will turn left, given the state
  - Often learn probabilities from past experiences (more later)
  - New evidence updates beliefs (more later)
Uncertainty Everywhere

- Not just for games of chance!
  - I’m sick: will I sneeze this minute?
  - Email contains “FREE!”: is it spam?
  - Tooth hurts: have cavity?
  - 60 min enough to get to the airport?
  - Robot rotated wheel three times, how far did it advance?
  - Safe to cross street? (Look both ways!)

- Sources of uncertainty in random variables:
  - Inherently random process (dice, etc)
  - Insufficient or weak evidence
  - Ignorance of underlying processes
  - Unmodeled variables
  - The world’s just noisy – it doesn’t behave according to plan!

Remindier: Expectations

- We can define function \( f(X) \) of a random variable \( X \)
- The expected value of a function is its average value, weighted by the probability distribution over inputs

- Example: How long to get to the airport?
  - Length of driving time as a function of traffic:
    - \( L(\text{none}) = 20 \)
    - \( L(\text{light}) = 30 \)
    - \( L(\text{heavy}) = 60 \)
- What is my expected driving time?
  - Notation: \( E[L(T)] \)
  - Remember, \( P(T) = \{\text{none: 0.25, light: 0.5, heavy: 0.25}\} \)
  - \( E[L(T)] = L(\text{none}) * P(\text{none}) + L(\text{light}) * P(\text{light}) + L(\text{heavy}) * P(\text{heavy}) \)
  - \( E[L(T)] = (20 * 0.25) + (30 * 0.5) + (60 * 0.25) = 35 \)

Utilities

- Utilities are functions from outcomes (states of the world) to real numbers that describe an agent’s preferences
- Where do utilities come from?
  - In a game, may be simple (+1/-1)
  - Utilities summarize the agent’s goals
  - In general, we hard-wire utilities and let actions emerge (why don’t we let agents decide their own utilities?)
- More on utilities soon…

Expectimax Search

- In expectimax search, we have a probabilistic model of how the opponent (or environment) will behave in any state
  - Model could be a simple uniform distribution (roll a die)
  - Model could be sophisticated and require a great deal of computation
  - We have a node for every outcome out of our control: opponent or environment
  - The model might say that adversarial actions are likely!
  - For now, assume for any state we magically have a distribution to assign probabilities to opponent actions / environment outcomes

Expectimax Pseudocode

```
def value(s):
    if s is a terminal node return evaluation(s)
    if s is a max node return maxValue(s)
    if s is an exp node return expValue(s)
    if s is a terminal node return evaluation(s)

def maxValue(s):
    values = [value(s') for s' in successors(s)]
    return max(values)

def expValue(s):
    values = [value(s') for s' in successors(s)]
    weights = [probability(s, s') for s' in successors(s)]
    return expectation(values, weights)
```
### Expectimax Evaluation

- Evaluation functions quickly return an estimate for a node’s true value (which value, expectimax or minimax?)
- For minimax, evaluation function scale doesn’t matter
  - We just want better states to have higher evaluations (get the ordering right)
  - We call this **insensitivity** to monotonic transformations
- For expectimax, we need **magnitudes** to be meaningful

![Expectimax Diagram](image)

### Mixed Layer Types

- E.g. Backgammon
- Expectiminimax
  - Environment is an extra player that moves after each agent
  - Chance nodes take expectations, otherwise like minimax

\[
\text{Expectiminimax-Value}(\text{state}):
\]
\[
\begin{align*}
\text{if } \text{state} \text{ is a MAX node then} & \quad \text{return the highest } \text{Expectiminimax-Value of Successors}(\text{state}) \\
\text{if } \text{state} \text{ is a MIN node then} & \quad \text{return the lowest } \text{Expectiminimax-Value of Successors}(\text{state}) \\
\text{if } \text{state} \text{ is a chance node then} & \quad \text{return average of } \text{Expectiminimax-Value of Successors}(\text{state})
\end{align*}
\]

### Stochastic Two-Player

- Dice rolls increase b: 21 possible rolls with 2 dice
  - Backgammon – 20 legal moves
  - Depth 4 = 20 x (21 x 20)^4 1.2 x 10^9
- As depth increases, probability of reaching a given node shrinks
  - So value of lookahead is diminished
  - But pruning is less possible...
- TDGammon uses depth-2 search + very good eval function + reinforcement learning: world-champion level play

### Maximum Expected Utility

- Principle of maximum expected utility:
  - A rational agent should choose the action which maximizes its expected utility, given its knowledge

- Questions:
  - Where do utilities come from?
  - How do we know such utilities even exist?
  - Why are we taking expectations of utilities (not, e.g. minimax)?
  - What if our behavior can’t be described by utilities?

### Utilities: Unknown Outcomes

- **Going to airport from home**
- **Take freewway**
- **Take surface streets**
- **Clear, 10 min**
- **Traffic, 50 min**
- **Clear, 20 min**
- **Arrive early**
- **Arrive late**
- **Arrive on time**
Preferences

- An agent chooses among:
  - Prizes: A, B, etc.
  - Lotteries: situations with uncertain prizes

\[ L = [p; A; (1-p); B] \]

- Notation:
  - \( A > B \): A preferred over B
  - \( A \sim B \): indiscernible between A and B
  - \( A \geq B \): B not preferred over A

Rational Preferences

- We want some constraints on preferences before we call them rational

\[ (A > B) \land (B > C) \Rightarrow (A > C) \]

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

  - If \( B > C \), then an agent with \( C \) would pay (say) 1 cent to get \( B \)
  - If \( A > B \), then an agent with \( B \) would pay (say) 1 cent to get \( A \)
  - If \( C > A \), then an agent with \( A \) would pay (say) 1 cent to get \( C \)

Rational Preferences

- Preferences of a rational agent must obey constraints:
  - The axioms of rationality:
    - Orderability
      \[ (A > B) \lor (B > A) \lor (A \sim B) \]
    - Transitivity
      \[ (A > B) \land (B > C) \Rightarrow (A > C) \]
    - Continuity
      \[ A > B > C \Rightarrow \exists p \in [0,1] \text{ s.t. } [p; A; 1-p; C] \sim B \]
    - Substitutability
      \[ A \sim B \Rightarrow [p; A; 1-p; C] \sim [p; B; 1-p; C] \]
    - Monotonicity
      \[ A > B \Rightarrow (p \geq q \Rightarrow [p; A; 1-p; B] \geq [q; A; 1-q; B]) \]

- Theorem: Rational preferences imply behavior describable as maximization of expected utility

Utility Scales

- Normalized utilities: \( u_\text{min} = 1.0 \), \( u_\text{max} = 0.0 \)
- Micromorts: one-millionth chance of death, useful for paying to reduce product risks, etc.
- QALYs: quality-adjusted life years, useful for medical decisions involving substantial risk

- Note: behavior is invariant under positive linear transformation

\[ U'(x) = k_1 U(x) + k_2 \text{ where } k_1 > 0 \]

- With deterministic prizes only (no lottery choices), only ordinal utility can be determined, i.e., total order on prizes

Human Utilities

- Utilities map states to real numbers. Which numbers?

- Standard approach to assessment of human utilities:
  - Compare a state A to a standard lottery \( L_u \) between
    - “best possible prize” \( u_\text{max} \) with probability p
    - “worst possible catastrophe” \( u_\text{min} \) with probability 1-p
  - Adjust lottery probability p until A \( \sim L_u \)
  - Resulting p is a utility in \([0,1]\)

\[ \text{pay$30} \sim \]

\[ \text{continue as before} \]

\[ \text{instant death} \]

\[ 0.99999 \]

\[ 0.00001 \]
Money

- Money does not behave as a utility function, but we can talk about
  the utility of having money (or being in debt)

- Given a lottery $L = [p, \$X; (1-p), \$Y]$
  - The expected monetary value $EMV(L) = p*X + (1-p)*Y$
  - Typically, $U(L) < U(EMV(L))$: why?
  - In this sense, people are risk-averse
  - When deep in debt, we are risk-prone

- Utility curve: for what probability $p$
  am I indifferent between:
  - Some sure outcome $x$
  - A lottery $[p, \$M; (1-p), \$0]$, $M$ large

Example: Insurance

- Consider the lottery $[0.5, \$1000; 0.5, \$0]$
  - What is its expected monetary value? ($\$500$)
  - What is its certainty equivalent?
  - Monetory value acceptable in lieu of lottery
  - $\$400$ for most people
  - Difference of $\$100$ is the insurance premium
    - There’s an insurance industry because people will pay to
      reduce their risk
    - If everyone were risk-neutral, no insurance needed!

Example: Insurance

- Because people ascribe different utilities to different
  amounts of money, insurance agreements can increase
  both parties’ expected utility

  You own a car. Your lottery:
  $L_Y = [0.8, \$0 ; 0.2, -$200]$
  i.e., 20% chance of crashing

  You do not want -$200!
  $U_Y(L_Y) = 0.2*U_Y(-\$200) = -200$
  $U_Y(-\$50) = -150$

  Insurance company buys risk:
  $L_I = [0.8, \$50 ; 0.2, -$150]$
  i.e., $\$50$ revenue + your $L_Y$

  Insurer is risk-neutral:
  $U(L_I) = U(0.8*50 + 0.2*(-150))$
  $= U(\$10) > U(\$0)$

Example: Human Rationality?

- Famous example of Allais (1953)
  - A: $[0.8, \$4k; 0.2, \$0]$
  - B: $[1.0, \$3k; 0.0, \$0]$
  - C: $[0.2, \$4k; 0.8, \$0]$
  - D: $[0.25, \$3k; 0.75, \$0]$

  - Most people prefer B > A, C > D
  - But if $U(\$0) = 0$, then
    - B > A $\Rightarrow$ $U(\$3k) > 0.8 U(\$4k)$
    - C > D $\Rightarrow$ 0.8 $U(\$4k) > U(\$3k)$